Earth Observations for Geohazards

Volume 2

Special Issue Editors
Zhenhong Li
Roberto Tomás
This edition is a reprint of the Special Issue published online in the open access journal Remote Sensing (ISSN 2072-4292) from 2015–2017 (available at: http://www.mdpi.com/journal/remotesensing/special_issues/earth_geohazards).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Author 1; Author 2; Author 3 etc. Article title. Journal Name. Year. Article number/page range.


Articles in this volume are Open Access and distributed under the Creative Commons Attribution license (CC BY), which allows users to download, copy and build upon published articles even for commercial purposes, as long as the author and publisher are properly credited, which ensures maximum dissemination and a wider impact of our publications. The book taken as a whole is © 2017 MDPI, Basel, Switzerland, distributed under the terms and conditions of the Creative Commons by Attribution (CC BY-NC-ND) license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Table of Contents

About the Guest Editors............................................................................................................................ vii

Preface to “Earth Observations for Geohazards”............................................................................................ ix

Roberto Tomás and Zhenhong Li
Earth Observations for Geohazards: Present and Future Challenges
http://www.mdpi.com/2072-4292/9/3/194............................................................................................... 1

Chapter 1: Land Subsidence Hazards

Mi Chen, Roberto Tomás, Zhenhong Li, Mahdi Motagh, Tao Li, Leyin Hu, Huili Gong,
Xiaojuan Li, Jun Yu and Xulong Gong
Imaging Land Subsidence Induced by Groundwater Extraction in Beijing (China) Using Satellite Radar Interferometry
Reprinted from: Remote Sens. 2016, 8(6), 468; doi: 10.3390/rs8060468
http://www.mdpi.com/2072-4292/8/6/468............................................................................................... 13

Peng Liu, Qingquan Li, Zhenhong Li, Trevor Hoey, Guoxiang Liu, Chisheng Wang,
Zhongwen Hu, Zhiwei Zhou and Andrew Singleton
Anatomy of Subsidence in Tianjin from Time Series InSAR
Reprinted from: Remote Sens. 2016, 8(3), 266; doi: 10.3390/rs8030266
http://www.mdpi.com/2072-4292/8/3/266............................................................................................... 34

Yonghong Zhang, Hong’an Wu, Yonghui Kang and Chuanguang Zhu
Ground Subsidence in the Beijing-Tianjin-Hebei Region from 1992 to 2014 Revealed by Multiple SAR Stacks
Reprinted from: Remote Sens. 2016, 8(8), 675; doi: 10.3390/rs8080675
http://www.mdpi.com/2072-4292/8/8/675............................................................................................... 56

Lin Bai, Liming Jiang, Hansheng Wang and Qishi Sun
Spatiotemporal Characterization of Land Subsidence and Uplift (2009–2010) over Wuhan in Central China Revealed by TerraSAR-X InSAR Analysis

Jesús Pacheco-Martínez, Enrique Cabral-Cano, Shimon Wdowinski, Martín Hernández-Marin,
José Ángel Ortiz-Lozano and Mario Eduardo Zermeño-de-León
Application of InSAR and Gravimetry for Land Subsidence Hazard Zoning in Aguascalientes, Mexico
http://www.mdpi.com/2072-4292/7/12/15868 ........................................................................................... 87

Roberta Bonì, Giorgio Pilla and Claudia Meisina
Methodology for Detection and Interpretation of Ground Motion Areas with the A-DInSAR Time Series Analysis
Reprinted from: Remote Sens. 2016, 8(8), 686; doi: 10.3390/rs8080686
http://www.mdpi.com/2072-4292/8/8/686............................................................................................... 103
Cheng-sheng Yang, Qin Zhang, Qiang Xu, Chao-ying Zhao, Jian-bing Peng and Ling-yun Ji
Complex Deformation Monitoring over the Linfen–Yuncheng Basin (China) with Time Series InSAR Technology

Giuseppe Cianflone, Cristiano Tolomei, Carlo Alberto Brunori and Rocco Dominici
InSAR Time Series Analysis of Natural and Anthropogenic Coastal Plain Subsidence: The Case of Sibari (Southern Italy)
http://www.mdpi.com/2072-4292/7/12/15812 .......................................................... 140

Fabiana Caló, Davide Notti, Jorge Pedro Galve, Saygin Abdikan, Tolga Görüm, Antonio Pepe and Füsun Balıktan Şanlı
DInSAR-Based Detection of Land Subsidence and Correlation with Groundwater Depletion in Konya Plain, Turkey
http://www.mdpi.com/2072-4292/9/1/83 ............................................................... 160

Bing Xu, Guangcai Feng, Zhiwei Li, Qijie Wang, Changcheng Wang and Rongan Xie
Coastal Subsidence Monitoring Associated with Land Reclamation Using the Point Target Based SBAS-InSAR Method: A Case Study of Shenzhen, China
Reprinted from: Remote Sens. 2016, 8(8), 652; doi: 10.3390/rs8080652
http://www.mdpi.com/2072-4292/8/8/652 ............................................................... 185

Chao Ma, Xiaojian Cheng, Yali Yang, Xiaoke Zhang, Zengzhang Guo and Youfeng Zou
Investigation on Mining Subsidence Based on Multi-Temporal InSAR and Time-Series Analysis of the Small Baseline Subset—Case Study of Working Faces 22201-1/2 in Bu’ertai Mine, Shendong Coalfield, China
Reprinted from: Remote Sens. 2016, 8(11), 951; doi: 10.3390/rs8110951
http://www.mdpi.com/2072-4292/8/11/951 ............................................................... 207

Jun Hu, Qijie Wang, Zhiwei Li, Rong Zhao and Qian Sun
Investigating the Ground Deformation and Source Model of the Yangbajing Geothermal Field in Tibet, China with the WLS InSAR Technique

Wei Zhou, Shaolin Li, Zhiwei Zhou and Xiaolin Chang
InSAR Observation and Numerical Modeling of the Earth-Dam Displacement of Shuibuya Dam (China)
Reprinted from: Remote Sens. 2016, 8(10), 877; doi: 10.3390/rs8100877
http://www.mdpi.com/2072-4292/8/10/877 ............................................................... 247

Chapter 2: New EO Techniques and Services

Kejie Chen, Natalia Zamora, Andrey Y. Babeyko, Xingxing Li and Maorong Ge
Precise Positioning of BDS, BDS/GPS: Implications for Tsunami Early Warning in South China Sea
http://www.mdpi.com/2072-4292/7/12/15814 .......................................................... 265
Chao Ding, Guangcai Feng, Zhiwei Li, Xinjian Shan, Yanan Du and Huiqiang Wang
Spatio-Temporal Error Sources Analysis and Accuracy Improvement in Landsat 8 Image Ground
Displacement Measurements
Reprinted from: Remote Sens. 2016, 8(11), 937; doi: 10.3390/rs8110937
http://www.mdpi.com/2072-4292/8/11/937 ................................................................. 279

Guoqing Zhou, Tao Yue, Yujun Shi, Rongting Zhang and Jingjin Huang
Second-Order Polynomial Equation-Based Block Adjustment for Orthorectification of
DISP Imagery
Reprinted from: Remote Sens. 2016, 8(8), 680; doi: 10.3390/rs8080680
http://www.mdpi.com/2072-4292/8/8/680 .................................................................................. 301

Martina Cignetti, Andrea Manconi, Michele Manunta, Daniele Giordan, Claudio De Luca,
Paolo Allasia and Francesca Ardizzone
Taking Advantage of the ESA G-POD Service to Study Ground Deformation Processes in High
Mountain Areas: A Valle d’Aosta Case Study, Northern Italy
Reprinted from: Remote Sens. 2016, 8(10), 852; doi: 10.3390/rs8100852
http://www.mdpi.com/2072-4292/8/10/852 .................................................................................. 319

Meizhang He, Qing Zhu, Zhiquiang Du, Han Hu, Yulin Ding and Min Chen
A 3D Shape Descriptor Based on Contour Clusters for Damaged Roof Detection Using Airborne
LiDAR Point Clouds
http://www.mdpi.com/2072-4292/8/3/189 .................................................................................. 344

Anand Vetrivel, Markus Gerke, Norman Kerleand George Vosselman
Identification of Structurally Damaged Areas in Airborne Oblique Images Using a Visual-Bag-of-
Words Approach
http://www.mdpi.com/2072-4292/8/3/231 .................................................................................. 367

Yong Ma, Fu Chen, Jianbo Liu, Yang He, Jianbo Duan and Xinpeng Li
An Automatic Procedure for Early Disaster Change Mapping Based on Optical Remote Sensing

Shuai Xie, Jianbo Duan, Shbin Liu, Qin Dai, Wei Liu, Yong Ma, Rui Guo and Caibong Ma
Crowdsourcing Rapid Assessment of Collapsed Buildings Early after the Earthquake Based on
Aerial Remote Sensing Image: A Case Study of Yushu Earthquake
Reprinted from: Remote Sens. 2016, 8(9), 759; doi: 10.3390/rs8090759
http://www.mdpi.com/2072-4292/8/9/759 .................................................................................. 421

Austin J. Cooner, Yang Shao and James B. Campbell
the 2010 Haiti Earthquake
Reprinted from: Remote Sens. 2016, 8(10), 868; doi: 10.3390/rs8100868
http://www.mdpi.com/2072-4292/8/10/868 .................................................................................. 437

Lixia Gong, Chao Wang, Fan Wu, Jingfa Zhang, Hong Zhang and Qiang Li
Earthquake-Induced Building Damage Detection with Post-Event Sub-Meter VHR TerraSAR-X
Staring Spotlight Imagery
Reprinted from: Remote Sens. 2016, 8(11), 887; doi: 10.3390/rs8110887
http://www.mdpi.com/2072-4292/8/11/887 .................................................................................. 454
Wei Zhai, Huanfeng Shen, Chunlin Huang and Wansheng Pei
Building Earthquake Damage Information Extraction from a Single Post-Earthquake PolSAR Image
http://www.mdpi.com/2072-4292/8/3/171
About the Guest Editors

Zhenhong Li is Professor of Imaging Geodesy in the School of Civil Engineering and Geosciences at Newcastle University. Professor Li has approximately twenty years of research experience in Space Geodesy and Remote Sensing (mainly InSAR and GNSS) and their application to geohazards (e.g., earthquakes, landslides and land subsidence). He specialises in the development of InSAR atmospheric corrections and time-series algorithms for precisely mapping surface movements, and has made several original contributions to the direct estimation and/or mitigation of the effects of atmospheric water vapour on InSAR measurements. His recent major achievements include the generation of the first interferogram from Chinese Gaofen-3 mission together with some collaborators, which is also the first interferogram from Chinese civilian radar missions. He has investigated a series of large earthquakes (e.g., Sumatra (Indonesia, 2007), Wenchuan (China, 2008), Yushu (China, 2011), Van (Turkey, 2011), Tohoku (Japan, 2011), and Gorkha (Nepal, 2015)) and active landslides (e.g., Huangtupo, Shuping, and Daguangbao landslides, China). He is a Fellow of the International Association of Geodesy (IAG), a Senior Member of the Institute of Electrical and Electronics Engineers (IEEE), an Associate Editor for Advances in Space Research and Remote Sensing, and a Guest Editor for three Special Issues of Remote Sensing and one Special Issue of Sensors.

Roberto Tomás is a Civil and Geological engineer. Currently, he is Associate Professor in Geotechnical Engineering at the Department of Civil Engineering of the University of Alicante. He is co-director of the Associate Research Unit IGME-UA of ground movements monitoring using radar interferometry; leader of the Geotechnical and Structural Engineering research group; member of the Geohazard InSAR Laboratory and Modeling Group of the Geological Survey from Spain; member of the Institute of the Water and the Environmental Sciences of the University of Alicante; member of the UNESCO Working Group on Land Subsidence and chair of the Spanish Working Group on land subsidence (SubTer). He has worked on many research projects and has published numerous contributions focused on geohazards (i.e., landslides and land subsidence) monitoring using remote sensing techniques (i.e., LiDAR, InSAR and photogrammetry), modelling, characterization of the effects induced by these phenomena on infrastructures and geotechnical characterization of soils/rocks. For more information, please visit https://personal.ua.es/en/roberto-tomas.
Preface to “Earth Observations for Geohazards”

Earth Observations (EO) encompasses different types of sensors (e.g., SAR, LiDAR, Optical and multispectral) and platforms (e.g., satellites, aircraft, and Unmanned Aerial Vehicles) and enables us to monitor and model geohazards over regions at different scales in which ground observations may not be possible due to physical and/or political constraints. EO can provide high spatial, temporal and spectral resolution, stereo-mapping and all-weather-imaging capabilities, but not by a single satellite at a time. Improved satellite and sensor technologies, increased frequency of satellite measurements, and easier access and interpretation of EO data have all contributed to the increased demand for satellite EO data. EO, combined with complementary terrestrial observations and with physical models, have been widely used to monitor geohazards, revolutionizing our understanding of how the Earth system works.

This book is the second volume of the 2-volume collection of scientific contributions focusing on innovative EO methods and applications for monitoring and modeling geohazards, consisting of two chapters: (1) land subsidence hazards, and (2) New EO techniques and services. The first volume of this book series contains another two chapters: (1) earthquake hazards, and (2) landslide hazards.

Finally, we would like to take this opportunity to thank all authors, editors, reviewers, and supporters for their hard work and dedication that made this Special Issue possible.

Zhenhong Li and Roberto Tomás

Guest Editors
Editorial

Earth Observations for Geohazards: Present and Future Challenges

Roberto Tomás 1 and Zhenhong Li 2,*

1 Departamento de Ingeniería Civil, Escuela Politécnica Superior, Universidad de Alicante, P.O. Box 99, E-03080 Alicante, Spain; roberto.tomas@ua.es
2 COMET, School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne NE1 7RU, UK
* Correspondence: Zhenhong.Li@newcastle.ac.uk; Tel.: +44-191-208-5704

Academic Editor: Prasad S. Thenkabail
Received: 22 February 2017; Accepted: 22 February 2017; Published: 24 February 2017

Abstract: Earth Observations (EO) encompasses different types of sensors (e.g., Synthetic Aperture Radar, Laser Imaging Detection and Ranging, Optical and multispectral) and platforms (e.g., satellites, aircraft, and Unmanned Aerial Vehicles) and enables us to monitor and model geohazards over regions at different scales in which ground observations may not be possible due to physical and/or political constraints. EO can provide high spatial, temporal and spectral resolution, stereo-mapping and all-weather-imaging capabilities, but not by a single satellite at a time. Improved satellite and sensor technologies, increased frequency of satellite measurements, and easier access and interpretation of EO data have all contributed to the increased demand for satellite EO data. EO, combined with complementary terrestrial observations and with physical models, have been widely used to monitor geohazards, revolutionizing our understanding of how the Earth system works. This Special Issue presents a collection of scientific contributions focusing on innovative EO methods and applications for monitoring and modeling geohazards, consisting of four Sections: (1) earthquake hazards; (2) landslide hazards; (3) land subsidence hazards; and (4) new EO techniques and services.

Keywords: earth observation; EO; geohazards; earthquake; landslide; land subsidence; InSAR; LiDAR; optical; images; displacement; deformation; damage assessment; satellite; monitoring

1. Introduction

Geohazards are often defined as the events related to the geological state and processes that pose potential risks to people, properties and/or the environment, which can be classified within two main categories: natural hazards (such as earthquakes, landslides, volcanic eruptions, tsunamis, and floods) and human-induced hazards (such as land subsidence due to groundwater-extraction, water contamination, and atmosphere pollution). Geohazards could cause enormous human and economic losses and disruption, which continue to grow worldwide. In the past decades, the annual cost of natural hazards has increased dramatically [1]. Earthquakes represent one of the most devastating geohazards in terms of human suffering and economic damage, but the major cause of casualties, infrastructural damage, and economic losses, is the secondary hazard of landslides [2]. Volcanic eruptions also represent a significant proportion of geohazards [2], and major eruptions can modulate regional or global atmospheric composition and climate in detrimental ways. Land subsidence due to anthropogenic processes, such as extraction of groundwater, gas, oil, and coal, is another worldwide geohazard that affects wide areas, causing infrastructure damage and increasing flood risk [3,4]. Better decisions require better knowledge to characterize, monitor and model geohazards and then mitigate their impacts on people and the environment. During the past decades, Earth Observation (EO) has been widely applied to disaster risk management (including
Remote Sens. 2017, 9, 194

disaster preparation, response, recovery and mitigation), especially disaster response [5], since it provides extremely useful information for researchers, decision makers and plan makers.

EO is the gathering of information about the Earth using remote sensing technologies, which are often supported by ground surveying techniques. EO has considerably changed our ways of seeing the world, providing a framework to precisely map and monitor large-scale phenomena in a timely way. EO from space and aircraft, combined with complementary terrestrial observations and physical models, have been used to monitor geohazards. An important aspect of space-based (and airborne) EO is that we can investigate areas in which ground observations are not possible due to physical or political constraints. EO techniques can be classified according to sensor types, e.g., passive or active, optical, radar (radio detection and ranging), LiDAR (Laser Imaging Detection and Ranging), or multispectral/hyperspectral. They can be also classified according to the platforms in which the sensors are installed: satellite-based, aircraft-based, Unmanned Aerial Vehicle (UAV) based and ground-based.

This Special Issue contains a collection of articles focusing on the use of EO techniques for the investigation of geo-hazards. We received a total of 79 manuscripts for consideration of publication, which were carefully reviewed by external and independent experts in their respective fields. Forty-three of these manuscripts were accepted for publication. These studies utilize the state-of-the-art EO techniques to map, characterize, monitor and model a range of geohazards, including earthquakes, landslides, land subsidence, and tsunamis. The following sections include a study of present and future trends and challenges on the use of EO and an overview of the 43 contributions in this Special Issue.

2. Use of EO for Geohazards throughout Bibliometric Data

Although EO techniques were firstly used for military and security applications in the Cold War [6], they have become available for a wide range of applications and users during the past few decades. In this section, we illustrate the increasing use of EO techniques throughout a bibliometric study on the occurrence of the term “Earth Observation” in the Web of Science’s bibliographical database similar to previous studies (e.g., [7]). Our searches are restricted to the “Remote sensing” and “Geology” segments. Figure 1 shows the increasing number of yearly publications about the above-mentioned terms. It is clear in Figure 1 that the number of EO publications has started to increase since the 1990s with an apparent exponential-like trend, indicating that EO literature has been growing considerably. Although EO techniques were born last century, it is still a topical issue probably due to its inherent high technological components and wide range of potential applications.

![Figure 1. Evolution of the number of publications per year (dots) in the database Web of Science (WOS) containing the term “Earth observation” in the field of TOPICS between 1968 and 2015. Note that the searches are restricted to (a) “Remote Sensing”; and (b) Geology in the Web of Science. The dashed lines represent the best fit power function of the dots. The dotted lines correspond to the logistic fit performed by [8] for the terms “LiDAR” or “Laser scanning” in the GEOREF database.](image-url)
3. Overview of Contributions

As mentioned earlier, forty-three papers were published in this Special Issue. These studies cover a range of geohazards including earthquakes, landslides, land subsidence, and tsunamis (Figure 2a). Figure 2b shows all the different types of EO sensors used in the published papers such as Synthetic Aperture Radar (SAR), LiDAR, optical, multispectral, GPS and Altimetry, and Figure 2c shows all the different platforms such as satellites, aircraft, Unmanned Aerial Vehicles (UAV) and ground-based platforms. It is clear in Figure 2 that satellite SAR has been widely employed to investigate geohazards due to its all-weather imaging capabilities. It should also be noted that these 43 studies cover different EO usages including geohazard mapping, monitoring and modelling. An overview of all the contributions are presented in the following sections.

![Figure 2](image)

**Figure 2.** Number of papers published in this Special Issue (SI) classified according to: (a) the types of geohazards; (b) the types of sensors employed; and (c) the types of the sensor platforms. Note that (i) the total number in (a) is less than 43 since one article addresses an EO method that could be applied to various geohazards; and (ii) the total numbers in (b,c) are greater than 43 because multiple sensors and/or platforms were utilised in some papers.

3.1. Earthquake Hazards

Earthquakes represent an increasing risk of human loss and severe economic damage as vulnerable populations grow in areas of seismic hazard. Observations of the seismic cycle not only give insight into the mechanics of a fault, but also play key roles in estimating the likelihood of future earthquakes. Interseismic motions build up stress and lead to earthquakes. Zhu et al. [9] employ two tracks of ENVISAT ASAR images to determine the deformation rate maps of the Altyn Tagh Fault (ATF), and then calculate the regional strain rate field using a multi-scale wavelet method. Their results suggest a left-lateral slip rate of $8.0 \pm 0.7$ mm/year and a locking depth of $14.5 \pm 3$ km, which is in agreement with previous GPS and ERS InSAR results.

Postseismic transient deformation is a process that contributes to regional stress evolution, modifying the background tectonic plate motions. Liu et al. [10] present a method to model postseismic deformation time series with the combination model of afterslip and viscoelastic relaxation, and then simultaneously estimate the time-dependent afterslip distribution and the viscosity beneath the earthquake zone. It is reported that (i) the preferred time-dependent afterslip of the 2009 Mw 6.3 Dachaidan, China earthquake mainly occurs in the upper 9.1 km, and increases with time; and (ii) the preferred lower bound of the viscosity beneath the Qaidam Basin’s northern side is $1 \times 10^{19}$ Pa·s, close to that beneath its southern side but different from those in other parts of the Tibetan Plateau, indicating that the viscosity structure beneath the Tibetan Plateau may vary laterally.

The deformation occurring during an earthquake is referred to as coseismic. Observations of coseismic deformation are often used to determine earthquake source parameters, slip distributions, and even the rupture histories. Xu et al. [11] generate continent-wide two-dimensional (2D) (east–west
and vertical) coseismic displacement maps for the 2015 Mw 8.3 Illapel, Chile earthquake using Sentinel-1A TOPS imagery, suggesting that the east–west component (up to 2 m) dominates the 2D surface displacement. Using similar Sentinel-1A data, Solaro et al. [12] also produce coseismic displacement maps for the 2015 Mw 8.3 Illapel, Chile earthquake. Their joint Okada inversion with multiple Sentinel-1A interferograms suggests that most of the slip occurs northwest of the epicenter with a maximum located in the shallowest 20 km; their Finite Element Model indicates that (i) its estimated maximum slip is comparable to the Okada model; and (ii) the von Mises stress distribution agrees with the depth distribution of the aftershock hypocenters. InSAR observations are utilised to investigate the 2003–2004 Bange, China earthquake sequence, involving a series of normal faulting events with Mw > 5.0, indicating that InSAR can provide reliable source parameters of shallow, moderate-sized earthquakes in areas that lack dense seismic networks [13]. Li et al. [14] use Sentinel-1A interferograms to model the 2016 Mw 5.9 Menyuan, China earthquake; they find that the 2016 event has a different focal mechanism from a previous Ms 6.5 earthquake although both are at the two ends of a secondary fault, which is believed to reflect the left-lateral strike-slip characteristics of the Lenglongling fault zone. Multi-platform InSAR observations are employed to determine the coseismic and postseismic slip distributions of the 2011 Mw 7.1 Van, Turkey earthquake, indicating that the upper 7–9 km of the fault, unruptured during the coseismic phase, underwent afterslip in the postseismic phase that may have reduced the seismic potential in its whole length from NW to SE [15].

3.2. Landslide Hazards

Landslides can be triggered by many different mechanisms, such as sudden large earthquakes, constant seismic activity in a tectonically active area, monsoonal rainfall or storms. There are a range of factors that affect landslide motion, including topography, geology, vegetation, precipitation and anthropogenic factors (building roads, deforestation or agricultural terraces). In total, 13 papers on landslides are published in this Special Issue. Mapping is the first topic covered by these papers. Landslide inventory maps document the extent of landslide phenomena in an area and contain relevant information that can be exploited in different ways [16]. Therefore, EO techniques can play an important role in landslide mapping. Al-Rawabdeh et al. [17] present an automated approach to use unmanned aerial vehicles (UAVs) and Semi-Global dense Matching techniques to identify and extract landslide scarps. Watanabe et al. [18] use airborne L-band fully polarimetric SAR to detect landsliding areas induced by Typhoon Wipha on 16 October 2013 on Izu Oshima Island (Japan). Plank et al. [19] propose a fast procedure for mapping landslides based on change detection between pre-event optical imagery and the polarimetric entropy derived from post-event very high resolution (VHR) polarimetric SAR data.

Monitoring landslides is a crucial task to understand the mechanisms, adopt preventive measures and reduce casualties and infrastructure damage [20]. This assignment has been revolutionized by EO techniques. Several contributions are focused on landslide monitoring, proposing and applying novel procedures, using different sensors and platforms and evaluating the quality of the results. Qu et al. [21] develop the hybrid-SAR procedure to combine both amplitude-based and phase-based methods to map and monitor large landslides exhibiting different deformation magnitudes, sliding modes and slope geometries. Using the Slumgullion landslide (southwestern Colorado, USA) as an example, Wang et al. [22] propose a fully polarimetric SAR offset tracking method to improve the precision of landslide movement monitoring. Taking the Shuping landslide (Three Gorges, China) as a case study, Sun and Muller [23] demonstrate the capability of sub-Pixel Offset Tracking techniques to monitor relatively fast slope movements in densely vegetated areas with and without the presence of artificial corner reflectors. Bardi et al. [24] integrate ground-based and satellite InSAR data to study the Åknes rockslide (western coast of Norway). Kropáček et al. [25] monitor the displacements of a large landslide on the western escarpment of the Main Ethiopian Rift (Debre Sina) by means of a multisensor and multitechnique approach. Fernández et al. [26] use Unmanned Aerial Vehicles imagery and high resolution photogrammetry to monitor horizontal and vertical displacements of a landslide affecting...
olive groves in La Guardia de Jaén (Spain). Finally, in Hsieh et al. [27], the errors of digital terrain models (DTMs) derived using different techniques are discussed.

Modelling is also a hot topic in the landslide field. The relationships between triggering factors and the landslide kinematics are a key aspect for the subsequent prediction of future episodes of activities. Jiang et al. [28] propose a sequential data assimilation method (i.e., Ensemble Kalman filter) to couple the surface displacements of the Shuping landslide (Three Gorges, China) derived from the Pixel Offset Tracking technique with hydrological factors. De Novellis et al. [29] develop three-dimensional (3D) Finite Element Models (FEM) of the Ivancich landslide located in Assisi town (Central Italy) through the integration of geological, geotechnical and satellite datasets.

3.3. Land Subsidence Hazards

Land subsidence is an increasing problem worldwide that has strongly attracted the attention of the InSAR community due to the high capability of InSAR techniques for the study of this type of phenomena. This is evidenced by the fact that all the contributions dedicated to this topic in this Special Issue use InSAR techniques for characterizing and monitoring land subsidence.

Beijing and Tianjin are regional economic drivers in Northern China. The aquifers systems in this region have been massively exploited and land subsidence has become evident. Three papers published in this Special Issue are focused on the spatio-temporal distribution pattern and the characterization of land subsidence as well as its triggering and conditioning factors. Zhang et al. [30] study land subsidence in the Beijing-Tianjin-Hebei region from 1992 to 2014 using ERS-1/2, ENVISAT ASAR and RADARSAT-2 images. Similarly, Liu et al. [31] present their results from C-band ENVISAT ASAR and L-band ALOS PALSAR imagery covering the period 2007–2009, implying Line of Sight (LOS) displacements up to 170 mm/yr. Chen et al. [32] employ Small Baseline InSAR technique to process ENVISAT ASAR and TerraSAR-X stripmap images collected from 2003 to 2011 and observe a maximum subsidence in the eastern part of Beijing with a rate greater than 100 mm/year; they also find some relationships between land subsidence and different conditioning and triggering factors (e.g., groundwater levels, soft soil thickness and active faults). This contribution has attracted attention of a wide range of prestigious international media (e.g., The Guardian, The Telegraph, Huffington Post, Forbes and BBC), and is ranked in the top 5% of all research outputs ever tracked by Altmetric, a system that tracks the online attention for a specific piece of research (See: https://mdpi.altmetric.com/details/8441790#score). This contribution is also selected as TOP 10 published articles in 2016 by MDPI (http://blog.mdpi.com/2017/02/20/mdpi-altmetrics-top-10-published-articles-in-2016).

Many coastal areas in the world are experiencing land subsidence due to various factors. The combination of land lowering with rising water levels due to global sea-level rise can make coastal areas especially vulnerable to flooding [33]. Two papers in this Special Issue study land subsidence in coastal areas [34,35]. Cianflore et al. [34] analyze different causes contributing to land subsidence observed in the ancient Greek colony of the Sibari Plain (Southern Italy) using ENVISAT ASAR and Cosmo-SkyMED images. Xu et al. [35] focus on the land subsidence affecting the land reclaimed from the sea in Shenzhen (SE China). These authors use a Point Target based Small Baseline Subset InSAR approach to process ascending and descending ENVISAT ASAR images acquired during the period from 2007 to 2010, and observe subsidence rates up to 25 mm/year.

Yang et al. [36] perform correlation analyses between potential triggering and conditioning factors and land subsidence in the Linfen–Yuncheng basin (China) derived from ENVISAT ASAR data collected in 2009–2010. The authors conclude that the observed land subsidence occurs within the fault-controlled basin and is mainly caused by groundwater withdrawal. Similarly, Bai et al. [37] observe a maximum subsidence rate of up to −67.3 mm/year in Wuhan, China using TerraSAR-X images from 2009 to 2010, which are believed to be mainly caused by anthropogenic activities, natural compaction and karst dissolution. A seasonal pattern of displacements is also noticed near the Yangtze River by these authors. Calò et al. [38] use eight-year ENVISAT ASAR images to investigate land subsidence mainly due to groundwater overextraction in the Konya plain, Turkey, and a joint analysis
with GRACE data suggests that the groundwater depletion is not limited to the study site but affects a wider region in the Anatolian Plateau.

Land subsidence information from InSAR observations can be exploited for different purposes. Pacheco-Martínez et al. [39] present a novel approach to combine InSAR and gravimetric surveys for risk management related to land subsidence and surface ground faulting generation. Boni et al. [40] describe a novel methodology for the exploitation of InSAR data to support geological interpretation in areas affected by land subsidence, uplift and seasonal displacements.

Mining-induced subsidence is another important anthropogenic geohazard. Ma et al. [41] study mining subsidence in Shendong Coalfield (China), suggesting that the extent of subsidence exhibits a progressive increase of 13.09 km² per month during the period from 2012 to 2013.

Surface displacements associated with the geothermal field of Yangbajing (China) are studied and modelled by Hu et al. [42] using ENVISAT ASAR images collected from 2006 to 2010, allowing for the interpretation of the volume changes produced in the geothermal field.

Zhou et al. [43] focus on the combination of InSAR observation and numerical modelling to obtain the physical parameters of the Earth-dam displacements of the Shuibuya Dam (China), and these parameters are then used to predict future behaviors of the dam.

3.4. New EO Techniques and Services

Several papers in this Special Issue are focused on new EO techniques and services for geohazard management. Ding et al. [44] develop new methods to reduce stripe artifacts (SA) and Topographic Shadowing Artifacts (TSA) in surface displacement maps retrieved from Landsat 8 optical images. Their experiments indicate that their algorithms could improve the precision of surface displacement maps (near 15%).

Since parameters of satellite orbits of historical declassified intelligence satellite photography (DISP) imagery are not available and ground control points (GCPs) are lacking, Zhou et al. [45] develop a second order polynomial equation-based block adjustment model for orthorectification that provides accuracy in the level of 2.0 pixels (i.e., approximately 2.0–4.0 m) in the assembling of the imagery for geohazard mapping. Chen et al. [46] illustrate the potential of the BeiDou Navigation Satellite System (BDS) to serve as a fast and reliable early warning system of tsunamis. Cignetti et al. [47] propose an iterative procedure which is applied through the Parallel-SBAS web-tool within the Grid Processing-on-Demand (G-POD) environment to improve SAR data selection and processing by minimizing the temporal decorrelation effects over high mountain regions to obtain mean deformation velocity maps and displacement time series.

Following large geohazards, especially those associated with widespread destruction and high mortality, rapid, accurate and reliable damage assessment is essential to obtain information to guide response activities in the critical post-event hours. Using airborne LiDAR data, He et al. [48] develop a 3D shape descriptor to detect surface- and structure-damaged roofs. Using airborne oblique images, Vetrivel et al. [49] develop a Visual-Bag-of-Words (BoW) based damage classification to detect structure-damaged areas. Ma et al. [50] present an automatic procedure to generate cloudless backdrop and disaster change-detection maps from optical imagery, whilst Xie et al. [51] demonstrate a framework to combine aerial remote sensing imagery with crowdsourcing to support wide-area assessments of building collapses. Using high resolution multispectral and panchromatic remote sensing data, Cooner et al. [52] demonstrate the effectiveness of multilayer feedforward neural networks, radial basis neural networks, and Random Forests in detecting earthquake damage. It is demonstrated in [53] that a combination of a post-event very high resolution SAR image with a pre-event building footprint map can be effectively used to detect damaged buildings, and Zhai et al. [54] demonstrated the feasibility to use a single post-event PolSAR image to assess building damage.

In this section, we analyze the evolution of the annual production of EO publications recorded in the Web of Science’s bibliographical database in the fields of Remote sensing and Geology. Several previous studies [55–57] conclude that a power model (\(q > 0, K > 1\)) can properly explain the growth of publications and new authors in different fields of science. Therefore, to explore the evolution of EO publications, we fit power functions to the available data (Figure 1). It is clear in Figure 1 that the first work in this discipline was published in the middle of the 1970s, but the take-off of this discipline took place in the 1990s. Since then, the number of EO papers has exhibited an exponential-like growth. Furthermore, the fittings show that the growth rate (K) in the number of publications is higher in the remote sensing segment than in the Geology one, because remote sensing includes all the applications of EO not restricted to geology. In order to provide a reference value to be compared with our data, we also plot the best-fit logistic functions of the publications focused on LiDAR or Laser scanning proposed by Derron and Jaboyedoff [8]. Although the data are not directly comparable since their publication data were extracted from the GEOREF database rather than from the Web of Science (WOS), which is usually more restrictive, and the LiDAR discipline is much younger that EO, the initial trends of both EO and LiDAR publications are quite similar.

The high number of manuscripts (i.e., 79) received to be considered for this Special Issue is also a good indicator of the increasing prosperity of the EO community. Most of the papers published in this Special Issue are focused on the post-event exploitation of EO products, i.e., they use EO data to investigate disasters after the event occurs. There is a delay between the event occurrence and the delivery of the useful EO-derived information for decision-makers. Therefore, a great effort is needed in the coming years to reduce the response time after disasters. It should be noted that EO can also be widely applied to disaster preparation, recovery and mitigation. On 1 January 2016, the 17 Sustainable Development Goals (SDGs) of Transforming Our World: the 2030 Agenda for Sustainable Development adopted by world leaders at the 2015 UN Sustainable Development Summit officially came into force, and will run through 2030 and applies to every country. EO and its derived information are specifically demanded to serve the 2030 Agenda by monitoring the 17 SDGs and associated 169 targets, planning and tracking progress, and helping nations and other stakeholders make informed decisions.

5. Conclusions

To conclude, EO has reached some degree of maturity although it is still growing fast following a power trend. The recent success of EO for the investigation of geohazards is mainly due to the development of new EO sensors and techniques, the improved capabilities of EO data acquisition and analysis, and the advance in the production of standardized products to be used by planners and decision makers. It is expected that more and more real-time EO products and services will be emerging in the near future, which will enable us to better manage geohazards.

Acknowledgments: Guest editors would like to take this opportunity to thank all authors, editors, reviewers, and supporters for the hard work and dedication that made this Special Issue possible. Part of this work was supported by the UK Natural Environmental Research Council (NERC) through the Centre for the Observation and Modelling of Earthquakes, Volcanoes and Tectonics (COMET, ref.: come30001) and the LICS and CEDRRIC projects (refs. NE/K010794/1 and NE/N012151/1, respectively), European Space Agency through the ESA-MOST DRAGON-4 projects (ref. 32244) and the Spanish Ministry of Economy and Competitiveness and EU FEDER funds under projects TIN2014-55413-C2-2-P and ESP2013-47780-C2-2-R.

Author Contributions: R.T. and Z.L. are guest editors of this Special Issue, and both wrote and edited the editorial.

Conflicts of Interest: The authors declare no conflict of interest.

References


10. Liu, Y.; Xu, C.; Li, Z.; Wen, Y.; Chen, J.; Li, Z. Time-dependent afterslip of the 2009 mw 6.3 dachaidan earthquake (China) and viscosity beneath the Qaidam basin inferred from postseismic deformation observations. *Remote Sens.* 2016, 8, 784. [CrossRef]


27. Hsieh, Y.-C.; Chan, Y.-C.; Hu, J.-C. Digital elevation model differencing and error estimation from multiple sources: A case study from the meiyuan shan landslide in taiwan. Remote Sens. 2016, 8, 199. [CrossRef]


34. Cianflone, G.; Tolomei, C.; Brunori, C.; Dominici, R. Insar time series analysis of natural and anthropogenic coastal plain subsidence: The case of sibari (southern Italy). Remote Sens. 2015, 7, 15812. [CrossRef]


36. Yang, C.-S.; Zhang, Q.; Xu, Q.; Zhao, C.-Y.; Peng, J.-B.; Ji, L.-Y. Complex deformation monitoring over the Linfen–Yuncheng basin (China) with time series insar technology. Remote Sens. 2016, 8, 284. [CrossRef]


40. Boni, R.; Pilla, G.; Meisina, C. Methodology for detection and interpretation of ground motion areas with the A-dinsar time series analysis. Remote Sens. 2016, 8, 686. [CrossRef]


42. Hu, J.; Wang, Q.; Li, Z.; Zhao, R.; Sun, Q. Investigating the ground deformation and source model of the yangbajing geothermal field in tibet, china with the wls insar technique. Remote Sens. 2016, 8, 191. [CrossRef]


50. Ma, Y.; Chen, F.; Liu, J.; He, Y.; Duan, J.; Li, X. An automatic procedure for early disaster change mapping based on optical remote sensing. Remote Sens. 2016, 8, 272. [CrossRef]
52. Cooner, J.A.; Shao, Y.; Campbell, B.J. Detection of urban damage using remote sensing and machine learning algorithms: Revisiting the 2010 Haiti earthquake. Remote Sens. 2016, 8, 868. [CrossRef]
54. Zhai, W.; Shen, H.; Huang, C.; Pei, W. Building earthquake damage information extraction from a single post-earthquake polsar image. Remote Sens. 2016, 8, 171. [CrossRef]
57. Gupta, B.M.; Sharma, P.; Kumar, S. Growth of world and Indian physics literature. Scientometrics, 1999, 44, 5–16. [CrossRef]

© 2017 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
Chapter 1:
Land Subsidence Hazards
Imaging Land Subsidence Induced by Groundwater Extraction in Beijing (China) Using Satellite Radar Interferometry

Mi Chen 1,2,*, Roberto Tomás 2,3, Zhenhong Li 2, Mahdi Motagh 4, Tao Li 2,5, Leyin Hu 2,6, Huili Gong 1, Xiaojuan Li 1, Jun Yu 7 and Xulong Gong 7

1 College of Resources Environment and Tourism, Capital Normal University, Beijing 10048, China; gonghl@263.net (H.G.); xiaojuanli@vip.sina.com (X.L.)
2 Center for Observation & Modeling of Earthquakes, Volcanoes & Tectonics (COMET), School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne, Tyne and Wear NE1 7RU, UK; roberto.tomas@ua.es (R.T.); Zhenhong.Li@newcastle.ac.uk (Z.L.); TaoLi@whu.edu.cn (T.L.); huleyin@hotmail.com (L.H.)
3 Department of Civil Engineering, Escuela Politécnica Superior, University of Alicante, P.O. Box 99, Alicante 03080, Spain
4 GFZ German Research Centre for Geosciences, Department of Geodesy, Potsdam 14473, Germany; motagh@gfz-potsdam.de
5 GNSS Research Center, Wuhan University, Wuhan 430079, China
6 Earthquake Administration of Beijing Municipality, Beijing 100080, China
7 Key Laboratory of Earth Fissures Geological Disaster, Ministry of Land and Resources (Geological Survey of Jiangsu Province), Nanjing 210018, China; njhzmyj@163.com (J.Y.); xulonggong@126.com (X.G.)

* Correspondence: mi.chen@cnu.edu.cn; Tel.: +86-10-6890-3321

Academic Editors: Zhong Lu, Salvatore Stramondo and Prasad S. Thenkabail
Received: 6 April 2016; Accepted: 23 May 2016; Published: 2 June 2016

Abstract: Beijing is one of the most water-stressed cities in the world. Due to over-exploitation of groundwater, the Beijing region has been suffering from land subsidence since 1935. In this study, the Small Baseline InSAR technique has been employed to process Envisat ASAR images acquired between 2003 and 2010 and TerraSAR-X stripmap images collected from 2010 to 2011 to investigate land subsidence in the Beijing region. The maximum subsidence is seen in the eastern part of Beijing with a rate greater than 100 mm/year. Comparisons between InSAR and GPS derived subsidence rates show an RMS difference of 2.94 mm/year with a mean of 2.41 ± 1.84 mm/year. In addition, a high correlation was observed between InSAR subsidence rate maps derived from two different datasets (i.e., Envisat and TerraSAR-X). These demonstrate once again that InSAR is a powerful tool for monitoring land subsidence. InSAR derived subsidence rate maps have allowed for a comprehensive spatio-temporal analysis to identify the main triggering factors of land subsidence. Some interesting relationships in terms of land subsidence were found with groundwater level, active faults, accumulated soft soil thickness and different aquifer types. Furthermore, a relationship with the distances to pumping wells was also recognized in this work.

Keywords: radar; InSAR; Beijing; land subsidence; triggering factors; conditioning factors

1. Introduction

Located at the northern end of the North China Plain, Beijing is ranked as the 5th most water-stressed city in the world [1], and groundwater is the main water source for industrial, agricultural and household activities. With its rapid urban growth, there has been increasing water demand in Beijing. Previous studies [2,3] reveal that the Beijing region has been suffering from land subsidence due to over-exploitation of groundwater since 1935, and more seriously, the rate and extent...
of land subsidence shows an increasing trend. Land subsidence is a severe geohazard threatening the safety of the public and urban infrastructure. Hence, continuous monitoring of land subsidence is critical for detecting potential hazards and designing compensation strategies.

Interferometric Synthetic Aperture Radar (InSAR) is a powerful tool widely used to map land subsidence over wide regions with high spatial-temporal resolutions. InSAR time series techniques such as Persistent Scatterer Interferometry (PSI) [4,5] and the Small Baseline Subset (SBAS) [6–8] can minimise the limitations of traditional InSAR (e.g., spatial and temporal decorrelation and atmospheric effects) and have been demonstrated to be able to map ground displacements with precision comparable to traditional geodetic techniques such as leveling [9–12]. In this work, the Small Baseline InSAR technique has been employed to investigate land subsidence in the Beijing region and its relationship with different triggering and conditioning factors.

Gong et al. [13] explored the application of PSI method to monitor land subsidence in Beijing with C-band Envisat ASAR data acquired from 2003 to 2006. Their results showed that the spatial pattern of the subsidence is influenced by Quaternary faults, and land subsidence often develops in areas with clay layer thicker than 50 m. Li et al. [14] employed SBAS InSAR to investigate city subsidence in Beijing with two adjacent tracks of C-band Envisat ASAR images, reporting a maximum subsidence rate reaching 100 mm/year in Chaoyang district. Chen et al. [15] studied the spatial and temporal evolution of the groundwater depression funnels in the Beijing plain and the 2D and 3D spatial evolution of land subsidence from a model based on the combination of piezometric, GPS, and radar data. Liang et al. [16] monitored the accumulated crustal deformation and its characteristics in Beijing and its surrounding areas by traditional D-InSAR method with L-band SAR data. They reported that the maximum deformation rate along Line of Sight (LOS) was about 125 mm/year between 2007 and 2010. Ng et al. [17] applied PS InSAR to 41 Envisat ASAR images (acquired from 2003 to 2009) and 24 ALOS PALSAR images (acquired from 2007 to 2009) to investigate the land deformation rate maps over Beijing showing that the vertical displacement rates were in the range of –115 to 6 mm/year. Zhu et al. [18] analyzed the spatial relationship between land subsidence and three factors (groundwater level, compressible soil thickness and building areas) by means of remote sensing and Geographical Information System (GIS) tools and used the Back Propagation Neural Network (BPN) model combined with Genetic Algorithm (GA) to simulate regional distribution of land subsidence. Hu et al. [19] used 52 Envisat ASAR images acquired from 2003 to 2010 to derive land deformation information in Beijing, and found a maximum subsidence velocity of up to 110 mm/year. Zhang et al. [20] utilized stress-strain analysis and oedometer tests to characterize the hydraulic and mechanical properties of the five hydrogeologic units at different depths in the Beijing plain. The results revealed that the second (64.5–82.3 m) and third (102–117 m) aquitards contributed 39% of the total compression deformation, and the second (82.3–102 m) and third (117–148 m) confined aquifers exhibited elasto-plastic mechanical behaviour. Zhu et al. [21] used SAR images, optical images and hydrogeological data to study the land subsidence in the northern Beijing plain. PS InSAR was applied to C-band Envisat ASAR images acquired from 2003 to 2010 to estimate land subsidence on this area. The results indicated that the largest subsidence rates reached 52 mm/year and the silty clay layers contribute to the larger land subsidence.

Compared with medium resolution SAR imagery such as C-band Envisat ASAR and L-band ALOS PALSAR data, recent advanced imaging radar sensors such as the TerraSAR-X SAR system (X-band radar wavelength and a shorter revisiting interval) can provide higher spatial resolution and geometric accuracy imagery, which makes it possible to monitor land subsidence at a finer scale [22,23]. Chen et al. [24] utilized C-band Envisat ASAR data and X-band COSMO-SkyMED images to detect land subsidence rates in Shanghai and validated the results by spirit leveling measurements. The results indicated the location of a significant subsidence funnel and the root mean square error of differences of the point targets between two band SAR images was 3 mm/year. Luo et al. [25] extracted land deformation information from L-band ALOS PALSAR data between 2007 and 2010 and X-band TerraSAR-X data from 2009 to 2010. The obtained results indicated that the subsidence rate was in...
the range of −90 mm/year to −10 mm/year from TerraSAR-X data and −190 to −10 mm/year from PALSAR data, which was in good agreement with leveling and GPS data.

Although existing investigations have tried different ways of monitoring land subsidence in the Beijing region [13–21], the consistency of the land subsidence monitoring results in this densely populated area in the world acquired by multi-wavelength InSAR time series analysis and validation by GPS data is rarely covered. Furthermore, the comprehensive and quantitative spatio-temporal analysis to identify the main conditioning and triggering factors of land subsidence still need to be systematically researched further. This paper aims to use TerraSAR-X (X-band) stripmap images collected from 2010 to 2011 and Envisat ASAR (C-band) images acquired between 2003 and 2010 to provide a new insight into the spatial and temporal distribution characteristics and main conditioning and triggering factors of land subsidence in the Beijing plain. The joint cross analysis of radar-derived deformation data and existing geo-information will help to improve the knowledge of the mechanisms that govern land subsidence to be used for the development of suitable groundwater management policies in this area. The paper is organized as follows: Section 2 describes the geographical and geological setting of the study area. SAR images and InSAR processing used in this work is described in Section 3. Section 4 shows the main InSAR-derived products, validation of the results and the spatial-temporal characteristics of land subsidence over the study area. Section 5 discusses and analyses the effect of different conditioning and triggering factors on land subsidence. The main conclusions are summarized in Section 6.

2. Description of the Beijing Basin

Beijing is the political, cultural and economic centre of China. As one of the most populated cities in the world, its population has reached 20 million [26]. As one of the most water-scarce cities in the world, the major water source for the Beijing municipality is groundwater, accounting for about two thirds of water use [27].

Beijing is located in northern China (Figure 1a), at 39°28′–41°05′ north latitude and 115°25′–117°30′ east longitude, with an area of 16,807 square kilometers, of which 6390 square kilometers are the plain. The northern and north-eastern parts of Beijing are dominated by (Figure 1b) the Jundu Mountains, while the western part of Beijing is framed by the Taihang Mountains. The southeastern part of Beijing is alluvial-pluvial plains crossed by five rivers (Figure 1b): Chaobai, Yongding, Wenyu, Ju and Juma. The terrain of Beijing is high in the west-north area but low in south-east part, sloping from the front of the mountain to the south-east. The elevations range from around 60–80 m.a.s.l. in the front of mountains to around 20–60 m.a.s.l. in the plains.

The study area is characterized by a monsoon-influenced semi-arid and semi-humid continental climate [20]. The average annual temperature is around 11.7 °C. Annual precipitation, mainly concentrated in summer months (from June to September), is around 570 millimetres and constitutes the main source of water resources in Beijing [27].

The thickness of quaternary deposits, whose lithologies change from single gravel to multilayer structures of clay with sand interbeds and whose grain-size turns from coarse to fine, varies from a few meters over the mountain-front area to hundreds of meters in the plain zone [20,21]. Groundwater levels change from deep to shallow and the groundwater varies greatly from phreatic to multilayer confined water throughout the study area, increasing from tens of meters in the mountain-front zone to several hundred meters in the central or southeast plain area [28].

The sediment, forming the principal aquifer system units, predominantly consists of gravel, fine-to-coarse sand, silt and clays (Figure 1c). Figure 1b,c illustrate the gradual transition from fine-to-coarse sand and gravel in the proximal fan to fine sand, silt or clay in the mid and distal-fan areas.
The Quaternary aquifers of Beijing plain can be divided into three aquifer groups [2, 20, 29]: (a) The first aquifer group (late Pleistocene-Holocene) is widespread over the Beijing plain. The depths of the bottom of this group are less than 100 m. In this aquifer group, the phreatic aquifer with single gravel structure located in the top area of the alluvial-diluvial fan and over the middle and lower parts of the alluvial-proluvial fan, distributes phreatic aquifer and the shallow confined water with a multilayer structure; (b) The second aquifer group (middle Pleistocene) is mainly located over the middle and lower segments of the alluvial-diluvial fan, which is the middle and deep confined aquifers with depths of the roof from 80 to 100 m, and depths of the bottom plate around 300 m. This aquifer group presents a multilayer structure, which is mainly made up of medium-coarse sand and small amounts of gravel; (c) The third aquifer group (early Pleistocene) consists on deep confined aquifers with the depth of the top over 300 m. This aquifer group, mainly distributed over the north-east part and the south-east part of the centre of the depression land in Beijing plain, is mainly composed of medium-coarse sand and gravel, with clay aquiclude thicker than 30 m covering the top.

The unconfined aquifer and the first confined aquifer are intensely exploited for irrigation providing about 68% of the total groundwater pumpage of the aquifer system. However, the deeper aquifer units are not significantly exploited [20].
Corresponding to the classification of the aquifers, three compressible layers responsible of land subsidence can be recognized in Beijing [2,29]. The first compressible layer is widely distributed over the Beijing plain. The depth of the bottom of this layer, which is composed of normally consolidated silt, silty clay and clay with plastic or hard plastic states, is less than 100 m. The second compressible layer is widely distributed over the middle and lower parts of the alluvial-proluvial fan in Beijing. The main lithologics of this layer are silt, silty clay and clay. The depth of the bottom over the south-west part of Beijing is less than 150 m, while the depth of the bottom over the east and north parts of Beijing can reach a depth of 280 m. The compressible over-consolidated layers with plastic or hard plastic clays account for 60–80 percent of the overall thickness. The third compressible layer is distributed mainly over the centre of the depression of the Beijing plain. The depth of the roof of this layer is larger than 270 m. This layer is mainly composed of over-consolidated clays with a solid state [28].

3. InSAR Processing

The SAR dataset used in this study consists of 41 images from Envisat-ASAR acquired from 2003 to 2010 and 14 images from TerraSAR-X (TSX) acquired from 2010 to 2011. The Envisat ASAR images were collected in Stripmap mode from descending track with VV polarisation, and the TSX images were obtained in Stripmap mode from ascending track with HH polarisation (Figure 1). The interferograms were generated from single-look complex (SLC) images using DORIS (Delft object oriented radar interferometric software) [30]. Among all the total amount of possible interferograms formed by pairs of images, only 91 interferograms from Envisat ASAR images and 36 interferograms from TerraSAR-X images were selected for further processing. The selection of interferograms was restricted to those interferograms with spatial and temporal baselines smaller than 1070 m and 1500 days for Envisat ASAR images, and temporal baselines smaller than 200 m and 200 days for TerraSAR-X images. The external Shuttle Radar Topography Mission (SRTM) DEM with 90 m resolution was used to remove the topographic component of the interferometric phase and geocode the interferograms.

In this study, land subsidence information was obtained by using a small baseline subset of interferograms with the StaMPS (Stanford Method for Persistent Scatterer/Multi-Temporal InSAR) technique [31]. The processing includes three main steps:

Firstly, a network of multi-master differential interferograms was created by taking into account small spatial and temporal baselines and reduced Doppler centroid frequency differences to minimize the spatial and temporal decorrelation and topographic errors for the group of small baseline interferograms. Figure 2 shows the network of the small baseline interferograms used for the SBAS analysis of the land subsidence in Beijing.

![Figure 2](image-url)
Secondly, after filtering the azimuth and range spectra to reduce decorrelation effects due to geometry and non-overlapping Doppler frequencies, the slowly decorrelating filtered phase (SDFP) pixels whose phase shows slow decorrelation over short time intervals were selected. For computational efficiency, the SDFP candidates were initially selected by setting a threshold for the amplitude difference dispersion, which is the standard deviation of the amplitude difference between the master and slave divided by the mean amplitude. Afterwards, an iteratively conducted phase analysis yielded the phase stability estimation of each candidate to generate the final set of SDFP pixels [31]. The amplitude dispersion index threshold was set at 0.6 to generate the largest set of candidate pixels.

Thirdly, after the iteratively robust process of SDFP selection and correcting for spatially uncorrelated look angle errors, a three-dimensional phase unwrapping was applied on the final sets of SDFP pixels by using a statistical cost flow algorithm [32,33]. The time series of displacement for each SDFP pixel was derived through a least-squares inversion method [34]. This technique estimates displacement time series without prior consideration of a temporal model for the deformation and thus allows the derivation of temporally varying deformation processes. From these values, long-wavelength atmospheric effects and orbit error are estimated from the SDFP pixels using a spatial-temporal filter method.

4. Results

4.1. InSAR-Derived Land Subsidence Maps

Figure 3 shows the result of deformation pattern along line of sight (LOS) for the multi-temporal SBAS processing of the available Envisat ASAR dataset. In this area, a high concentration of SDFP pixels (i.e., an average density 400 SDFP pixels per square kilometre) is recognized mainly due to the dense urbanization of the scene. The Beijing plain area exhibits high displacement rates with an average velocity of up to 110 mm/year in the LOS direction from 2003 to 2010.

![Figure 3. Mean LOS displacement rates derived from Envisat ASAR (2003–2010) in the Beijing basin.](image-url)
Figure 4 illustrates the spatio-temporal evolution of the accumulated LOS surface displacements at approximately one-year intervals derived from Envisat ASAR images. It is clear that land subsidence gradually increases with time, mainly in those areas located at the northeast part of the city in which subsidence rates are higher (Figure 3).

**Figure 4.** Spatio-temporal evolution of InSAR-derived accumulated LOS displacement in Beijing from Envisat-ASAR images. Note that for simplicity, the displacement patterns of only 8 of the 41 acquisition dates are shown (*i.e.*, one per year).
4.2. InSAR Data Validation

To evaluate the quality of the InSAR derived land subsidence rate map from Envisat ASAR, a comparison with \textit{in situ} instrumental data acquired by means of GPS measurement was conducted. Six GPS points that approximately cover the studied period (\textit{i.e.}, from 2007 to 2010) were utilized as checkpoints. The GPS data were projected along the LOS according to [35]. Figure 5 shows the scatter plot of GPS and InSAR solutions over the six GPS points. Note that InSAR-derived displacement rates considered for the comparison of both techniques are here calculated while averaging displacement rate values from all SDFP pixels contained in a 50 m vicinity for the regional comparison of deformation patterns. It can be seen that the maximum discrepancy (MaxD) is 5.34 mm/year, and the minimum one (mD) is 0.14 mm/year. The mean absolute discrepancies (MD) at six checkpoints is 2.41 mm/year, the standard deviation of the discrepancies is $\pm 1.84$ mm/year and the root mean square (RMS) is 2.94 mm/year. The data correlation between GPS data and InSAR measurement is over 0.98. This fact indicates that the overall SBAS results agree well with GPS, suggesting the reliability of InSAR-derived land subsidence rates.

![Figure 5](image)

\textbf{Figure 5.} Comparison of GPS and Envisat InSAR results for points G1–G6. See GPS stations location in Figure 3.

Ten thousand samples of InSAR-derived subsidence rates between multi-wavelength SAR data (\textit{i.e.}, C- and X- bands) were also compared for further validation. Figure 6 shows the correlation and differences between the two data sets used in this work. A high correlation coefficient of 0.92 was observed with an RMS difference of 7.48 mm/year, further indicating the reliability of InSAR retrievals of land subsidence rate maps.

The overall good agreement at least suggests (1) the study region exhibited similar linear motions during these two time intervals; and (2) it is sensible to use InSAR derived subsidence rates (rather than displacement time series) for further analysis (\textit{e.g.}, to identify controlling factors and \textit{etc.}).
4.3. Spatio-Temporal Analysis of Land Subsidence

There are more than 2.5 million high coherent points extracted in the study region with a total area of around 8000 km². The reference point, located in the centre of Beijing, shows a relative stable state without significant subsidence appearance, as shown in Figure 7. The settlement amounts are relatively minor over the downtown, west, south-east and south-west parts of Beijing city with a mean subsidence rate of around 10 mm/year. The fillings covering the West part of Beijing mainly consists of coarse-grain and medium-grain sediments (Figure 1) from proximal fans with low compressibility [20], resulting in small settlements. Nevertheless, the sediments over the east, north and south parts of Beijing mainly consist of high compressibility fine-grain sediments [20] with thick quaternary deposits (Figure 1). The aquifers units of these areas have been intensely over-exploited, generating severe land subsidence.

The main subsidence bowls are distributed over the Chaoyang, Changping, Shunyi and Tongzhou Districts. It can be observed in Figure 7 and Figure 8 that the subsidence bowls located in the east, north-east and north parts of Beijing have been gradually connected, indicating the serious situation in these areas. Consequently, this bowl, which appeared for the first time in 1975 [15], has gradually increased in depth and extension until the present.
Figure 7. (a) Subsidence bowls located in the NE of Beijing; (b) extension of the subsidence bowls. Mapped InSAR results along LOS have been derived from Envisat ASAR images acquired from 2003 to 2010.

The Xianninghou-Shuangqiao subsidence area, located in the eastern suburbs of Chaoyang District, is in the maximum settlement centre from the East bowl. The maximum subsidence rate in this area for the period 2003–2010 is 110 mm/year (Figure 7). The maximum subsidence from the Northeast bowl is located in Jinzhan-Dongba area with a subsidence rate over 80 mm/year (Figure 7).
The Laiguangying subsidence area over Chaoyang District is one of the areas affected by severe land subsidence with displacement rates over 80 mm/year. As mentioned earlier, the subsidence bowls over the east, northeast and north parts of Beijing have gradually merged, indicating the serious subsidence situation in these districts.

Figure 8. LOS displacements difference in Beijing obtained from Envisat ASAR images between the acquisition dates selected in Figure 4.
In the north part of the Haidian-Changping subsidence area, the maximum displacement rate reaches 60 mm/year. The Dongbalizhuang-Dajiaoting subsidence area (including Sanjianfang, Guanzhuang) over the eastern suburbs of Chaoyang District measured settlements exceeding 50 mm/year (with local values up to 80 mm/year) from 2003 to 2010. The subsidence rate in the Shunyi subsidence area reaches 40 mm/year.

Figure 7 and Figure 8 clearly indicate that the bowls of subsidence in Beijing present a growth trend. The mean subsidence rate around the centres of subsidence in Chaoyang district, located in the eastern part of the city, reaches 80 mm/year, and a maximum subsidence rate higher than 100 mm/year.

Figure 8 illustrates the temporal evolution of the accumulated displacements with an approximately one-year interval, showing that the magnitudes of the annual accumulated displacement vary from one year to another. However, the subsidence rate clearly decreases in 2008. For example, comparing the average accumulated subsidence for a group of SDFP pixels located within the main deformation area (39.898°N, 116.572°E) in different years (December 2007–September 2008 with annual displacement of −99.86 mm and September 2008–October 2009 with annual displacement of −119.77 mm), a decrease of the annual accumulated subsidence of 16.6% is observed.

5. Discussion

In this section, different InSAR products including mean displacement rates and displacement time-series are used for the analysis of InSAR-derived surface displacement in the Beijing plain and compared with different triggering and conditioning factors.

5.1. Relationships between Land Subsidence and Groundwater Level

The over-exploitation of groundwater in the study region has caused a regional drop of piezometric levels, decreasing pore pressure and increasing effective stresses according to Terzaghi’s [36] theory. Consequently, when groundwater is pumped from the aquifer system, both the aquifers and the aquitards that constitute the aquifer system undergo deformation, but to different degrees, as most of the permanent subsidence occurs due to the irreversible compression or consolidation of aquitards [37]. Thus, the deformation of an aquifer-system can be inelastic (the aquifer-system undergoes permanent rearrangement resulting in irreversible compaction and only a part of the total deformation is recovered when water level rises) or elastic (when water level rises, the skeleton of the aquifer-system expands, recovering from the previous occurred deformation) depending on whether or not the applied stresses are beyond its previous maximum level (preconsolidation stress threshold) [37].

When land subsidence data are used in conjunction with good well logs and water-level data from adjacent observation wells, the deformation history can provide the basis for stress-strain analysis [38] and inverse modeling that defines the average compressibility and vertical hydraulic conductivity of the aquitards [39]. In the study region, hundreds of observation wells are used to monitor the evolution of groundwater levels. Figure 9 shows the relationship between piezometric level and land subsidence for some piezometers from the study region. It is worth noting that the displacements measured by InSAR on the ground surface are the addition of the deformations occurred in the different aquitard and aquifer layers underlying the ground surface and, thus, the stress-strain curves shown in Figure 9 correspond to the whole aquifer system. The analysis of the different stress-strain relationship clearly indicate that the aquifer system presents different behaviour (Figure 9).

If we analyse the general trend of water level and InSAR time series from P1 (see location in Figure 3) we can realise that from June 2003 to February 2007 the water level fell from 36.1 to 26.87 m.a.s.l. (i.e., −9.23 m) and the accumulated subsidence was up to 38.2 mm. However, after this period, the water level recovered up to 33.5 m.a.s.l. in May 2008 (i.e., it increased +6.6 m, near the original water level) and the ground surface also rose 43.0 mm (the previous deformation was recovered). So, according to the general trend of both time series (water level and displacements), we
can see that they both follow a similar trend, i.e., they mainly exhibit an elastic behaviour as when the support provided by the fluid is removed or partially removed (i.e., when water level falls), the support provided by the pore pressure is gradually transferred to the soil skeleton, causing subsidence. Conversely, when the pore pressure recovers (i.e., increase due to water level recovery), the support provided by the skeleton is gradually transferred to the pore pressure, which in turn leads to soil expansion (i.e., uplift) [37].

Figure 9. Time series of piezometric level and land subsidence measured by means of InSAR (left); and strain-stress curves, for different areas of Beijing (right). The location of the observation wells is shown in Figure 3. Note that the wells are ordered from north (top) to south (bottom). A1 to A4 refers to different aquifers units described in Section 2.
On the other hand, observation wells P4, P10, P7, P15 exhibit a predominantly inelastic behaviour as the aquifer system presents a continuous subsidence associated with the gradual water level decrease (Figure 9). As the fine sediment layers of the study area are normally consolidated \[21\] \(i.e.,\) the present effective overburden pressure is the maximum pressure that the soil was subjected to in the past, then the geomechanical behaviour of the aquifer-system is characterized by the inelastic storage coefficient \((Skv)\). \(Skv\) is a dimensionless parameter which is proportional to the deformability of the aquifer system \(i.e.,\) the largest the \(Skv\) value, the higher the settlement induced by a water level change. This parameter has been calculated as the inverse slope of the stress-displacement data trend line, according to the graphical methodology proposed by Riley \[38\], for the observational wells P4 to P18 shown in Figure 9, providing values varying from \(4.2 \times 10^{-3}\) in P4 to \(5.6 \times 10^{-2}\) in P18 and confirming that the deformability of the aquifer apparently increases from the north towards the south of the studied area.

5.2. Relationships between Land Subsidence and Geological Structures

The distribution and development trend of the land subsidence over the study region are controlled by geological structures, shown in Figure 10a. Tectonic movements cause tectonic units to decline entirely and slowly under the regional stress field and, complementarily, the relative displacement of the hanging walls and the footwalls induces differences in the land subsidence patterns on both sides of the fault \[40\]. There are several active high angle normal faults in the Beijing plain. Huangzhuang-Gaoliying, Shunyi and Nanyuan-Tongzhou faults present a SW-NE direction. The Huangzhuang-Gaoliying fault dips to the SE with an angle of \(55^\circ \sim 75^\circ\). The Shunyi fault, with a SE dipping angle is a boundary fault controlling the Houshayu Quaternary depression, and the Nanyuan-Tongzhou fault dips towards the NW. The Nankou-Sunhe fault is the unique active NW-SE oriented fault in the Beijing basin. The Northwestern segment of the Nankou-Sunhe fault (from Nankou to Baishan) dipping SW controls the Machikou Quaternary depression with a maximum thickness of Quaternary sediments about 600 m, and the Southeastern segment of the Nankou-Sunhe fault (from Baishan to Tongzhou) dipping towards NE controls the depocenter of Dongba, in the southern part of the Shunyi Quaternary depression \[27,41–46\]. In the Beijing basin, Dongbalizhuang-Dajiaoting subsidence area over the eastern subsidence zone is collectively controlled by the Shuiy fault and the Nanyuan-Tongzhou fault. The Laiguangying subsidence area over northeast subsidence district is mainly controlled by the Shuiy fault and the northwest subsidence zone is mostly led by the Nankou-Sunhe and Huangzhuang-Gaoliying faults, for which the main deformation direction is basically parallel to the fracture direction.

The examples of four profiles of the average subsidence rate for Nankou-Sunhe fault can be found in Figure 10b. Profiles AA’ and BB’ parallel to the Nankou-Sunhe fault and Profiles CC’ and DD’ perpendicular to the Nankou-Sunhe fault with a length of about 10 km are chosen to analyze the spatial characteristics of land subsidence along this fault zone. It can be found that the ground surface displacement along Profile AA’ located at the north of the Nankou-Sunhe fault shows a relative stable displacement rate of around 25 mm/year. However, the profile of average subsidence rate for BB’ located in the south of the fault exhibits a significant increasing trend with the maximum subsidence rate exceeding 70 mm/year. The profiles of the average displacement rate for CC’ and DD’ perpendicular to the Nankou-Sunhe fault also show a high ground surface displacement gradient along the profile with maximum values up to 77 mm/year, indicating the markedly differential settlement existing on both sides of the fault.
5.3. Relationships between Land Subsidence and Compressible Soil Thickness

As is well known, when piezometric levels decrease, the thicker the accumulated compressible soil is, the greater the land subsidence rates are [47,48]. Land subsidence over the Beijing plain mainly develops in the sand, gravel and clayey layer of Quaternary system. Among the biggest contributors to land subsidence is clayey layer. Researches have revealed that given the same underground water exploitation situation, the clayey units (aquitards made of fine soils) are two or three times more deformable than those of sand and gravel (coarse soils) [2]. Figure 11a shows the thickness distribution of the clayey layer existing in the first 100 m depth from surface over the Beijing plain area which is
superimposed on the land subsidence rate map. Figure 11b shows the statistics of InSAR subsidence rates for the considered soft soil ranges. As can be seen in Figure 11a, the subsidence bowls, which present displacement rates ranging from $-110$ to $-60$ mm/year (i.e., about 123329 SDFP pixels within this area from ASAR data), are mostly situated in the zone with a clayey layer thickness of around 50 to 70 m (i.e., about 98190 SDFP pixels within this zone from ASAR data), accounting for 79.6% of the whole amount of subsidence.

![Figure 11](image)

**Figure 11.** (a) Distribution of clayey layers on the study area [28]; (b) Statistics of InSAR-derived land subsidence rate data for the different soft soil ranges.

5.4. Relationships between Land Subsidence and the Aquifer Structure

The structure of the aquifer system in the Beijing plain consists of [20]: (a) coarse-grain sediments with single sandy gravel; (b) median-grain sediments with multiple layers of pebbles and gravels; and (c) fine-grain sediments with multiple sand layers and gravel layers. Figure 12 shows the statistics of
InSAR subsidence data for the different types of aquifers systems. It can be seen that the subsidence bowls with displacement values varying from $-110$ to $-60$ mm/year are located over predominantly fine-grain sediments, accounting for 8% of the total subsidence rate data.

Figure 12. (a) Distribution and (b) statistics of InSAR derived data for the different types of aquifers.

5.5. Relationships between Land Subsidence and Main Groundwater Extraction Areas

Groundwater extraction is the main water supply source in the study region [21]. Groundwater is extracted from different depths for agricultural, industrial and domestic uses. In this section, we evaluate the effect of groundwater extraction from the Huairou emergency wells drilled in 2003 located along the banks of Huai, Sha and Yanxi Rivers and the well field located along the Chaobai River which were drilled in the 70s (Figure 13).

To evaluate the influence of the pumping wells on land subsidence, different buffers at specified distances were created as shown in Figure 13a. The subsidence rates varied from $-35$ to $-5$ mm/year.
in the buffer areas. The statistics of InSAR subsidence rates against the distance to the wells are plotted in Figure 13b. It shows that the proportion of maximum subsidence rates from −35 to −25 mm/year decrease as the distance to the normal pumping wells increases. A similar relationship is found for the emergency pumping wells (Figure 13b). It indicates that, as expected, land subsidence rates decrease with the distance to pumping wells according to the shape of the cone of depression that develops around active pumping wells.

![Image](image.png)

**Figure 13.** (a) Location of the Huairou emergency wells and the well field located along the Chaobai River overlapped to the InSAR subsidence rate map; (b) Statistics of InSAR-derived data against the distance to the Chaobai River (left) and Huairou (right) pumping wells.

6. Conclusions

In this study, land subsidence due to over-extraction of groundwater in the Beijing region was investigated with SBAS InSAR using 41 Envisat ASAR and 14 TSX images. The results reveal that the Beijing region has experienced significant ground subsidence from 2003 to 2010 with a maximum accumulative displacement of 790 mm. The SBAS results have been validated by GPS measurements with a mean difference of 2.41 ± 1.84 mm/year and a RMS of 2.94 mm/year, demonstrating that SBAS InSAR can effectively monitor and detect complicated settlements. A high correlation coefficient of 0.92
and a standard deviation of 7.48 mm/year between both TerraSAR-X and Envisat ASAR InSAR-derived subsidence rates was observed, indicating the reliability of InSAR derived land subsidence rate maps.

The spatiotemporal analysis of land subsidence indicates an increasing trend in the rate and extent of land subsidence. The join spatial and/or temporal analysis of InSAR data and conditioning and triggering factors shows that land subsidence is correlated with groundwater levels, active faults, different soft soil ranges and aquifer types. Furthermore, a relationship between land subsidence and the distance to the pumping wells has been found.

The analysis of different stress-strain relationships shows that the aquifer-system exhibits different behaviour, such as quite elastic behaviour and predominantly inelastic behaviour. Additionally, the analysis of the variation of the inelastic storage coefficient (Skv) of the aquifer system computed from the stress-strain curves shows that the deformability of the aquifer system increases from the north towards the south.

The distribution and development trends of the land subsidence over the study area are obviously controlled by its geological structures. Among the biggest contributors to land subsidence are the clayey layers. The subsidence bowls located in the east, north-east and north are mostly placed over a zone with a clayey layer underneath, with a thickness around 50 to 70 m, and those subsidence bowls are located over predominantly fine-grain sediments. The InSAR data statistics and the distance to the pumping wells show that land subsidence rates are higher near the pumping wells, as expected according to the shape of a well’s cone of depression.

To summarize, accurate InSAR-derived data from two different sensors have allowed us to perform a comprehensive spatial and temporal study of the land subsidence in Beijing basin from 2003 to 2011. Complementarily, these data and existing geo-information have been used to perform a study of the role of the main conditioning and triggering factors controlling land subsidence mechanisms in this area, providing useful information for improved development of future models for the prediction of land subsidence affecting this area.

Acknowledgments: This work was supported by the National Natural Science Foundation of China under Grant 41201419 and a China Scholarship Council (CSC) scholarship to Mi Chen. Roberto Tomás was supported by the Ministry of Education, Culture and Sport through the project PRX14/00100. Part of this work is also supported by the Spanish Ministry of Economy and Competitiveness and EU FEDER funds under projects TIN2014-55413-C2-2-P, by the UK Natural Environmental Research Council (NERC) through the LICs and IRNHIC projects (ref. NE/K010794/1 and NE/N012151/1, respectively), the ESA-MOST DRAGON-3 projects (ref. 10607 and 10665) and the Open Fund from the Key Laboratory of Earth Fissures Geological Disaster, Ministry of Land and Resources (Geological Survey of Jiangsu Province). TerraSAR-X data are copyright of German Aerospace Agency (DLR) and were provided under the research proposal motagh_GEO1916.

Author Contributions: Mi Chen and Roberto Tomás performed experiments, analyzed the data and prepared the manuscript. Zhenhong Li provided crucial guidance and support through the research. Mahdi Motagh made important suggestions on data processing and analysis. Tao Li, Leyin Hu, Huili Gong, Xiaojuan Li, Jun Yu and Xulong Gong contributed significantly to the validation work and data interpretation.

Conflicts of Interest: The authors declare no conflict of interest.

References


9. Prati, C.; Ferretti, A.; Perissin, D. Recent advances on surface ground deformation measurement by means of repeated space-borne SAR observations. *J. Geodyn.* **2010**, *49*, 161–170. [CrossRef]


17. Ng, A. H.; Ge, L.; Li, X.; Zhang, K. Monitoring ground deformation in Beijing, China with persistent scatterer SAR interferometry. *J. Geod.* **2012**, *86*, 375–392. [CrossRef]


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
Anatomy of Subsidence in Tianjin from Time Series InSAR

Peng Liu 1,2, Qingquan Li 1,2,*, Zhenhong Li 3, Trevor Hoey 4, Guoxiang Liu 5, Chisheng Wang 1,2, Zhongwen Hu 1,2, Zhiwei Zhou 4 and Andrew Singleton 3,4

1 Key Laboratory for Geo-Environment Monitoring of Coastal Zone of the National Administration of Surveying, Mapping and GeoInformation & Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen 518060, China; peng.liu.geodesy@gmail.com (P.L.); wangchisheng@szu.edu.cn (C.W.); zwhoo@szu.edu.cn (Z.H.)
2 College of Information Engineering, Shenzhen University, Shenzhen 518060, China
3 Center for Observation & Modeling of Earthquakes, Volcanoes & Tectonics (COMET), School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne, Tyne and Wear NE1 7RU, UK; Zhenhong.Li@newcastle.ac.uk (Z.L.); asingleton@hotmail.co.uk (A.S.)
4 School of Geographical & Earth Science, University of Glasgow, Scotland G12 8QQ, UK; Trevore.Hoey@glasgow.ac.uk (T.H.); zhwei8848@gmail.com (Z.Z)
5 Department of Surveying Engineering, Southwest Jiaotong University, Chengdu 610031, China; rsgxiu@swjtu.edu.cn

* Correspondence: liqq@szu.edu.cn; Tel.: +86-755-2653-6298

Academic Editors: Richard Gloaguen and Prasad Thenkabail

Received: 15 December 2015; Accepted: 14 March 2016; Published: 22 March 2016

Abstract: Groundwater is a major source of fresh water in Tianjin Municipality, China. The average rate of groundwater extraction in this area for the last 20 years fluctuates between 0.6 and 0.8 billion cubic meters per year. As a result, significant subsidence has been observed in Tianjin. In this study, C-band Envisat (Environmental Satellite) ASAR (Advanced SyntheticAperture Radar) images and L-band ALOS (Advanced Land Observing Satellite) PALSAR (Phased Array type L-band Synthetic Aperture Radar) data were employed to recover the Earth’s surface evolution during the period between 2007 and 2009 using InSAR time series techniques. Similar subsidence patterns can be observed in the overlapping area of the ASAR and PALSAR mean velocity maps with a maximum radar line of sight rate of ~170 mm·year⁻¹. The west subsidence is modeled for ground water volume change using Mogi source array. Geological control by major faults on the east subsidence is analyzed. Storage coefficient of the east subsidence is estimated by InSAR displacements and temporal pattern of water level changes. InSAR has proven a useful tool for subsidence monitoring and displacement interpretation associated with underground water usage.

Keywords: InSAR; underground water extraction; subsidence

1. Introduction

Subsidence caused by groundwater extraction is a worldwide problem [1–3]. Estimated subsidence rates range up to several hundred mm per year, highlighting the importance of continuous displacement mapping. With great demand for fresh water and a limited river water resource, ground water use seems inevitable for Tianjin City (Figure 1a), which has a population of 12.93 million and is currently growing at 2.6% per year till 2010 [4]. The maximum pumping rate recorded was in 1981 with the annual rate of 1.2 × 10⁹ m³. The annual rates from 1981 to 2007 are in the range 0.6–0.8 × 10⁹ m³·year⁻¹ (Figure 1b) [5]. As a consequence of this sustained high level of abstraction, an overall drop in the water table has been observed in Tianjin. Water extraction has increasingly been from deeper underground [6]. This extensive use of ground water could account for the subsidence in

Remote Sens. 2016, 8, 266
Tianjin, which has been observed since 1959. The subsidence can be divided into three different stages: slow subsidence at 50 mm·year⁻¹ from 1959 to 1977, accelerating subsidence from 1978 to 1982 at 100 mm·year⁻¹ and attenuating subsidence at 20–30 mm·year⁻¹ from 1983 to 1997 [7]. The maximum total subsidence from 1959 to 2008 has reached 3.2 meters [5]. Locally, higher rates have been recorded: in the 1980s, the maximum subsidence rate is given as 250 mm·year⁻¹ in Tianjin [6]. The current subsidence rate has been estimated as 8–56 mm·year⁻¹ [6] and 30–40 mm·year⁻¹ [5].

InSAR (Interferometric Synthetic Aperture Radar) has been used to map subsidence due to groundwater extraction in Tianjin [8–10] and other areas like Shanghai [11], Mekong delta [12], Ganges-Brahmaputra delta [13], Iran [14–17], Reykjanes Peninsula [18] and so on. In this study, an InSAR time series technique is also used to recover the surface displacement field over the urban area of Tianjin between 2007 and 2009. The patterns of observed subsidence are then interpreted in details and modeled. Subsidence is concentrated in particular areas of the city, and these subsidence centers are analyzed in terms of their geological controls and association with groundwater levels. The effects of groundwater volume changes on surface displacement are investigated for one subsidence center through inverse modeling of InSAR derived subsidence.

Figure 1. (a) The tectonic framework of the Northern China Plain. Boundary of North China Plain is in thick red line. Piedmont plain, alluvial fan and flood plain, and coastal plain in North China Plain are bounded by thick yellow and blue lines. Major faults and subsidiary faults in North China Plain are shown in thin black and red solid lines, respectively [19]. Palaeogene lacuna (missing line of Palaeogene system) is in thin dotted line, while further details of the fault system in Tianjin are shown by thin yellow solid lines [20]. Locations of major cities in this area are shown. Envisat ASAR frame, ALOS PALSAR frame, and political boundary of Tianjin are outlined with dashed lines; (b) Annual groundwater pumping volume and rainfall in Tianjin City from 1980 to 2007 (after [5]); (c) Stratigraphy in Bohai Bay Basin (after [21–23]). The upper Tertiary formations can be divided into Neogene Minghuazhen (Nm) and Guantao (Ng) groups. The lower Tertiary formations can be divided into Eogene Dongying (Ed), Shahejie (Es), and Kongdian (Ek) formations.
2. Geological Background

2.1. Tectonic Divisions

Tianjin Municipality is located 130 km southeast of Beijing and is on the western shore of Bohai Gulf. Tianjin lies on the North China Plain (NCP), also known as Huabei Plain. The NCP is dominated by Haihe and Cangdong fault systems (Figure 1a). The NNE-SSW oriented alignment of Cangdong fault, Tianjin fault, and the Paleogene lacuna, forms a horst-graben structure, which is transversely cut by WNW-ESE aligned faults including Baodi, Hangu, and Haihe faults [24].

In summary, Tianjin is separated by Baodi Fault into the northern Yanshan Folded Belt and the southern Bohai Bay Basin. The Bohai Bay basin contains three tectonic divisions: Jizhong depression, Cangxian uplift, and Huanghua depression. Paleogene lacuna separates Jizhong depression and Cangxian uplift. Cangdong fault separates the Cangxian uplift and the Huanghua depression [25].

2.2. Sediments

The stratigraphy in the Bohai Bay Basin is summarized in Figure 1c. The Bohai Bay Basin is a tectonically subsiding region filled with continental Tertiary and Quaternary sediments [21]. The Quaternary and Tertiary sediments and aquifer system range in thickness from less than 1000 m in the Cangxian uplift [22] to 1200–9000 m in the Jizhong depression, and 900–5000 m in the Huanghua depression [20].

2.3. Hydrogeological Conditions

The ground water system in Tianjin has three main aquifers, the Quaternary, Tertiary and Lower Paleozoic-Sinian aquifers [22]. Above the depth of 700 m in Tianjin, the ground water system can be sub-divided into six aquifer layers (Table 1) [26]. Thickness of Aquifers I to V are in the range of 40–60, 25–60, 30–55, 30–50, and 20–30 m, respectively [6]. Aquifer II is in an over consolidated late Pleistocene strata and Aquifer III is in a normally consolidated middle Pleistocene strata. In Tanggu district of Tianjin, even though aquifer II and III account for 46% and 24% of ground water exploitation, respectively, Aquifer III is the greatest deformation contribution source [7]. Layers between 136 m and 300 m in depth account for more than 65% of subsidence by extensometer observation in Tanggu district of Tianjin [7].

The ground water extracted from a depth greater than 300 meters is mainly used for geothermal exchange. The Tianjin geothermal field covers approximately 11,000 km² [22]. The horst where Tianjin is located is the most productive geothermal field and the highest measured geothermal gradients are found here. More than 200 geothermal wells, of which 10–15 are injection wells, are producing waters from 30 °C to boiling for room heating. Shallow geothermal wells range from 300 m in Quaternary formations to 600 m in Tertiary formations for lower temperatures. Deep wells can reach up to 2600 m in the Tertiary formation [22].

Table 1. Bottom depth of aquifers in Tianjin. Baodi fault and Plain of Tianjin are depicted in Figure 1a. Tanggu district is shown in Figure 2.

<table>
<thead>
<tr>
<th>Location</th>
<th>Tanggu District of Tianjin</th>
<th>Plain to the South of Baodi Fault</th>
<th>Plain of Tianjin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature</td>
<td>[6] (unit: m)</td>
<td>[27] (unit: m)</td>
<td>[26] (unit: m)</td>
</tr>
<tr>
<td>Aquifer I</td>
<td>60–70</td>
<td>&lt;70</td>
<td>40–123</td>
</tr>
<tr>
<td>Aquifer II</td>
<td>170–190</td>
<td>160–220</td>
<td>191–202</td>
</tr>
<tr>
<td>Aquifer IV</td>
<td>370–400</td>
<td>360–430</td>
<td>390–405</td>
</tr>
<tr>
<td>Aquifer V</td>
<td>~550</td>
<td>500–550</td>
<td>505–544</td>
</tr>
<tr>
<td>Aquifer VI</td>
<td>N/A</td>
<td>N/A</td>
<td>608–613</td>
</tr>
</tbody>
</table>
3. Data and Methods

3.1. Data

In order to investigate the subsidence in Tianjin, an InSAR (Interferometric Synthetic Aperture Radar) method is used [28,29], exploring C-band Envisat-ASAR data from European Space Agency (ESA) and L-band ALOS-PALSAR data from Japan Aerospace Exploration Agency (JAXA). ASAR C band images (Table 2) are retrieved with a wavelength of 56 mm compared with PALSAR L band images (Table 3) of 236 mm. The radar line of sight look angles are 23 and 34 degrees off nadir for ASAR and PALSAR, respectively, which means a different sensitivity towards the vertical deformation. The azimuth angles are 166 and 13 degrees counter clockwise from north for ASAR and PALSAR, respectively.

Table 2. Envisat/ASAR archive with acquisition time and length of perpendicular baseline relative to the reference image on 10 October 2008.

<table>
<thead>
<tr>
<th>SAR Image</th>
<th>Acquisition Time</th>
<th>Baseline Perpendicular</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4 January 2008</td>
<td>−401 m</td>
</tr>
<tr>
<td>2</td>
<td>8 February 2008</td>
<td>289 m</td>
</tr>
<tr>
<td>3</td>
<td>14 March 2008</td>
<td>−75 m</td>
</tr>
<tr>
<td>4</td>
<td>18 April 2008</td>
<td>270 m</td>
</tr>
<tr>
<td>5</td>
<td>23 May 2008</td>
<td>32 m</td>
</tr>
<tr>
<td>6</td>
<td>1 August 2008</td>
<td>129 m</td>
</tr>
<tr>
<td>7</td>
<td>1 October 2008</td>
<td>164 m</td>
</tr>
<tr>
<td>8</td>
<td>10 October 2008</td>
<td>0 m</td>
</tr>
<tr>
<td>9</td>
<td>23 January 2009</td>
<td>348 m</td>
</tr>
<tr>
<td>10</td>
<td>27 February 2009</td>
<td>235 m</td>
</tr>
<tr>
<td>11</td>
<td>3 April 2009</td>
<td>624 m</td>
</tr>
<tr>
<td>12</td>
<td>8 May 2009</td>
<td>−48 m</td>
</tr>
<tr>
<td>13</td>
<td>12 June 2009</td>
<td>308 m</td>
</tr>
</tbody>
</table>

Table 3. ALOS/PALSAR archive with acquisition time and length of perpendicular baseline relative to the reference image on 7 December 2008.

<table>
<thead>
<tr>
<th>SAR Image</th>
<th>Acquisition Time</th>
<th>Baseline Perpendicular</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17 January 2007</td>
<td>295 m</td>
</tr>
<tr>
<td>2</td>
<td>4 March 2007</td>
<td>1963 m</td>
</tr>
<tr>
<td>3</td>
<td>20 July 2007</td>
<td>2700 m</td>
</tr>
<tr>
<td>4</td>
<td>20 October 2007</td>
<td>3154 m</td>
</tr>
<tr>
<td>5</td>
<td>5 December 2007</td>
<td>3288 m</td>
</tr>
<tr>
<td>6</td>
<td>20 January 2008</td>
<td>3764 m</td>
</tr>
<tr>
<td>7</td>
<td>21 April 2008</td>
<td>4866 m</td>
</tr>
<tr>
<td>8</td>
<td>6 June 2008</td>
<td>4829 m</td>
</tr>
<tr>
<td>9</td>
<td>22 July 2008</td>
<td>1624 m</td>
</tr>
<tr>
<td>10</td>
<td>6 September 2008</td>
<td>−746 m</td>
</tr>
<tr>
<td>11</td>
<td>7 December 2008</td>
<td>0 m</td>
</tr>
<tr>
<td>12</td>
<td>22 January 2009</td>
<td>484 m</td>
</tr>
<tr>
<td>13</td>
<td>9 March 2009</td>
<td>938 m</td>
</tr>
<tr>
<td>14</td>
<td>24 April 2009</td>
<td>1490 m</td>
</tr>
<tr>
<td>15</td>
<td>9 September 2009</td>
<td>1756 m</td>
</tr>
</tbody>
</table>

3.2. Methods

Single look complex image was generated using ROI PAC [28]. Interferograms were generated using Doris [30]. SRTM (NASA’s Shuttle Radar Topography Mission) with a 3 arc-second resolution was used for topography removal [31]. The StaMPS (Stanford Method for Persistent Scatterers) software was used to process the images using Persistent Scattering (PS) methods inside the package [32,33].
In contrast to most other PS methods, StaMPS uses phase spatial correlation to identify PS pixels instead of amplitude analysis [30,34,35]. The advantage of this strategy is the capability to detect persistent scattering (PS)/slowly decorrelating filtered phase (SDFP) pixels with low amplitude, which is often the case in natural terrains. The probability for a pixel to be PS/SDFP is estimated through phase analysis, which is successively refined in a series of iterations. Without any prior assumption about the temporal nature of ground deformation, StaMPS relies on the spatial correlation of deformation rather than any assumption of the temporal dependence of deformation [36].

For selected PS points, the two dimensional phases are still modulo $2\pi$. Phase unwrapping is therefore implemented in order to derive continuous displacement fields. On each interferogram, phase unwrapping is implemented using the SNAPHU (Statistical-Cost, Network-Flow Algorithm for Phase Unwrapping) algorithm [37]. SNAPHU is a statistical-cost network flow algorithm. The algorithm aims to compute the most likely unwrapped solution with maximum posterior probability estimation given the observable input data.

The topography of the NCP (North China Plain) is quite gentle because it is formed in a coastal area by sediment aggradations. Large DEM errors are not expected here. The look angle error is estimated in time series using perpendicular baselines of each scene [32].

The atmospheric effects exhibit spatially correlated pattern. Deformation pattern are also spatially correlated. However, displacement usually stays in the same area while the atmospheric phase screen (APS) pattern is variable from one image to another. For a single pixel, the deformation should be continuous in time, APS will cause unexpected signal in time series. The APS can be estimated according to its different spatial and temporal nature from displacement. The master atmospheric effect and orbital error is estimated by its presence in every interferogram. The slave atmospheric effect and orbital error is estimated by temporal high pass and spatial low pass filtering [32]. As a flood plain, our research area in NCP shows gentle topography with height ranges between $-16$ and $+68$ m. As a result, the topography dependent APS [38–40], which could stay in the same area and be misunderstood as displacement, are believed to be small. Long-wavelength atmospheric effects and imprecise orbit errors were reduced by a planar phase gradient.

4. Results

4.1. Mean Velocities from ALOS and Envisat

The mean phase values of each image are firstly used as the reference phase in time series analysis because no ground truth data are available. The area with obvious uplift displacements on the resulting mean velocity map was then chosen as the reference area (small triangle in Figure 2). The Envisat and ALOS mean velocity rates are not uniform across the study area with similar pattern of individual subsidence zones on both maps. Two subsidence areas can be identified from the mean velocity map (Figure 2).

The west subsidence area is located between Tianjin and its northwest neighbor city Langfang. The velocity swath in the west subsidence area exhibits similar pattern between ALOS and ENIVSAT measurements (Figure 3a), which suggest that the mean velocities in the west subsidence area are reliable.

The other east subsidence area is located around the Tuanbo Reservoir to the east of Jinghai County. The newly built Jing-Hu High Speed Railway also passes through this area (Figure 2). The mean velocity swath profile along JH Railway (Figure 3b) indicates the overlap of Jing-Hu High Speed Railway and the east subsidence area e.g., between 38.8° and 39.05° N.
Figure 2. LOS (line of sight) rates from ALOS (17 January 2007 to 9 September 2009) and Envisat (4 January 2008 to 12 June 2009) satellites. Negative values mean moving away from the satellite (e.g., subsidence). The straight line “TJLF (Tianjin Langfang)” is located between Tianjin City and Langfang City. The poly line “HSR (High Speed Railway)” is the path of Jing-Hu high speed railway. The displacement rates along the two swaths are shown in Figure 3. BM2010 and BM2011, along with other benchmarks, are shown in white dots. Leveling results for all benchmarks are compared with InSAR displacements in Figure 4. Reference points for InSAR and leveling are shown in black triangle and dot respectively. Model area is outlined by rectangle. Counties and districts in Tianjin are outlined by thin dashed line. Geothermal fields with gradient above 3.5 (°C/100 m) are outlined by thick dashed line.

Figure 3. (a) LOS rates against longitude in TJLF swath with buffer distance of 500 meters to both sides of the line; and (b) LOS rates against latitude in a swath along the Jing-Hu High Speed Railway with buffer distance of 500 meters to the railway at both sides.
4.2. InSAR and Leveling Comparison

Leveling data are available in May 2007, March 2008, September 2008 and March 2009 from 13 benchmarks along a 20 km section of the Jing-Hu High Speed railway in this area. InSAR results are compared with the leveling survey results. Comparison between InSAR and leveling is not straightforward due to different reference systems. Leveling reference point for Tianjin is the Liqizhuang bedrock point located in Tianjin institution of surveying and mapping. Liqizhuang bedrock point is a deep stratum benchmark, with its base 1088 m underground in Cambrian formations [41]. The InSAR results are referenced to the mean value of phases first. Then, the areas with the greatest uplift signals are chosen as the InSAR reference area. Comparison between leveling and InSAR displacements can be made through a double difference in time and space [42]. The spatial reference for InSAR and leveling comparison is leveling point BM2010, which exhibited the least subsidence among all leveling points. The temporal reference for InSAR and leveling comparison is set at the first leveling campaign in May 2007. Firstly, LOS InSAR displacements are linearly interpolated at the leveling times. Secondly, Interpolated LOS InSAR results are double differenced using the spatial and temporal reference. Finally, LOS Envisat and ALOS results are decomposed into east-west and vertical components by ignoring north-south horizontal movement for a polar orbit SAR satellite [43]. To be consistent with InSAR results, the leveling results are re-referenced to BM2010 at May 2007 as well. It should be noted that first Envisat acquisition is later than first leveling campaign, so InSAR and leveling results are temporally re-referenced to March 2008 when necessary.

The RMS differences are 9.7 mm between ASAR LOS and leveling, 8.8 mm between PALSAR and leveling, and 8.3 mm between vertical component and leveling (Figure 4a–c). Slightly improvement is seen after decomposition of horizontal movement from vertical (Figure 4c,f). The reason lies in

![Figure 4](image)
the fact that the accuracy of Leveling data, interpolated ALOS and Envisat displacement is limited. On the one hand, the accuracy of interpolated ALOS and Envisat displacement is limited by seasonal movement. For example, ALOS displacement on 1 March 2009 is interpolated from displacements of 22 January and 9 March 2009, while Envisat displacement on 1 March 2009 is interpolated from displacements of 27 February and 3 April 2009. The water level began to drop in late March, so the interpolated Envisat displacement is greater than ALOS in March 2009. It seems that the decomposition of horizontal and vertical movement is not highly effective in the presence of seasonal displacement. On the other hand, the leveling dates available are only accurate to months, and the first day of the month is assumed as the leveling date for a campaign. However, it still can be seen that interpolated ALOS and Envisat displacements generally follow the displacement trend of benchmarks, although with less displacement (Figure 4d,e). It is likely that the vertical displacement is dominant to horizontal displacement along the leveling line, and the projection from vertical to LOS results in displacement reduction for both ascending and descending tracks.

4.3. Model of West Subsidence

The western subsidence is modeled using Mogi source solution in a semi-analytical approach [44,45].

$$u_z (x, y) = \sum_i (1 - \nu) \frac{\Delta V_i}{\pi} g_z (r_i, d)$$

(1)

$$g_z (r, d) = \frac{d}{(r^2 + d^2)^{3/2}}$$

(2)

$$r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2}$$

(3)

where $u_z (x, y)$ is the vertical displacement at $(x, y)$, $\Delta V_i$ is the volume change at reservoir element $i$, $g_z (r, d)$ is the vertical green/influence function, $\nu$. is the Poisson’s ratio, $r_i$ is horizontal distance between surface point $(x, y)$ and reservoir element $(x_i, y_i)$, $d$ is reservoir depth. The reservoir is divided into equal squares.

A semi analytical inversion is performed for Mogi source array to determine the best fit distribution of source volume change for LOS annual displacements (annual rates times one year). Modeled surface displacements are the ensemble contributions from each source. Poisson’s ratio of 0.25 is set for this study. Reservoir depth is fixed at about 200 meters as ground water is pumped from a depth of 100–300 m in Tianjin [7]. The linear least-squares inversion procedure has been adopted to estimate the source volume change.

It is well known that InSAR observations may contain global bias due to uncertainties in reference level or other uniform signals (e.g., sediment compaction). Accordingly, a constant offset is allowed in model inversion to accommodate the global bias.

In the process of inversion, regularization on the source strength is often needed to ensure a reliable inversion with a faithful representation of the source. Directly modeled reservoir volume changes (Figure 5a) seem noisier than the regularized source (Figure 5b). This can be caused by displacement uncertainty due to noise and other error sources. Besides, the semi analytical approach allows approximation of reservoir change and efficient computations, but it does not assure the physical process that must involve a continuous contracting/inflating volume. A Laplacian regularization based process was applied to smooth the source (Figure 5b).

$$\begin{bmatrix} G & 1 \\ LF & 0 \end{bmatrix} \begin{bmatrix} S \\ O \end{bmatrix} = \begin{bmatrix} D \\ 0 \end{bmatrix}$$

(4)

where $G$ is the Green’s function, $L$ is the Laplacian smooth operator, $F$ is the smoothing factor (weight), and their product $LF$ is the smoothing matrix used in model inversion. $S$ is the source volume change,
O is the offset term for global bias, and $D$ is the observed displacement. In order to balance the roughness of source volume change and model fit to displacement, the L curve method (Figure 6) is adopted to find the best smoothing factor [46].

![Figure 5. Mogi reservoir model: (a) no smooth applied; and (b) Laplacian smooth.](image)

![Figure 6. L curve to determine source smoothness. The horizontal axis denotes the modeled source smoothness $\log(||LS||_2)$, where $L$ and $S$ are defined in Equation (4). The vertical axis denotes source fit to InSAR observations $\log(||GS - D||_2)$, where $G$ and $D$ are defined in Equation (4). The curve is formed by 13 different Laplacian regularized sources $S$. Each $S$ corresponds to the best-fit result of each inversion. Smoothing strength for the source is controlled by $LF$ in model inversion. There are 13 smoothing factors $F$ range between $10^{-4}$ to $10^{-1}$ with exponential step of 0.25. Note that horizontal axis $LS$ is calculated without multiplying smoothing factor $F$. The square centered with an asterisk represents that a value of $10^{-3}$ is employed as the smoothing factor for final result.](image)

The modeled source volume change is about $-7694$ to $+1678$ m³ for each element (300 m square) per year (Figure 7), equivalent to water volume change of $-85,000$ to $+18,000$ m³/km²/year, further equivalent to water storage change of $-85$ to $+18$ mm·year$^{-1}$ in height. Unfortunately, to the best of our knowledge, in situ measurements of fluid volume from production wells is unpublished and is not available for comparison with our modeling results. Ground water storage (GWS) change from GRACE satellite measurement is about $-17$ to $-22$ mm·year$^{-1}$ in North China Plain [47–49], equivalent to $-17,000$ to $-22,000$ m³/km²/year. The GRACE measurement covers Beijing, Tianjin, Hebei and Shanxi of $-370,000$ km², while our model is in an area of 860 km² in Tianjin. Simulated
recoverable groundwater storage depletion is $-30,000 \text{ m}^3/\text{km}^2/\text{year}$ (equivalent to $30 \text{ mm \ year}^{-1}$ in height) from 1970 to 2008 [50], equivalent to $-2700 \text{ m}^3$ for each element (source) per year. Contour line of $-2700 \text{ m}^3$ is superimposed on source volume change (Figure 7d). Line $-2700 \text{ m}^3$ coincides with sharp reduction of both the water extraction volume and subsidence observed. Modeled source volume change show local water storage increase up to $1678 \text{ m}^3$. InSAR observations show subsidence of $-168.5$ to $-2.9 \text{ mm}$ in model area (Figure 7a). The volume increase seems to contradict with the fact that only subsidence is observed in this area. From the perspective of semi-analytical modeling, any surface deformation is the ensemble contribution of all model sources via their influence functions. Therefore, for a surface point, the superposition of a remote contracting source with large volume decrease, and a nearby inflating source with small volume increase, may still generate a surface subsidence. From the perspective of physical process, aquifer can benefit from precipitation infiltration, irrigation return, or other vertical or horizontal recharge. For instance, rainfall in summer recharge local ground water aquifers in North China Plain by $12–29 \text{ mm}$ seen from GRACE and $25 \text{ mm}$ seen from ground boreholes [48].

Figure 7. (a) ALOS LOS annual displacements (annual rates times one year); (b) modeled ALOS LOS annual displacement; (c) residuals between observed and modeled displacements for ALOS; (d) Envisat LOS annual displacements; (e) modeled Envisat LOS annual displacement; (f) residuals between observed and modeled displacements for Envisat; and (g) modeled Source array volume changes from ALOS displacement. Dashed line is a contour of $-2700 \text{ m}^3$.

The horizontal displacement can be modeled using horizontal Green’s function and the best-fit source volume change.

$$g_r (r, d) = \frac{r}{(r^2 + d^2)^{3/2}}$$

where $g_r (r, d)$ is the horizontal green/influence function. Modeled horizontal displacements can be substantial, with maximum annual horizontal displacements reach up to $74 \text{ mm}$ (Figure 8). Maximum horizontal displacements are located along a semi closed oval around the modeled source center. The modeled maximum horizontal displacement is 36% of the maximum vertical displacement. Theoretical ratio of maximum horizontal to maximum vertical displacement is 38% for a single Mogi source [51]. A field based aquifer test in Nevada shows that the horizontal displacements reach $8 \text{ mm}$.
and vertical displacements reach 12 mm within the first 22 days of pumping before reaching the steady state pumping [52]. Thus, the horizontal displacement induced by aquifer abstraction can be significant. Horizontal displacements are believed to be a cause of ground fissures [53]. Ground fissures associated with subsidence has caused severe infrastructure damage in Taiyuan Basin of China [54]. Hence, it is worth measuring the horizontal displacements due to ground water depletion, including Tianjin.

Figure 8. Modeled annual: (a) horizontal displacements; (b) horizontal displacements in E-W direction; and (c) horizontal displacements in N-S direction.

5. Discussion

5.1. Impact of Reservoir Grid Size

Grid-like patterns stand out in the modeled displacement (Figure 9b). The hypotheses that these are related with reservoir grid size are examined. The 300 m reservoir is now resampled into 270 m, 225 m, and 75 m grids (Figure 9c–e). At the same time, reservoir volume changes are resampled by size to fit the surface displacement. It can be seen that surface displacements are well represented when finer reservoir grids are adopted without alternating surface grid size. Hence, the grid patterns that obscure the modeled displacement are mainly due to the reservoir grid size in this case.

It should be noted that Mogi-type modeling assumes that source radius $\alpha$ is much smaller than source depth $(\alpha/d \approx 0)$. A radius of 1 km is usually assumed for volcano cases [51]. The maximum displacement for a finite sphere is about $1 + (\alpha/d)^3$ times of that for a point sphere [51]. Accordingly, the grid-like patterns appear when the ratio of cavity radius to depth is significant (Figure 9). Hence, water storage change can be underestimated when reservoir grid size is similar to its depth. However, fine
reservoir grid causes quadratic growth of source numbers, greatly increasing computational burden in model inversion.

Figure 9. Observed LOS surface displacement (a) and Modeled LOS displacements with reservoir grid size of: (b) 300 m; (c) 270 m; (d) 225 m; and (e) 75 m.

5.2. Tectonic Division and Its Effect on Subsidence

There are two kinds of explanations when fault act as a subsidence barrier. First, sediment deposits separated by the fault exhibit different compressibility [55]. Second, ground water hydraulic conductivity is low around the fault [56]. In this section, major faults in Tianjin are examined to see if they are related with subsidence observed. The tectonic division could help address the subsidence differences seen in Tianjin.

The east subsidence, outlined by B1 in Figure 10b, is controlled by Tianjinbei fault to the west, and Dasi fault to the east. Most of the east subsidence is located in the area of Shuangyao salient. Velocity differences (abrupt InSAR mean velocity changes) can be found between the areas to the west and to the east of lines F1, F2, and F3 (The F lines in Figure 10b are drawn by visual inspection) (Figure 10b). Line F1 coincides with a fault to the west of Baitangkou depression [57]. Line F2 matches the central segment of Dasi fault [58]. Dasi fault is also known as Baitangkou fault. Line F3, however, bends away from Dasi fault and is oriented in NE-SW direction. It remains unclear if a fault exists along F3. Further geological investigation is needed to confirm this.

The Tianjinbei fault is striking NNE (strike is a line representing the intersection of a fault with a horizontal plane) and dipping NW (dip is the angle of descent of a fault relative to a horizontal plane). The Tianjinnan fault is also striking NNE while dipping SE. The Tianjinbei fault and the Tianjinnan fault are 6–7 km apart. Tianjinbei fault joins Haihe fault to the north in Tianjin city center [59]. A velocity strip (a long and narrow zone with mean velocities different from neighboring area) matches the location of the Tianjinnan fault (Figure 10). Dislocated stratum appears at about 47 m in depth from
TJ08 seismic profile of Tianjinnan fault [59]. Boreholes TS1 and TS4 near profile TJ08 show that a 9.5 m fault throw (the vertical component of the dip separation) appears at 88.3–90.3 m in depth to the west of Tianjinnan fault and at 98.9–99.7 m in depth to the east of Tianjinnan fault [60].

Haihe fault separates Tianjinbei fault into north and south segments. A velocity strip matches the north segment of Tianjinbei fault (Figure 10). In the north segment, dislocated stratum appears at 132 m and 125 m in depth from TJ01 and SY02 seismic profiles of Tianjinbei fault [59]. Weak velocity differences can also be observed along line F4 (Figure 10b), which can be the south segment of Tianjinbei fault.

Figure 10. Faults and LOS subsidence rates in Tianjin: (a) Tectonic system of the area includes Dacheng Salient (DCS), Shuangyao Salient (SYS), Panzhuang Salient (PZS), Xiaohanzhuang Salient (XHZS), Litan Depression (LTD), Banqiao Depression (BQD), and Baitangkou Depression (BTKD) [57–59,61,62]. Major faults include the Tianjinnan Fault (TJNF), Tianjinbei Fault (TJBF), Dasi Fault (DSF), Haihe Fault (HHF) and Cangdong Fault (CDF); (b) TJ01 and SY02 are seismic profile of TJBF. TJ08 is seismic profile of TJNF, and TS1 and TS4 are boreholes near TJ08. TJ23 is seismic profile of CDF [59].

Cangdong fault is striking NNE and dipping SE. Tianjin segment of Cangdong fault is cut by Baodi fault and haihe fault [62]. Velocity differences are seen along lines F6 and F7 (Figure 10b). Dislocated stratum appears at 152 m in depth from TJ23 seismic profile of Cangdong fault [59].

Haihe fault is striking NW to NWW and dipping SSW [59]. Distinguishable minimum depth of Haihe fault is 60–70 m [59]. It crosses the Tianjin urban area where uniform velocities are observed.

The east subsidence area shows greater subsidence rates than urban areas of Tianjin on both Envisat and ALOS mean velocity maps. This is consistent with the fact that the area of ground water extraction expanded to the suburban area since 1990s and the ground water pumping volume in urban area has declined from more than 100 million in 1981 to less than 10 million cubic meters in 2007 [5].

Within a geological uplift, the east subsidence is likely the result of heavy underground water use. Geological control of the east subsidence is seen along DSF and TJBF. Subsidence is likely to be related with stratum below 125 m seen from TJBF.

5.3. Water Level and Subsidence

In the east subsidence, the time series displacement of JH-12 Well is plotted against contemporary precipitation and its water level records of Aquifer II (Figure 11). It can be seen that seasonal InSAR displacements near JH-12 generally follows the underground water levels of JH-12. North China Plain is one of the largest agricultural production area in China, accounting for about 50% of wheat and 33%
of maize grain production in China [63]. The boot growth stage of winter wheat is April, when water demand by irrigation reaches its peak level. The water table was the lowest from May to June each year. The irrigation return and abundant precipitation in July and August infiltrated and recharged ground water [64]. The water level then rose to the highest level in March or April the following year.

Figure 11. Envisat and ALOS LOS displacement at JH-12 Well (the triangle in Figure 12b) in east subsidence area. Precipitation data are available from China Meteorological Administration, specifically meteorology station No. 54527 in Tianjin [65]. Water level of JH-12 Well in Aquifer II is available to the end of 2008 [27].

The water level pattern of JH-12 resembles the lateral recharge-runoff-discharge pattern in the piedmont of North China Plain [66]. However, JH-12 is at least 100 km away from the piedmont plain zone besides Yanshan Mountain and Taihang Mountain, and is located in alluvial fan and flood plain [67]. In central North China Plain, water level is normally low in winter and spring with strong evaporation and less precipitation [66], while JH-12 show consistently high water level in winter and spring except a tiny fluctuation of ca. 0.5 m around December (Figure 11). It is likely that JH12 is laterally recharged since Tuanbo Reservoir is nearby (Figure 12b).

Water level of JH-12 declines sharply from $-62.57$ m to $-69.03$ m between 23 March and 26 April 2008. Consequent subsidence between 21 April and 22 July in 2008 reaches about 63 mm in ALOS LOS direction. The short term aquifer compaction should be elastic as the water level recovers to a nearly identical level next march. Hence we evaluate the elastic storage coefficient $S_{ke}$ near JH-12 by relating aquifer compaction and effective stress change. Aquifer compaction and stress change are reflected by ALOS displacement and hydraulic head change respectively [68]. The seasonal LOS deformation, satellite incidence angle, and water level drop, yield an elastic storage coefficient of 0.0125 for JH-12 from ALOS. Although 18 mm LOS subsidence is observed between 23 May and 1 August 2008 from Envisat, no ASAR image can be found in June or July, when subsidence reaches its maximum value seen from ALOS.

In a numerical study of flow system simulation in the North China Plain, storage parameter set for Aquifer III is in the range of 0.0001–0.005 with geometric mean value of 0.0012, and storage parameter set for Aquifers I and II are in the range of 0.04–0.25 with geometric mean of 0.075 [50].
The elastic storage near JH-12 Well triples the maximum storage parameter set for Aquifer III, but is merely one-third of the minimum storage parameter set for Aquifers I and II [50]. Hence Aquifer II and other aquifer layers are likely contribute together to the observed surface subsidence.

Subsidence rates are compared with water level of Aquifer III and Aquifer IV in 2008 [26] (Figure 12). Notably, the water levels are relative because the depth of Aquifer III and Aquifer IV are in the range of 180–400 m for the east subsidence area in their paper. The water level in Aquifer III of the east subsidence area is −70 to −60 m in depth, while it is >−55 m in the urban area of Tianjin in 2008. Higher subsidence rates are associated with lower water level in Aquifer III. Aquifer III is likely to be a subsidence source to the east subsidence due to similarity in map patterns between subsidence and lower water levels. Correlation between the water level of Aquifer IV and subsidence pattern are not seen.

A wavelet analysis on InSAR displacement and water level time series is implemented (Appendix). Displacement periodicities and time shift between water level and displacement time series are analyzed.

Figure 12. Water levels in 2008 superimposed on LOS subsidence rates from Envisat (4 January 2008 to 12 June 2009): (a) Aquifer III; and (b) Aquifer IV. Location of JH-12 Well is marked in triangle. Tuanbo Reservoir is labeled.

6. Conclusions

Subsidence up to 170 mm·year−1 is observed in the Tianjin Municipality from Envisat/ASAR and ALOS/PALSAR images collected between 2007 and 2009. The subsidence rates are similar across the overlap area between Envisat and ALOS results. Subsidence along the JH High Speed Railway has also been verified by four leveling campaigns between 2007 and 2009 with a RMS difference of 8–9 mm between InSAR and leveling.

The west subsidence is modeled using Mogi source array. Best-fit water extraction volume is obtained through model inversion. Estimated water volume change is −85,000 to +18,000 m3/km2/year (equivalent to −85 to +18 mm·year−1 in height). High gradients of subsidence and water storage change are associated with the recoverable water volume change of −30,000 m3/km2/year (equivalent to −30 mm·year−1 in height) from other numerical simulation study.

Geological control of Tianjinman, Tianjinbei and Dasi faults on the east subsidence can be seen from InSAR displacement rates. The east subsidence can be related with Aquifer II, seen from temporal association of subsidence and water level of JH12. The east subsidence can also be related with Aquifer III, seen from elastic storage coefficient and spatial association of subsidence and water level of Aquifer III.
This study demonstrates the capability of time series InSAR to map subsidence, model ground water volume change, identify geological control factors, and estimate storage coefficients in Tianjin. InSAR can be an effective tool to investigate ground water use conditions in Tianjin and NCP.

Acknowledgments: This work is supported by a NSFC (National Natural Science Foundation of China) Program to Peng Liu (No. 41504003), a NSFC program to Chisheng Wang (No. 41504004), and Shenzhen-Hong Kong Innovation Circle Program (No. SGLH201502016152559032) to Peng Liu and Chisheng Wang. The ENVISAT images were supplied through the ESA-MOST Dragon 2 Cooperation Program (ID: 5343). Qingquan Li is supported by grants from Shenzhen Scientific Research and Development Funding Program (No. ZDSY20121019111146499 and No. JSGG20121026111056204), Shenzhen Dedicated Funding of Strategic Emerging Industry Development Program (No. JCYJ20121019111128765), and Shenzhen Future Industry Development Funding Program (No. 201507211219247860). Zhenhong Li is supported by the UK Natural Environmental Research Council (NERC) through the LICP project (ref. NE/K010794/1), the ESA-MOST DRAGON-3 projects (ref. 10607 and 10665) and Open Fund from the Key Laboratory of Earth Fissures Geological Disaster, Ministry of Land and Resources (Geological Survey of Jiangsu Province). We thank JPL/Caltech for the use of ROI PAC, TU-Delft for DORIS and Andy Hooper for StaMPS in our data processing and analysis. Part of the figures were prepared using the public domain Generic Mapping Tools.

Author Contributions: Peng Liu performed experiments and drafted the manuscripts. Qingquan Li, Zhenhong Li, Trevor Hoey and Guoxiang Liu contributed to InSAR data analysis. Chisheng Wang, Zhongwen Hu, Zhiwei Zhou and Andrew Singleton contributed to model calculations.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

- InSAR: Interferometric Synthetic Aperture Radar
- ASAR: Advanced Synthetic Aperture Radar
- ALOS: Advanced Land Observing Satellite
- PALSAR: The Phased Array type L-band Synthetic Aperture Radar
- NCP: North China Plain
- ROI_PAC: Repeat Orbit Interferometry PACkage
- DORIS: The Delft Object-oriented Radar Interferometric software
- StaMPS: Stanford Method for Persistent Scatterers
- SRTM: Shuttle Radar Topography Mission
- PS: Persistent Scatter
- SDFP: Slowly-Decorrelating Filtered Phase
- APS: Atmospheric Phase Screen
- GWS: Ground Water Storage

Appendix

Wavelet tools are used to analyze the seasonal displacements observed from Envisat and ALOS and the associations between displacements and water level time series at JH-12 Well. The wavelet tools employed in this study are continuous wavelet transform (CWT), cross wavelet transform (XWT), and wavelet coherence (WTC). CWT decomposes one dimensional InSAR time series into two dimensional time frequency space to determine the local periodicities of seasonal variations [69]. The cross wavelet transform (XWT) and wavelet coherence (WTC) tools indicate the common power and relative phase between the two time series [70,71].

Common interval for Envisat displacement and water level time series is between 4 January 2008 and 10 October 2008. Linear component of Envisat displacements is estimated by least square method and subtracted from the time series. The Envisat image number is limited to eight, and down sampling of water level time series at image acquisition time results in very coarse power spectrum. Oppositely, the non-linear component of Envisat displacement is over sampled by linear interpolation at water level time. However, the interpolation also reduces the accuracy of displacements. Envisat
time series show a two-month cycle (64-day period) between April and May 2008 from CWT analysis (Figure A1b). No effective cycle can be observed for water level time series (Figure A1d). XWT analysis shows a phase rotation from about 120° out of phase in the one and a half month cycle (48-day period) to about 120° in phase in the two month cycle (64-day period) between April and May 2008 (Figure A1e). Abrupt phase changes occur between adjacent 48-day band and 64-day band. This implies strong seasonal variations in a short time. It corresponds to the quick water level drop in April. WTC analysis shows about 90° out of phase in cycles of 4–8 days and 16–32 days (Figure A1f). It means the displacement leads water level by 1–2 days or 4–8 days. Actually, this is due to lower temporal resolution of Envisat images compared with water level time series. The lowest water level is between the 4th and 5th Envisat image acquisition time, so it looks as if displacement from 4th to 5th image happens before the water level rise. In fact, the 4th image corresponds to a water drop stage while the 5th image corresponds to a water rise stage.

Figure A1. Cont.
Figure A1. (a) Detrend of Envisat time series at JH-12 Well; (b) continuous wavelet power of the Envisat time series; (c) water level of JH-12 Well; (d) continuous wavelet power of water level time series; (e) cross wavelet transform of Envisat and water level time series at JH-12 Well; and (f) wavelet coherence of Envisat and water level time series for JH-12. Images (g–l) are the same for ALOS and water level time series at JH-12 Well. The thick contour is the 95% confidence level. The cone of influence (COI) is shown in light shadow.

Common interval for ALOS and water level time series is from 17 January 2007 to 7 December 2008. CWT analysis of ALOS time series shows a four-month cycle (128 days) between the 8th (6 June 2008) and the 10th (6 September 2008) ALOS image. CWT analysis for water level time series also shows a 128-day cycle, though at a lower confidence level. XWT analysis show 90° in phase in the 80-day band and 90° out of phase in the 128-day band, from the 6th (20 January 2008) to the 9th (22 July 2008) ALOS image. This means strong seasonality occurs between the two images. WTC analysis shows about 90° in phase in cycles of 4–8 days near the 7th ALOS image (21 April 2008). This means water level leads displacement by 1–2 days, which is beyond the 46-day repetition frequency of ALOS acquisition. WTC analysis also shows 90° in phase in cycles of 64 days between 7th (21 April 2008) and 10th (6 September 2008) ALOS image, although with lower confidence level. It means water level leads displacement by 16 days. The displacement began to stabilize at the 10th (6 September 2008) ALOS image, while the water level’s recent rise stops at 8 August 2008, which is about 29 days ahead of the
10th ALOS image. The phase shift of 16 days between water level and displacement time series can be explained by the delay of 29 days. Moreover, it also implies the deformation may have already stabilized before 6 September 2008, because the phase (time) shift from continuous wavelet analysis is more robust than direct comparison of isolated time series points.

Wavelet tools are able to find the periodicities of InSAR time series and relative phase (time) shift between InSAR and water level time series. However, the wavelet power spectrum can be misleading if the temporal resolution of time series is low. Meanwhile, the phase (time) shift from continuous wavelet analysis enhances the robustness of time delay analysis between time series.

References


42. Leighton, J.M.; Sowter, A.; Tragheim, D.; Bingley, R.M.; Teferle, F.N. Land motion in the urban area of nottingham observed by ENVISAT-1. Int. J. Remote Sens. 2013, 34, 982–1003. [CrossRef]


50. Zhu, W.; Zhang, Q.; Ding, X.; Zhao, C.; Yang, C.; Qu, W. Recent ground deformation of Taiyuan basin (China) investigated with C-, L-, and X-bands SAR images. *J. Geodyn.* 2013, 70, 28–35. [CrossRef]


71. Tomás, R.; Li, Z.; Lopez-Sanchez, J.M.; Liu, P.; Singleton, A. Using wavelet tools to analyse seasonal variations from InSAR time-series data: A case study of the Huangtupo landslide. Landslides 2015. [CrossRef]
Ground Subsidence in the Beijing-Tianjin-Hebei Region from 1992 to 2014 Revealed by Multiple SAR Stacks

Yonghong Zhang1,*, Hong’an Wu1, Yonghui Kang1 and Chuanguang Zhu2
1 Chinese Academy of Surveying and Mapping, Beijing 100830, China; wuha@casm.ac.cn (H.W.); kangyh@casm.ac.cn (Y.K.)
2 Hunan Province Key Laboratory of Coal Resources Clean-utilization and Mine Environment, Hunan University of Science and Technology, Changsha 411201, China; zhuchuanguang@163.com
* Correspondence: yhzhang@casm.ac.cn; Tel.: +86-10-6388-0521

Academic Editors: Zhenhong Li, Roberto Tomas, Zhong Lu, Xiaofeng Li and Prasad S. Thenkabail
Received: 6 April 2016; Accepted: 15 August 2016; Published: 20 August 2016

Abstract: The coordinated development of the Beijing-Tianjin-Hebei has become a national strategy with Beijing and Tianjin as twin engines driving the regional development. However, the Beijing-Tianjin-Hebei region has suffered dramatic ground subsidence during last two to three decades, mainly due to long-term groundwater withdrawal. Although, annual spirit leveling has been conducted routinely in some parts of Beijing and Tianjin, and InSAR technique has also been used to monitor ground subsidence in some local areas of the region, there is a lack of a complete survey of ground subsidence over the whole region. In this paper, we report a research on mapping ground subsidence in the Beijing-Tianjin-Hebei region over a long time span from 1992 to 2014. Three SAR datasets from four satellites are used: ERS-1/2 SAR images from 1992 to 2000, ENVISAT ASAR images from 2003 to 2010, and RADARSAT-2 images from 2012 to 2014. An improved multi-temporal InSAR method, namely “Multiple-master Coherent Target Small-Baseline InSAR” (MCTSB-InSAR), has been developed to process the datasets. A unique feature of MCTSB-InSAR is the adjustment process useful for wide area monitoring which provides an integrated solution for both calibration of InSAR-derived deformation and the harmonization of the deformation estimates from overlapping SAR frames. Three maps of the subsidence rate corresponding to the three periods over the wide Beijing-Tianjin-Hebei region are generated, with respective accuracy of 8.7 mm/year (1992–2000), 4.7 mm/year (2003–2010), and 5.4 mm/year (2012–2014) validated by more than 120 leveling measurements. The spatial-temporal characteristics of the development of ground subsidence in Beijing and Tianjin are analyzed. This research represents a first-ever effort on mapping ground subsidence over very large area and over long time span in China. The result is of significance to serve the decision-making on ground subsidence mitigation in the Beijing-Tianjin-Hebei region.

Keywords: multi-temporal Interferometric SAR; ground subsidence; Beijing-Tianjin-Hebei

1. Introduction

Beijing, the Capital of China, and Tianjin, one of the four municipalities of China, both play very important roles in the nation’s political and economic framework. Especially since the coordinated development of Beijing-Tianjin-Hebei became a national strategy in 2013, in which Beijing and Tianjin are positioned as the twin engines driving the regional development. On the other side, due to insufficient rainfall in the North China plain, groundwater has been the main water resources for farmland irrigation, industrial activities, and daily water supply of urban inhabitants since 1970s [1]. Long-term over-withdrawal of groundwater has caused continuous and dramatic decline of groundwater level in the aquifer system of this region. For example, the groundwater level has
declined 2.7 m every year in Chaoyang district of Beijing from 2001 to 2010 [2]. Continuous pumping of aquifer water triggers large scale ground subsidence across the North China Plain. Currently, ground subsidence has been recognized as the most serious disaster hindering the sustainable development of this region. In 2012, the central government announced a 10-year plan (2012–2020 plan) swore to take measures fighting against ground subsidence in several main plains across the country including the North China Plain. Establishing an effective monitoring network over subsided areas was listed as a fundamental task of the plan.

Traditionally, the surveying and mapping agencies of Beijing and Tianjin have conducted spirit leveling once per year routinely since 1990s in some areas of Beijing and Tianjin. Therefore, ground subsidence information at some sparse leveling points has been obtained, which forms the basis of official reports on ground subsidence in Beijing and Tianjin. However, such a regular leveling campaign has never been conducted on a large scale in Hebei province, because the cost would be unaffordable given the area of more than 180,000 km² of Hebei province. The main drawback of the leveling method is that it cannot achieve dense spatial sampling due to cost limitations. For example, the area of Tianjin municipality is 11,946 km², while the amount of leveling points routinely measured every year is less than 1200, equivalent to about one leveling point within every 10 km². Therefore, ground subsidence information generated from leveling measurement is incomplete, thus having limited usefulness.

Interferometric SAR (InSAR) was considered an advantageous technique to monitor ground deformation over a large area in comparison with traditional ground-based methods when it emerged in late 1980s [3], but it did not become practically reliable until several kinds of advanced time series techniques came out around the end of last century, such as the permanent scatterer InSAR (PS-InSAR) [4–6] or persistent scattered InSAR [7], and the SBAS InSAR [8]. Since then, time series InSAR techniques have been widely applied to monitor various deformations across the geosciences community, including those associated with volcanic eruptions [9], earthquakes [10], landslides [11–13], and urban subsidence [14]. In China, given the significance of ground subsidence in the Beijing-Tianjin-Hebei region, time-series InSAR techniques have been employed to monitor ground subsidence in different cities, such as in Beijing [15–17], Tianjin [18–20], and Langfang [21] and Cangzhou [22] of Hebei province. All these studies concern subsidence over relatively small areas which can be covered by one SAR frame, and there is a lack of complete information on ground subsidence over the entire Beijing-Tianjin-Hebei region. In this paper, we report a research on monitoring ground subsidence using an improved time series InSAR technique for the wide Beijing-Tianjin-Hebei region over a 22-year time span dating back to 1992 when the European Remote Sensing (ERS)-1 SAR images became available. This research represents a first-ever effort on mapping ground subsidence over such a wide area in China. It involves not only technical innovations, but also significant output which can be used directly for decision-making on subsidence mitigation in the Beijing-Tianjin-Hebei region.

The paper is organized as follows. Section 2 describes the study area and the SAR images. The processing method is introduced in Section 3. The mosaics of ground subsidence rate during the three periods are presented in Section 4. The accuracy validation is also introduced in this section. In Section 5, the spatial-temporal characteristics revealed by the sequential subsidence rate over Beijing and Tianjin are described. Finally, conclusions and open problems are addressed in Section 6.

2. SAR Image Stacks and Study Area

We aim to investigate the ground subsidence of the Beijing-Tianjin-Hebei region for a period of 22 years. No single SAR system can provide images over such a long epoch. SAR images from four satellites are used: ERS-1/2 SAR images acquired between 1992 and 2000; Envisat ASAR images acquired between 2003 and 2010 and RADARSAT-2 images acquired between 2012 and 2014. Unfortunately, there are still three-year and two-year gaps between the sequential SAR datasets.
The plain area of the Beijing-Tianjin-Hebei region is targeted, because it has been confirmed that groundwater over-withdrawal is the main driving force of ground subsidence in this region and groundwater over-withdrawal mostly occurs in plain areas. Unfortunately, we could not obtain ERS and ENVISAT SAR images covering the south part of Hebei province. Therefore, for the period of 1992 to 2010, SAR images cover only the plain areas of Beijing, Tianjin, and Tangshan, Baoding and Langfang of Hebei province. For the period of 2012 to 2014, the whole plain area of the Beijing-Tianjin-Hebei region is covered by eight RADARSAT-2 wide mode frames. The study area and SAR coverages are shown in Figure 1. The number of SAR images over each SAR frame, together with the acquisition time span of the time series of images, are listed in Tables 1 and 2. We can see, the image stacks of ERS-1/2 are generally small, with the stack size less than 13 over 4 frames. The inadequacy causes difficulties in data processing and may result in less accurate deformation retrieval.

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Location of the study area and coverage of SAR data for subsidence monitoring in the Beijing-Tianjin-Hebei region. A blue rectangle stands for the coverage of one ERS/ENVISAT SAR scene, and a green rectangle for one RADARSAT-2 SAR image. The background is SRTM DEM. The red lines denote provincial boundaries, and the black lines are municipal boundaries.

**Table 1.** Image numbers and acquisition time span of ERS-1/2 SAR and ENVISAT ASAR stacks.

<table>
<thead>
<tr>
<th>Track/Frame</th>
<th>ERS-1/2 SAR Stacks</th>
<th>ENVISAT ASAR Stacks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stack Size</td>
<td>Acquisition Time Span</td>
</tr>
</tbody>
</table>
Table 2. Image numbers and acquisition time span of RADARSAT-2 stacks.

<table>
<thead>
<tr>
<th>Frame No.</th>
<th>Stack Size</th>
<th>Acquisition Time Span</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19</td>
<td>4 February 2012–11 July 2014</td>
</tr>
<tr>
<td>2</td>
<td>19</td>
<td>4 February 2012–11 July 2014</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>28 January 2012–4 July 2014</td>
</tr>
<tr>
<td>4</td>
<td>19</td>
<td>28 January 2012–4 July 2014</td>
</tr>
<tr>
<td>5</td>
<td>19</td>
<td>28 January 2012–4 July 2014</td>
</tr>
<tr>
<td>6</td>
<td>19</td>
<td>14 February 2012–27 June 2014</td>
</tr>
<tr>
<td>7</td>
<td>19</td>
<td>14 February 2012–27 June 2014</td>
</tr>
<tr>
<td>8</td>
<td>19</td>
<td>14 February 2012–27 June 2014</td>
</tr>
</tbody>
</table>

3. Methodology of Data Processing

An improved multiple InSAR method, which is entitled as “Multiple-master Coherent Target Small-Baseline InSAR” (MCTSB-InSAR), is developed to process time-series SAR images over a large area. The basic algorithms of MCTSB-InSAR are based on some classical multiple InSAR methods, like PS-InSAR [5,6], that presented in [23] and the PSP method in [24]. The general processing flowchart for deformation mapping over large area using the MCTSB-InSAR method is shown in Figure 2. When dealing with large area consisting of several SAR frames, the deformation over a single SAR frame is estimated first. Then, an adjustment process is employed to harmonize the deformation of overlapping frames, so that a deformation mosaic of the whole area can be generated. The adjustment process will be described in detail later. Some key steps of the method are overviewed in the paragraphs followed.

In the step of interferogram formation, interferograms with spatial baseline and temporal separation within specified thresholds are selected, not limited to one common master image. This is helpful when dealing with small dataset, like the ERS-1/2 stacks over those four frames in this research. The coherent targets include two kinds of pixels. The first type is those pixels characterized by point scatterers whose amplitude keeps stable over temporal acquisitions. These pixels are extracted by thresholding the dispersion index defined by Ferretti et al. [6]. The other type is those pixels associated with stable distributed targets which can be extracted by checking the mean coherence of the formed interferograms [23]. These two groups of pixels are then united and form the coherent target candidates over which the rest processing is conducted.

Similar with the PS-InSAR and the PSP method [24], the phase difference between two neighboring coherent targets are considered in MCTSB-InSAR. Therefore the coherent targets have to be networked firstly. The Delaunay triangulation has been widely employed in many multiple InSAR processing. However, the network generated by Delaunay triangulation is not robust [25]. MCTSB-InSAR adopts the local Delaunay triangulation proposed by Zhang et al. [26] which places small overlapping and regular patches over the entire study area, and then connects the coherent targets within each patch using Delaunay triangulation. The local Delaunay triangulation performs well in balancing the robustness and redundancy of connections. The final density of the arcs generated by local Delaunay triangulation can be controlled by setting the size of the patch and the overlapping rate of two adjoining patches. Figure 3 shows the performance of networking with Delaunay triangulation and with local Delaunay triangulation over an image subset consisting of 640 pixels in a row (about 12.8 km) and 600 pixels in a column (about 12.0 km) and 3363 coherent targets. The number of arcs generated by Delaunay triangulation is 9959, sparser than any of the three connections from local Delaunay triangulation. Generally, we will obtain dense connections if the patch size is small and the overlapping rate is high. Of course, to keep atmospheric phase cancelable along an arc, only arcs shorter than a threshold (1.2 km in this research) are kept.
Figure 2. Processing flow chart of mapping ground deformation over large area using the MCTSB-InSAR method.
Figure 3. Comparison of networking with Delaunay triangulation and local Delaunay triangulation. (a) is the network from Delaunay triangulation, which has 9959 arcs; (b) is from local Delaunay triangulation with patch size of 100 × 100 pixels and overlapping rate of 75%, and has 13,306 arcs; (c) is from local Delaunay triangulation with patch size of 100 × 100 pixels and overlapping rate of 50%, and has 11,375 arcs; (d) is from local Delaunay triangulation with patch size of 150 × 150 pixels and overlapping rate of 50%, and has 11,104 arcs.

Currently, the linear deformation model is used in MCTSB-InSAR for large area monitoring, that is, the main part of deformation is represented by a linear process with fixed rate, and the residual deformation is expressed as a non-linear process. More complicated deformation models such as high-order polynomial might be applicable for small region, but will cause problems in generating a mosaic of derived deformation for a large area consisting of several SAR frames. Under linear deformation assumption, the interferometric phase of the $i$th coherent target in the $k$th interferogram (The interferogram here is a differential interferogram with topographic phase having been subtracted using a DEM) is

$$\phi_i^k = \text{wrap}\{a_1 B_{i,\perp}^k \cdot \Delta h_i + a_2 t_i^k \cdot v_i + \phi_{i,\text{non-linear}}^k + \phi_{i,\text{atmo}}^k + \phi_{i,\text{noise}}^k\}$$  \hspace{1cm} (1)$$

where $\text{wrap}\{\}$ denotes phase wrapping operator, $B_{i,\perp}^k$ and $t_i^k$ are the perpendicular baseline and the temporal separation of the interferogram respectively, $\Delta h_i$ is the DEM error, $v_i$ is the linear deformation rate, $\phi_{i,\text{non-linear}}^k$ is the phase associated with non-linear deformation, $\phi_{i,\text{atmo}}^k$ is the phase caused by atmospheric artifacts, and $\phi_{i,\text{noise}}^k$ the phase accounting for noise, $a_1$ and $a_2$ are known coefficients. Consequently, the phase difference between two neighboring coherent targets along an arc can be expressed as

$$\Delta \phi_{i,j}^k = \text{wrap}\{a_1 B_{i,\perp}^k \cdot \Delta h_{i,j} + a_2 t_i^k \cdot \Delta v_{i,j} + \Delta \phi_{i,j,\text{res}}^k\}$$  \hspace{1cm} (2)$$

where $\Delta h_{i,j}$ and $\Delta v_{i,j}$ denote the difference of deformation rate and difference of DEM error between the two targets $i$ and $j$, the residual phase $\Delta \phi_{i,j,\text{res}}^k$ is the summation of atmospheric phase difference,
noise phase difference and non-linear phase difference over the two targets. For each arc, we will form a set of equations like (2) with each corresponding to an interferogram. The unknown variables of the equation set, namely $\Delta h_{ij}$ and $\Delta v_{ij}$, can be solved using the periodogram [3] or 2-D searching algorithm [27]. After all arcs in the network are processed, the absolute values of the linear deformation rate $v$ and DEM error $\Delta h$ for each coherent target can be calculated through an integration process.

After the deformation rate over every single SAR frame has been estimated, an adjustment process is employed to harmonize the deformation estimates from overlapping frames, which is a unique feature of the MCTSB-InSAR method, and has seldom been discussed by other multiple InSAR methods yet. Here only the adjustment of deformation rate is concerned. Accumulative deformations are difficult to be mosaiced, because the time span of SAR images over different tracks are usually different. Before the adjustment, the corresponding coherent targets from overlapping frames have to be recognized. This can be done easily on the basis of geocoded images. The underlying principle is that the geographic coordinates of two corresponding coherent targets should be similar, that is, the difference of their coordinates should be smaller than the grid size of geocoding. The adjustment is to estimate an offset adding to the calculated deformation rate for a frame, so that the adjusted deformation rate over corresponding coherent targets within overlapped area of neighboring frames is identical. This forms the first category of adjustment equations, which has the following form

$$v_L(x, y) + \Delta L = v_R(x, y) + \Delta R$$

where $v_L$ and $v_R$ are InSAR-derived deformation rate over a corresponding coherence target at position $(x, y)$ from overlapped frame $L$ and $R$, $\Delta L$ and $\Delta R$ are the unknown offset of deformation rate. On the other hand, the deformation rate calculated from leveling and/or GPS measurements can be used as reference of the adjustment. Especially, the adjusted deformation rate should be equal to that calculated from the reference data, which forms the second category of adjustment equations and has the following form

$$v_L(x, y) + \Delta L = v_{ref}(x, y)$$

where $v_{ref}(x, y)$ denotes deformation rate calculated from leveling and/or GPS measurements. The two groups of equations join together and form an over-determined linear system with the offset attached to every overlapped frame as unknown variables. Obviously, these unknowns can be solved under the least mean square criterion. Finally, a deformation rate mosaic is formed by averaging the adjusted deformation rates of overlapping frames. The adjustment process is powerful in large area monitoring. It provides an integrated solution for both calibration of InSAR-derived deformation rate for every frame and the harmonization of the deformation rates from overlapping frames.

Once the linear deformation rate is determined, the nonlinear deformation is estimated by temporal and spatial filtering of the residual phase. More information about nonlinear deformation estimation can be found in [5,23].

4. Result and Accuracy

The three sets of SAR stacks are processed and the deformation of the study area over the three time spans is calculated, respectively. ERS-1/2 and ENVISAT SAR images are multilooked with a pixel size on the ground of 40 m. RADARSAT-2 images are multilooked, with a pixel size of 25 m on the ground. The numbers of final coherent targets over the study area are 518,150 in ERS-1/2 images, 982,045 in ENVISAT images, and 4,906,691 in RADARSAT-2 images. The deformation rate mosaic of the three time spans is shown in Figures 4–6, respectively. Please note the deformation rate in SAR line-of-sight (LOS) direction has been projected into the up-down direction as subsidence rate according to Equation (5):

$$v_{u-d}(x, y) = v_{LOS}(x, y) / \cos \theta_{(x,y)}$$
where, \( v_{u-d}(x, y) \) and \( v_{LOS}(x, y) \) are the up-down subsidence rate and SAR LOS deformation rate respectively, \( \theta_{(x,y)} \) is SAR viewing angle at position \((x, y)\). Of course, Equation (5) holds only under an ideal condition that ground uplift or subsidence is the sole deformation. For this study, this assumption is reasonable, because it has been widely recognized that ground subsidence and uplift are the dominant ground deformations in the North China plain [1]. Another reason for doing so is to facilitate the generation of the subsidence rate mosaic for the wide Beijing-Tianjin-Hebei region. The study area is covered by SAR images from at least three tracks, and the SAR LOS varies with different tracks. Therefore, in order to generate a mosaic of the deformation rate, we have to choose a SAR LOS as the reference direction, and project the deformation rate of other LOS into this direction. This process also involves some assumptions. In conclusion, projecting the deformation rate of all SAR LOSs into the up-down direction may be a good solution. In the figures, positive value means uplift, and negative value means subsidence. We have to admit that the subsidence rate estimated from ERS-1/2 SAR stacks over some frames is not very accurate, which cause obvious inconsistency between neighboring frames, like the frame 2817 and 2799 of track 175 as shown in Figure 4. Also from Figure 4, there are some noisy points showing uplift at the rate of 6–15 mm/year, which is inconsistent with the neighboring subsidence. As mentioned in Section 2, the relatively small size of the image stack and the heterogeneous spatial and temporal baseline of ERS-1/2 images may result in a less accurate deformation estimate, thus producing inconsistency in the mosaic. In contrast, the subsidence rate mosaic generated from ENVISAT and RADARSAT-2 dataset shows very good consistency, which indirectly indicates that the deformation rate estimates are accurate. You may notice there are obvious uplifts in the south-western mountainous areas in Figure 6. This may be caused by some topography induced atmospheric artifacts, or real ground uplift. Unfortunately, there is no leveling measurements over there. Further research is needed to investigate the uplift over the mountainous areas.

**Figure 4.** Linear subsidence rate over the Beijing-Tianjin-Hebei region from 1992 to 2000 observed by time series ERS-1/2 SAR images. The red lines denote provincial boundaries, and the black lines are municipal boundaries.
Figure 5. Linear subsidence rate over the Beijing-Tianjin-Hebei region from 2003 to 2010 observed by time series ENVISAT ASAR images.

Figure 6. Linear subsidence rate over the Beijing-Tianjin-Hebei region from 2012 to 2014 observed by time series RADARSAT-2 images.
We have collected subsidence measurements over 204 leveling points from the surveying and mapping authority of Tianjin, which were measured every October from 1992 to 2014, and subsidence measurements over 169 leveling points measured annually through 2005 to 2012, and 34 leveling points measured annually through 2012 to 2014 from the surveying and mapping authority of Beijing. Of course, not all of these leveling points are usable, because there may not be a coherent InSAR target in the vicinity of a leveling point. Only those leveling points in a circle neighborhood with 80-m radius centered where there is at least one coherent target are used. Generally, around one third of those usable leveling measurements are used as reference in the adjustment process, and the rest are used to evaluate the accuracy of InSAR-derived subsidence rate. For example, totally 211 leveling points are used in the processing of RADARSAT-2 dataset. Out of them, 70 points are used as the references of adjustment, and 141 points are used in accuracy evaluation. The locations of the 211 leveling points are plotted in Figure 7a, where the blue points are for adjustment reference, and the red points for accuracy evaluation. The statistics of the evaluation result are listed in Table 3, including the standard deviation of the difference between subsidence rate derived from InSAR and leveling measurement, and the maximum and minimum values of the difference. The differences between the subsidence rate derived from RADARSAT-2 dataset and that from leveling measurement over the 141 red points in Figure 7a are shown in Figure 7b. The standard deviation in Table 3 indicates that the subsidence rate derived from ASAR and RADARSAT-2 images have an accuracy of 4.7 mm/year and 5 mm/year respectively, while the subsidence rate derived from ERS-1/2 images has an accuracy of 8.7 mm/year, not as good as the other two. Of course, the leveling measurements are back-projected into the SAR LOS direction in the accuracy evaluation in some publications. From Equation (5), it can be seen that the evaluation result of these two processes is linearly correlated. In this case, you may obtain an little better accuracy result if the leveling measurements are back-projected into SAR LOS, as the scaling factor $\cos \theta(x, y)$ tends to reduce the values of deformation. The evaluation results in Table 3 agree with the visual quality of the three mosaics of subsidence rate shown in Figures 4–6. This also confirms that the deformation rate derived from the three SAR dataset is generally accurate and the MCTSB-InSAR method is reliable.

### Table 3. Statistics of accuracy evaluation of InSAR-derived subsidence rate in three periods (mm/year).

<table>
<thead>
<tr>
<th>Time Spans</th>
<th>Number of Used Leveling Points</th>
<th>Maximum Difference</th>
<th>Minimum Difference</th>
<th>Standard Deviation of Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003-2010</td>
<td>166</td>
<td>11</td>
<td>–14</td>
<td>4.7</td>
</tr>
<tr>
<td>2012-2014</td>
<td>141</td>
<td>11</td>
<td>–16</td>
<td>5.4</td>
</tr>
</tbody>
</table>

Besides the subsidence rate, the accumulative subsidence over individual coherent targets can also be estimated from one SAR stack. As mentioned earlier, there are two-to-three year gaps between the three sequential SAR stacks, therefore we can only obtain three discrete time series of accumulative subsidence for the period of 1992 to 2014. However, by means of the continuous annual leveling measurements over some parts of the region, such as Tianjin, the three discrete time series of accumulative subsidence can be linked together to form a continuous time series of ground subsidence. Figure 8 shows the time series of subsidence over four coherent targets in Tianjin, together with the subsidence series from the nearest leveling point. The positions of the four coherent targets are marked in Figure 7b. We can see that the time series of InSAR-derived subsidence fit well with the subsidence evolution recorded by leveling measurements, which confirms the accuracy of the ground subsidence mapped by InSAR.
Figure 7. Locations of the 211 leveling points used for adjustment reference and accuracy evaluation of the subsidence rate generated from the RADARSAT-2 dataset (a): Blue points are for adjustment reference, and red points for accuracy evaluation. The errors of InSAR derived subsidence rate over the 141 red points are shown in (b). The four triangles in (b) mark the locations of the four coherent targets whose subsidence evolution is shown in Figure 8.
Remote Sens. 2016, 8, 675

Figure 8. (a–d) The time series of accumulative subsidence over four coherent targets in Tianjin, together with the subsidence series of the nearby leveling point.

5. Spatial—Temporal Variation of Ground Subsidence in Beijing and Tianjin

The subsidence rate of the Beijing plain area over the three time spans is shown in Figure 9. During the time span from 1992 to 2000, most of the Beijing plain is stable. Some small scale subsidence took place in Haidian, Changping, Chaoyang, and Mentougou districts. The maximum subsidence rate is 48 mm/year in this period which occurred at a mining site in Mentougou district marked by the redline circle in Figure 9b. The coal exploited in this mine had been the main energy source for heating Beijing during the winter for decades. The mine site was shut down in 2000. From Figure 9c, the previous subsiding sites in the Chaoyang district had evolved as large subsiding zones, with a maximum rate reaching 143 mm/year. Subsiding zones also existed in the north part of the Beijing plain, including Haidian, Changping, Shunyi, and Pinggu districts. To count the subsiding area, the subsidence rate
over non-coherent pixels is estimated using the simple Kriging interpolation method with the normal score transformation and spherical semivariogram model in ArcGIS software. Totally, the subsiding area with a subsidence rate of more than 50 mm/year was 265.41 km². A noticeable change is that the subsidence at the mining site in the Mentougou district disappeared, which turned out to be the result of the shutdown. During the period from 2012 to 2014 as shown in Figure 9d, the subsidence in the Beijing plain had continuously developed, with 433.25 km² area subsiding at a rate more than 50 mm/year. The subsiding zones of Chaoyang, Tongzhou, Haidian, Changping, and Shunyi districts kept growing and became geographically joined together. The maximum subsidence rate was 152 mm/year, even higher than that of the previous time span. Generally, during the last two decades, ground subsidence mainly occurred in the east and north part of the Beijing plain, with the east rim of Chaoyang district and the northwest part of Tongzhou district suffering the most severe subsidence. The urban center consisting of Xicheng and Dongcheng districts and the western part of Beijing kept relatively stable with little subsidence. More analysis on the relationship between the temporal-spatial variations of ground subsidence and the groundwater exploration and the hydro-geological environment in Beijing is needed in the future.

The districts and counties of Beijing municipality are shown as polygons in (a) with their names embedded. The redline rectangle in (a) corresponds to the area shown in (b–d). A mining site in Mentougou district marked by the redline cycle in (b) had the largest subsidence rate from 1992 to 2000.

Figure 9. (a–d) Subsidence rate in three stages in Beijing. (b) 1992–2000; (c) 2003–2010; (d) 2012–2014. The districts and counties of Beijing municipality are shown as polygons in (a) with their names embedded. The redline rectangle in (a) corresponds to the area shown in (b–d). A mining site in Mentougou district marked by the redline cycle in (b) had the largest subsidence rate from 1992 to 2000.

The spatial-temporal characteristic of subsidence in Tianjin is a little different from that of Beijing. The maps of subsidence rate of Tianjin in the three time spans are shown in Figure 10. During the time span from 1992 to 2000, ground subsidence took place in nearly all the suburb districts surrounding the urban center, including Wuqing, Beichen, Xiqing, Jinghai, Dongli, Jinnan, Tanggu, and Hangu districts. The majority of the subsidence rate was in the range of 30 mm/year to 60 mm/year. The maximum
The subsidence rate was 81 mm/year. The area with a subsidence rate more than 50 mm/year was 85.73 km², in comparison to none in Beijing at that period. During the period from 2003 to 2010, ground subsidence had developed very quickly. In the north, west, and south of the urban center, ground subsidence had expanded dramatically, especially in districts of Beichen, Xiqing, Jinghai, Dongli, and Jinnan as shown in Figure 10c. Only in the coastal area of Hangu district was ground subsidence relieved somehow. The maximum subsidence rate of Tianjin at this period reached 135 mm/year. The area with a subsiding rate more than 50 mm/year was 2749.1 km², 10 times of that in Beijing. During the period from 2012 to 2014, ground subsidence in Tianjin presented different scenario over the north part and south part which was approximately marked by the white solid line in Figure 10c,d. In the northern part, ground subsidence was exacerbated, especially in districts of Beichen and Wuqing. The subsided area became wider, and the subsidence rate became larger. The largest subsidence occurred at the town of Wangqingtuo of the Wuqing district reaching 153 mm/year. In the south part of Tianjin, ground subsidence showed substantial alleviation, which might be the result of control of groundwater exploration in recent years in this region. As a comparison, the area with a subsiding rate more than 50 mm/year decreased to 1117.55 km² in the period of 2012 to 2014, roughly 41% of the area in previous period. In summary, ground subsidence in Tianjin had already been serious in 1990s, had dramatically expanded during 2000s, and started to alleviate in recent years.

Figure 10. (a–d) Subsidence rate in the three periods in Tianjin; (b) 1992–2000; (c) 2003–2010; (d) 2012–2014. The districts and counties of Tianjin municipality are shown as polygons in (a) with the names on.
6. Conclusions

An application using multi-temporal InSAR technology to map ground subsidence over the wide Beijing-Tianjin-Hebei region is presented. Three sets of SAR images coming from four satellites are employed to cover a long time span from 1992 to 2014. An improved time series InSAR methodology, namely MCTSB-InSAR, has been developed to deal with small datasets consisting of only 11 to 12 images and to monitor deformation over wide area equivalent to the coverage of several SAR frames. The adjustment process of MCTSB-InSAR provides an integrated solution for both calibration of InSAR-derived deformation and the harmonization of the deformation estimates from overlapping frames. Of course, some aspects of MCTSB-InSAR need to be improved further. For example, the currently used linear deformation model may not fit a situation with strong non-linear deformation. More complicated deformation models, such as high-order polynomial and sinusoid model as presented in [28] should be incorporated into MCTSB-InSAR in the future.

Three subsidence rate maps over the Beijing-Tianjin-Hebei region on three time spans are generated from SAR images acquired by ERS-1/2, ENVISAT, and RADARSAT-2 satellites. Compared with more than 120 leveling measurements collected in Beijing and Tianjin, the subsidence rate maps have a standard deviation of 8.7 mm/year (1992–2000), 4.7 mm/year (2003–2010), and 5.4 mm/year (2012–2014) respectively. The comparison validates the accuracy the subsidence information derived from InSAR, and confirms the reliability and robustness of the MCTSB-InSAR method.

During the 22-year time span, the evolution of ground subsidence in Beijing and Tianjin shows different scenarios. In the period from 1992 to 2000, there were only some sites with small-scale subsidence mainly scattered in the eastern and northern suburbs of Beijing, with the largest rate being 48 mm/year; while in Tianjin, serious ground subsidence already took place in nearly all the suburb districts with 85.73 km² area subsiding at a rate more than 50 mm/year. In the period of 2003 to 2010, ground subsidence had expanded dramatically in both Beijing and Tianjin. In the period of 2012 to 2014, ground subsidence in Beijing had continuously developed, with the area subsiding at a rate more than 50 mm/year increasing 63% than that of the previous period; while the development of ground subsidence in Tianjin had slowed down with the area subsiding at a rate more than 50 mm/year decreasing 41% than that of the previous period. The different spatial-temporal characteristics of subsidence evolution in Beijing and Tianjin surely were related to the exploration of groundwater, precipitation, hydro-geological environment, and others. How these factors affected ground subsidence need further analysis in the future.

The output of the research provides first-hand, relatively complete and accurate information about ground subsidence over the Beijing-Tianjin-Hebei region. This information could directly serve the decision-making on ground subsidence mitigation and water resource management in the Beijing-Tianjin-Hebei region.

Acknowledgments: This research was supported by National Program on Key Basic Research Project (No. 2012CB719905) and National Natural Science Foundation of China (No. 41271430, No. 41304010). ESA is gratefully acknowledged for providing ERS-1/2 SAR and ENVISAT ASAR data to support this research. We also thank USGS for supporting the SRTM DEM used in this paper.

Author Contributions: Yonghong Zhang led this research, designed the framework of data processing, and developed the adjustment algorithm. Hong’an Wu contributed to algorithm programming, the Radarsat-2 data processing and accuracy evaluation. Yonghui Kang processed the ERS-1/2 SAR data and contributed to accuracy evaluation. Chuanguang Zhu contributed to the ENVISAT ASAR data processing.

Conflicts of Interest: The authors declare no conflict of interest.

References


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
Spatiotemporal Characterization of Land Subsidence and Uplift (2009–2010) over Wuhan in Central China Revealed by TerraSAR-X InSAR Analysis

Lin Bai 1,2, Liming Jiang 1,*, Hansheng Wang 1 and Qishi Sun 1,2

1 State Key Laboratory of Geodesy and Earth’s Dynamics, Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan 430077, China; bailin112@mails.ucas.ac.cn (L.B.); whs@whigg.ac.cn (H.W.); sunqishi14@mails.ucas.ac.cn (Q.S.)
2 University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: jlm@whigg.ac.cn; Tel.: +86-027-8677-8612

Academic Editors: Zhenhong Li, Roberto Tomas, Zhong Lu and Prasad S. Thenkabail
Received: 10 March 2016; Accepted: 14 April 2016; Published: 20 April 2016

Abstract: The effects of ground deformation pose a significant geo-hazard to the environment and infrastructure in Wuhan, the most populous city in Central China, at the intersection of the Yangtze and Han rivers. Prior to this study, however, rates and patterns of region-wide ground deformation in Wuhan were little known. Here we employ multi-temporal SAR interferometry to detect and characterize spatiotemporal variations of ground deformation in major metropolitan areas in Wuhan. A total of twelve TerraSAR-X images acquired during 2009–2010 are used in the InSAR time series analysis. InSAR-derived results are validated by levelling survey measurements and reveal a distinct subsidence pattern within six zones in major commercial and industrial areas, with a maximum subsidence rate up to −67.3 mm/year. A comparison analysis between subsiding patterns and urban developments as well as geological conditions suggests that land subsidence in Wuhan is mainly attributed to anthropogenic activities, natural compaction of soft soil, and karst dissolution of subsurface carbonate rocks. However, anthropogenic activities related to intensive municipal construction and industrial production have more significant impacts on the measured subsidence than natural factors. Moreover, remarkable signals of secular land uplift are found along both banks of the Yangtze River, especially along the southern bank, with deformation rates ranging mostly from +5 mm/year to +17.5 mm/year. A strong temporal correlation is highlighted between the detected displacement evolutions and the water level records of the Yangtze River, inferring that this previously unknown deformation phenomenon is likely related to seasonal fluctuations in water levels of the Yangtze River.

Keywords: multi-temporal InSAR; ground deformation; Wuhan city; Urban development; karst geology

1. Introduction

Accompanying large-scale urbanization and industrialization during the past 35 years, more than 95 megacities in China have undergone rapid land subsidence [1]. The potential consequences of land subsidence mainly include degradation of the aquifer system and damage to the utility infrastructures, buildings, railroads, highways and bridges [2,3]. In this context, intensive efforts and investigations related to land subsidence monitoring have been undertaken in four major subsiding regions in China: the Yangtze River Delta [4,5], the North China Plain [6,7], the Fenwei Basin [8] and the Pearl Delta [9]. As one of the most populous cities in China, Wuhan city has experienced a rapid urban expansion over the past decades. Consequently, various types of geohazards relevant to ground deformation have been frequently occurring in Wuhan metropolitan areas, due to anthropogenic activities (e.g.,
Remote Sens. 2016, 8, 350
groundwater pumping, tunnel excavation) and/or natural geological processes (e.g., alluvial soil consolidation, karst collapse and surface water loading) [10,11]. Prior to this study, however, rates and patterns of region-wide ground deformation in Wuhan were still little known. Consequently, there is an increasing demand for regular and large-scale monitoring of ground deformation to support the integrated sustainable development of Wuhan.

Over past decades, space-borne Synthetic Aperture Radar Interferometry (InSAR) has been proven as an effective remote sensing technique to detect ground deformation associated with earthquakes [12,13], volcano eruptions [14,15], permafrost degradation [16], glacier movement [17], landslides activity [18,19] and subsidence investigation [20–22]. Compared with point-measurement geodetic techniques such as Global Positioning System (GPS) and levelling surveys, InSAR provides deformation measurements at a significantly improved resolution over large areas. However, it is still a challenge as a practical tool for monitoring subtle ground deformation due to temporal and spatial decorrelation as well as atmospheric disturbances. To overcome the limitations of conventional InSAR methods, in recent years some advanced InSAR approaches have been developed based on time-series interferometric analysis of multi-temporal SAR acquisitions over the same areas, herein referred to as multi-temporal InSAR (MT-InSAR) techniques. The MT-InSAR techniques, mainly involving persistent scatterer InSAR (PSI) methods [23–25] and small baseline subset (SBAS) methods [26–28], identify and exploit a subset of image pixels which maintain a high coherence level over the study period, allowing the average rates and temporal evolutions of ground deformation to be estimated with millimeter-level accuracy. These advanced InSAR techniques have been widely used in monitoring urban ground deformation, such as in Macao [3], Shanghai [29], Guangzhou [30], Lisbon [31] and Mexico [32].

In this paper, we present results from the first application of multi-temporal InSAR in ground deformation monitoring in Wuhan. The StaMPS (Stanford Method for Persistent Scatterer) approach [24,33] is applied to 12 TerraSAR-X images acquired from October 2009 to August 2010 to retrieve and characterize spatiotemporal variations of ground deformation in major metropolitan areas in Wuhan. Moreover, a comparison between InSAR-based results and in situ data is carried out to validate the InSAR observations. Finally, the potential causes of the observed ground deformation are discussed through an integrated analysis of multidisciplinary information related to karst geology, surface hydrology, urban development and geotechnical engineering.

2. Study Area and Data Used

2.1. Geological Setting of Study Area

Wuhan, the capital of Hubei province and the most populous city in Central China, lies in the eastern Jianghan Plain at the intersection of the middle reaches of the Yangtze and Han rivers. The city contains many lakes and parks, including expansive East Lake, the largest urban lake in China. The geology of Wuhan is part of the Yangtze geosyncline sag with the characteristics of paraplatform, and developed well in an alluvial plain with monadnock. Climatically, Wuhan is in a subtropical monsoon climate zone, characterized by high temperature in summer, low temperature in winter and abundant precipitation. The average annual precipitation reaches up to 1261.2 mm, concentrated in the flood season from June to August [34].

The study area in this work is located at major metropolitan areas in Wuhan, with a total area of approximately 28 km by 20.5 km, and mainly contains Hankou and Wuchang parts (Figure 1). The Hankou region is characterized with an alluvial plain formed by flood siltation of Yangtze and Han rivers as well as lakes, and the Wuchang region is denuded hilly land except the alluvial plain along the Yangtze River. Soft soils are widely distributed throughout most areas of Hankou and in the riverside area of Wuchang. There are one or more layers of lens formed by silt clay and silt in the area 5~30 m beneath the ground surface, and the thickness of the lens reaches about 10~20 m. The materials of the lens, presenting a plastic flow state and strong hydrophilic and alteration features, easily lead to ground inhomogeneous subsidence under the influence of pumping groundwater and vibration [11].
In addition, there are insidious carbonate rocks in the second bottom along Yangtze River and the south of Changjiang Bridge and well-developed karst geology in the Wuhan urban areas (see Figure 1). A large number of karst collapses caused by human activities, such as pumping groundwater for water supply and mine dewatering, have been recorded since 1931 and pose a serious threat to buildings and human safety [35].

Figure 1. The geographic location and simplified geological setting of the study area. The distributions of carbonate belts are redrawn from Luo [36]. L1, L2 and L3 represent three major carbonate rock belts named Tianxingzhou, Daqiao and Baishazhou, respectively. The red rectangle is the study area’s extent. The red triangle represents the location of the levelling points, situated around Xunlimen Station in Hankou district and Mingdu Station in Hongshan district.

2.2. Datasets Used

In this study, 12 TerraSAR-X scenes at HH polarization, acquired between October 2009 and August 2010 along ascending track 142, are used to investigate the ground deformation in Wuhan. More parameters of the TerraSAR-X data are summarized in Table 1. The coverage of the TerraSAR scenes is shown as the red rectangle in Figure 1.
Table 1. The parameters of the TerraSAR-X data.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Track no.</td>
<td>142</td>
</tr>
<tr>
<td>Imaging Modes</td>
<td>StripMap</td>
</tr>
<tr>
<td>Polarization</td>
<td>HH</td>
</tr>
<tr>
<td>Orbit direction</td>
<td>Ascending</td>
</tr>
<tr>
<td>Looking direction</td>
<td>Right</td>
</tr>
<tr>
<td>Central incidence angle (degree)</td>
<td>34.9</td>
</tr>
<tr>
<td>Range resolution (m)</td>
<td>2.0</td>
</tr>
<tr>
<td>Azimuth resolution (m)</td>
<td>3.3</td>
</tr>
<tr>
<td>No. of images</td>
<td>12</td>
</tr>
<tr>
<td>Date of earliest image used</td>
<td>7 October 2009</td>
</tr>
<tr>
<td>Date of latest image used</td>
<td>11 August 2010</td>
</tr>
</tbody>
</table>

SRTM (Shuttle Radar Topography Mission) DEM with a resolution of 30 m are used to remove the topographic phase and geocode interferograms. This 30-m DEM was released in 2015 by USGS (U.S. Geological Survey) and can be downloaded at https://lta.cr.usgs.gov/SRTM1Arc. In addition, daily water level data on the Yangtze River acquired by the Hankou hydrological station from October 2009 to October 2010, provided by the Hubei Administration of hydrology and water resources, are used to analyze possible causes for ground uplift along both banks of the Yangtze River.

Finally, in situ measurements collected repeatedly by precise levelling campaigns between 2008 and 2010 are utilized to evaluate the InSAR-derived results. The levelling campaigns were carried out over two construction areas of Metro Line 2 (see Figure 1), situated around Xunlimen Station in Hankou district and Mingdu Station in Hongshan district, respectively. More detailed descriptions on the levelling surveys were reported in [37,38].

3. Multi-Temporal InSAR Data Processing

The StaMPS approach, an implementation of persistent scatterer (PS) Interferometry technique, is employed in this study to carry out multi-temporal InSAR analysis of the 12 TerraSAR-X data. The StaMPS approach uses spatial correlation of the interferometric phase to find PS pixels without prior knowledge of temporal variations in the deformation rate. This strategy can detect low amplitude pixels with phase stability in most terrain types, with or without buildings, and is applicable in areas undergoing temporally variable deformation with no prior knowledge [24]. Due to the large subsidence rates of the study area and the limitation of the quantity of SAR images, we applied the SB (small baseline) method implemented in the StaMPS approach [39] to derive ground deformation. The fundamental procedures of the StaMPS-SB method are briefly introduced below and more detailed descriptions of this method can be consulted in [24,33,39,40].

3.1. Interferogram Formation

The first key step of the StaMPS-SB method is to generate the interferograms. A coregistration algorithm is applied to reduce the coregistration errors owing to long perpendicular baselines [33]. In order to minimize the temporal and spatial decorrelation, interferometric pairs with small temporal and perpendicular baselines are selected to generate interferograms. Decorrelation is further reduced by spectral filtering in range and discarding of the non-overlapping Doppler frequencies in azimuth [39]. Specifically, we imposed 33 interferograms generated from the 12 TerraSAR-X datasets by using DORIS (Delft Object-oriented Radar Interferometric Software), with a perpendicular baseline constraint of 180 m and a maximum temporal baseline of 110 days (Figure 2). After the interferograms generation, we used 30 m resolution SRTM DEM to remove the topographic phase contribution from the interferometric phase and geocode interferograms.
3.2. SDFP Pixel Identification

StaMPS-SB method performs time-series analysis on the pixels whose filtered phase decorrelates little over short time intervals, referred to as slowly-decorrelating filtered phase (SDFP) pixels [39]. For the consideration of the computational cost, initial SDFP pixels candidates are selected by setting a threshold (0.6) for the amplitude difference dispersion, which is the standard deviation of the amplitude difference between the master and slave divided by the mean amplitude [39].

Then, the spatial correlation of interferometric phase analysis is applied to identify SDFP pixels from the candidates. The wrapped phase $\Phi_{int,x,i}$ of the $x$th pixel in the $i$th “flattened” and topographically corrected interferogram can be expressed as the sum of the phase change due to deformation in the satellite line-of-sight (LOS) direction $\Phi_{def,x,i}$, the phase due to the difference in atmospheric retardation between passes $\Phi_{atm,x,i}$, the residual phase due to orbit inaccuracies $\Delta \Phi_{orb,x,i}$, the residual phase due to look angle error $\Delta \Phi_{\theta,x,i}$, and the noise term due to variability in scattering, thermal noise, coregistration errors and uncertainty in the position of the phase center in azimuth $\Phi_{n,x,i}$:

$$\Phi_{int,x,i} = W\left\{ \Phi_{def,x,i} + \Phi_{atm,x,i} + \Delta \Phi_{orb,x,i} + \Delta \Phi_{\theta,x,i} + \Phi_{n,x,i} \right\} \quad (1)$$

where $W\{\cdot\}$ is a wrapping operator. The SDFP pixels are those for which $|\Phi_{n,x,i}|$ is small enough that it does not completely obscure the signal, so $\Phi_{n,x,i}$ is required to estimate accurately to identify SDFP pixels. For this purpose, the first four terms on the right-hand of Equation (1) must be estimated and subtracted from the wrapped phase. A bandpass filtering of surrounding pixels in the frequency domain is adopted to estimate the spatially-correlated look angle (SCLA) error (1) must be estimated and subtracted from the wrapped phase. A bandpass filtering of surrounding pixels in the frequency domain is adopted to estimate the spatially-correlated look angle (SCLA) error, including the phase due to ground deformation ($\Phi_{def,x,i}$), variation in atmospheric delay ($\Phi_{atm,x,i}$), orbital inaccuracies ($\Delta \Phi_{orb,x,i}$) and spatially correlated DEM error ($\Delta \Phi_{\theta,x,i}$). The spatially uncorrelated look angle (SULA) error term ($\Delta \Phi_{\theta,x,i}$) which is mainly caused by spatially uncorrelated DEM error and deviation of the pixel’s phase center from physical center of the backscattering object, is then estimated by an inversion associated with its correlation with perpendicular baseline. Subtracting these estimates from the interferometric phase, we get an estimate of $\Phi_{n,x,i}$ which is then characterized with a measure similar to coherence, referred to as $\gamma_x$, and the candidate pixels with low $\gamma_x$ are rejected. The procedure is adopted iteratively until $\gamma_x$ is converged, and a set of SDFP pixels are finally identified [33,39].
3.3. 3-D Phase Unwrapping and Time-Series Deformation Retrieval

The three-dimensional phase unwrapping algorithm proposed by Hooper [40,41] is applied on the sets of SDFP pixels to recover the unambiguous phase values. In order to unwrap correctly, the spatially uncorrelated contribution needs to be subtracted before unwrapping. After phase unwrapping, spatial and temporal bandpass filtering is applied to remove the spatially correlated contributions. Finally, the displacement time series for each SDFP pixel is obtained by least-squares inversion [39]. Deformation rate and displacement time series for each SDFP pixel are finally calibrated to a stable reference point which corresponds to the WUHN International GNSS Service (IGS) station located in Wuhan University (114.357259°E, 30.531654°N).

4. Results and Interpretations

4.1. InSAR-Derived Results and Validation

A total of more than one million SDFP pixels were ultimately identified and exploited by means of the StaMPS-SB time-series analysis for the 12 TerraSAR-X images, with an average density of ~1933 SDFP pixels/km². Figure 3 illustrates the spatial distribution of the SDFP pixels and the corresponding linear deformation velocity in the radar line-of-sight (LOS) direction. The annual deformation rates of SDFP pixels relative to the reference point range from $-67.3$ mm/year to $+17.5$ mm/year.

Heterogeneous land subsidence patterns are widely found in major urban areas of Wuhan, mostly concentrated in commercial zones in Hankou (Zone 1), industrial zones around Wuhan Iron and Steel Corp. in Qingshan (Zone 3) and the Donghu New Technology Development Zone (Zone 5) in Wuchang, as well as occurring in karst carbonate areas in Wuchang and Qingshan districts (Zone 2, 4 and 6, respectively). Causes for the observed subsidence could be generally divided into two categories of urban development (for Zone 1, 3 and 5) and karst geology (for Zone 2, 4 and 6). More detailed analysis of possible causes for land subsidence in these zones will be presented in next sub-sections. It is worth noting that exceptional land-uplift signals, with deformation rates ranging mostly from $+5$ to $+17.5$ mm/year, are detected along both banks of Yangtze River especially along the southern bank (Zone 7), and also in urban areas located on one sector of a carbonate rock belt between Yangtze River and East Lake (Zone 8). A possible explanation for this anomalous phenomenon in two uplift zones will be discussed in Section 5.

An accuracy assessment of the InSAR-derived deformation results was performed by comparing them against the precise levelling measurements as described in Section 2. To enable the comparison, the levelling measurements have been projected along radar LOS direction and interpolated via a linear fit within the acquisition period of the InSAR observations. A comparison method used in an InSAR validation project of European Space Agency named as PSIC4 [42] was applied to the accuracy assessment. In this study the comparison was carried out between each levelling point and all SDFP pixels located within a circle with 50 m radius around the corresponding levelling point. Results suggest that the InSAR-derived deformation rates agreed well with the levelling survey measurements with an average absolute difference of 3.1 mm/year and standard deviation of 1.8 mm/year.
Figure 3. The mean LOS deformation velocity map over the study area during the period from October 2009 to August 2010. The red star represents the location of the reference point. White rectangles indicate Zone 1, 3 and 5 where land subsidence was mainly caused by urban development. White ellipses indicate Zone 2, 4 and 6 where land subsidence was mainly caused by karst geology. Red dotted polygons indicate Zone 7 and 8 where the uplift occurs. This result is superimposed on a Landsat 8 image.

4.2. Subsidence Caused by Urban Development

As shown in Figure 3, major commercial zones in Hankou, industrial zones around Wuhan Iron and Steel Corp. and the Donghu New Technology Development Zone exhibit wide and continuous coverage of land subsidence, which are most likely attributed to ground water over-exploitation due to growing demand for intensive municipal constructions and industrial production in these regions.

In particular, a distinct subsidence bowl with a maximum subsidence rate of $-67.3 \text{ mm/ year}$ is detected in the central urban area in Hankou (Zone 1 as shown in Figure 3). Thus, we select this subsiding region for a more detailed analysis of the relationship between urban development and resulting land subsidence. Figure 4 illustrates a zoomed deformation rate map of this region (white rectangle 1 in Figure 3). A serious subsidence was detected in these areas in high densities of SDFP pixels from the InSAR results. In Zone 1, extensive urban construction activities were undertaken during the period of SAR imagery acquisitions, typically including the Metro Line 2 tunneling as well as large-scale building constructions such as the outpatient building of Wuhan Union Hospital and Wuhan International Plaza (see Figure 4b). Furthermore, Figure 4c presents the deformation time series corresponding to three pixels (labelled in Figure 4a as Point A, B and C), located on buildings near the
Wuhan Union Hospital, the Wuhan International Conference & Exhibition Center and in the vicinity of levelling benchmarks in the Xunlimen Station of Metro Line 2, respectively. The subsidence of Point B is also testified by an in situ photograph of the building (shown in Figure 4d), and the subsidence rate of Point C consists with measurements of levelling surveys at the Xunlimen Station (−14.7 mm/year vs. −11.6 mm/year).

Figure 4. Mean LOS deformation rate map superimposed on Google Earth image over (a) Zone 1; (b) the area highlighted in Figure 4a (blue box), engineering projects under construction are marked with light green lines; (c) Displacement time series corresponding to the SDFP pixels labeled as Point A, Point B and Point C in Figure 4a; (d) A photograph of structural damage on Wuhan International Conference & Exhibition Center caused by ground subsidence.

Moreover, most subsiding areas of Zone 1 as shown in Figure 4a are situated at an alluvial plane formed by soft sediment soils of Yangtze and Han rivers as well as lakes. The alluvial deposits in this region are characterized by a highly compressible layer of silt and clay typically with a depth of 5~30 m below the ground surface [11], and therefore can easily lead to heterogeneous subsidence as a result of external load on the soft soils due to intensive construction of infrastructure and buildings.

4.3. Subsidence Related to Carbonate Karstification

In Figure 3, we can notice a strong spatial correlation between land subsidence variations and distributions of the known karst inventory (see Figure 1). Subsiding areas associated with carbonate karst geology are mainly identified in Zone 2, 4 and 6 (see Figure 3) which are close to the Yangtze
River and located in three major carbonate rock belts named Tianxingzhou, Daqiao, and Baishazhou, respectively. In these carbonate belts, various karst types of solution fissures, grooves and sinks as well as small caves are extensively developed, where the karst cave encountering rate of boreholes is about 46.0%~50.1%, and about 30% of karst caves are unfilled or half-filled [43]. This shallow carbonate karstification can provide a channel for loss of the upper sandy soil if hydraulic variations of groundwater frequently occur in the vertical and horizontal karst seepage zones [44]. In particular, there is a close hydraulic connection between the karst subsurface water and the Yangtze River, and consequently, the active water level variations of the river, especially during wet/flood seasons [35], promote the dissolution of carbonate rocks and amplify karst subsidence or collapses.

Figure 5a,b illustrate the zoomed deformation rate maps of Zone 2 and Zone 4, two representative karst-impacted sites corresponding to residential and industrial contexts, respectively. In Figure 5a, heterogeneous subsidence signals (−5~−15 mm/year) are detected in Zone 2, which includes several urban blocks near Yellow Crane Tower situated on Sheshan (Snake Hill), one of the most renowned cultural towers in China. Moderate subsidence recoded in the deformation rate map is restricted to the western sector of the karst area, with maximum LOS displacement rates of −10 mm/year. This subsiding area has suffered severe karst collapses in recent years, triggered by groundwater variations due to building construction in the neighboring Jinduhangong community [45]. Some of the damage observed in the facades of the buildings and paved surfaces during the collapse event in 2011 are illustrated in Figure 5c,d.

Figure 5. Mean LOS deformation velocity map superimposed on Google Earth image over (a) Zone 2 and (b) Zone 4. The light green polygons represent the distributions of carbonate belts. The subsidence region in Zone 2 is outlined by a red dashed line. The red triangle represents the karst surface collapse that occurred in December 2011. The red cross represents the Yellow Crane Tower; (c,d) illustrate the photographs of structural damage caused by karst collapses that occurred in 2011.
Comparing with the heterogeneous subsidence pattern in the Zone 2, a relatively significant and spatially continuous ground settlement, with LOS deformation rates from $-10$ to $-30$ mm/year, can be found in Zone 4 (see Figure 5b). These karstification-related subsiding areas, located in a major industrial region of Wuhan, contain some of production bases of SINOPEC Wuhan Company and Wuhan Iron and Steel Corp. (Wuhan, China). Therefore, we suggest that dissolution of shallow carbonate rocks is a major cause for land subsidence in the Zone 4 that might be compounded by possible groundwater over-extraction for industrial production demands.

5. Discussion

It is noted that a remarkable ground uplift phenomenon has been detected in both bank areas of Yangtze River, especially along the southern bank (see Zone 7 in Figure 3), with deformation rates ranging from $+5$ to $+17.5$ mm/year. Here this exceptional uplift signal is highlighted in a zoomed deformation velocity map in Figure 6a for a representative site of the Zone 7 that is located on the bank sector between the River and the Sha Lake, and is further illustrated in Figure 6b by the deformation time series corresponding to the SDFP points (labelled in Figure 6a).

![Figure 6](image)

*Figure 6.* (a) Mean LOS deformation velocity map superimposed on Google Earth image over the bank sector between the Yangtze River and the Sha Lake; (b) Displacement time series relevant to the SDFP pixels labeled as Point D, E and F in Figure 6a vs. water level time series of the Yangtze River.

Due to a good spatial correlation between the deformation signals and the river locations, the most likely cause for the ground deformation observed in these areas is groundwater discharges and recharges related to seasonal fluctuations in water levels of the Yangtze River. In fact, this strong relationship between mass movements of river banks (e.g., ground deformations, bank erosions or landslides) and hydrological processes involving both the magnitude and frequency rainfall events and the seasonal fluctuations of river water levels have been emphasized by previous studies in China [46] and worldwide [47–50]. Consequently, comparing the LOS displacement histories of these mentioned SDFP points and the water level changes of Yangtze River during the study period, we can observe a coherent temporal correlation of variations and amplitudes between them (see Figure 6b). For this study, we notice a 5 m decrease of the river water level measured in the dry season between autumn 2009 and spring 2010, which corresponds to a mean cumulative deformation of about $-10$ mm at three SDFP points. However, a significant change of the deformation direction occurred around the lowest water stand in early March 2010, and an accelerating uplift from the beginning of the wet season in Wuhan lasted until 11 August 2010 when the maximum value of cumulative uplift reached approximately 15 mm, following the peak of water levels amounting to 27.3 m on 30 July 2010.

In addition, a similar deformation behavior with the southern bank of the Yangtze River is found in the Zone 8 near Hongshan Square (see Figure 3), which is situated in a sector of the carbonate rock belt named Daqiao between Yangtze River and East Lake. We suggest that the seasonal fluctuations of
river water levels might partially contribute to ground deformation variations in this area. However, it is difficult to provide a comprehensive explanation for this deformation phenomenon due to relatively complex hydrological-conditions in this area, because ground water and pore water pressure data are not available for this study.

6. Conclusions

In this work, we applied StaMPS-based multi-temporal InSAR methodology to high-resolution TerraSAR-X images from 2009 to 2010 to detect ground deformation in Wuhan, and provided detailed information on spatiotemporal characterization of land subsidence and uplift in this region, for the first time. The results of accuracy assessment suggest that the InSAR-derived deformation rates agreed well with the levelling survey measurements with an average absolute difference of 3.1 mm/year and standard deviation of 1.8 mm/year. We found that noticeable land subsidence occurred widely in major urban areas of Wuhan and the average measured rates of subsidence ranged from approximately −5 mm/year to −67.3 mm/year during the period of observation. These subsiding areas were mostly concentrated in six commercial and industrial zones in Hankou and Wuchang districts, as well as three karst carbonate belts in Wuchang and Qingshan districts. The main contributing factors to the detected land subsidence include anthropogenic activities, natural compaction of soft soil, and karst dissolution of subsurface carbonate rocks. However, anthropogenic activities, such as intensive municipal construction (e.g., building and subway tunneling) and industrial production, have more significant impacts on the measured subsidence rates than natural factors. Moreover, remarkably secular uplift signals were detected along both banks of the Yangtze River, especially along the southern bank, with deformation rates ranging mostly from +5 mm/year to +17.5 mm/year. A strong temporal correlation is highlighted between the detected displacement time series and the water level records of the Yangtze River, suggesting that this previously unknown deformation phenomenon in Wuhan is likely related to seasonal fluctuations in the water levels of the Yangtze River.

This study demonstrates the potential of multi-temporal InSAR analysis of high-resolution SAR datasets for ground deformation monitoring in Wuhan, which is characterized by vulnerable hydrological environments and complex deformation regimes. The InSAR-derived results also indicate an urgent demand for regular and large-scale monitoring of deformation-relevant geohazards in this region, which could help not only to better characterize the development of catastrophic hazards relevant to ground deformation but to recognize previously unknown deformation problems, such as instability and collapse of Yangtze River embankments. In the near future, more acquisitions of SAR images (e.g., TerraSAR-X, Sentinel-1) and more ground observations, especially hydrological and meteorological data, will be collected, meaning a further joint analysis of multidisciplinary data could be carried out to thoroughly study such complicated ground deformation phenomena in this study area.

Acknowledgments: The work in this study was supported by the National Natural Science Foundation of China (No. 41590854, 41274024, 41431070, 41321063), the National Basic Research Program of China (No. 2012CB957702) and the Hundred Talents Program of The Chinese Academy of Sciences (No. Y205771077). The authors would like to thank DLR for providing the TerraSAR-X images via the TerraSAR-X AO project (No. MTH1827), TU Delft for providing the DORIS software, Hooper at the University of Leeds for providing the StaMPS software, and Li Gang at the Chinese University of Hong Kong for the helpful advice in the data processing.

Author Contributions: Lin Bai designed the study and wrote the manuscript, Liming Jiang supervised the study and reviewed the manuscript, and Hansheng Wang and Qishi Sun contributed to the discussions.

Conflicts of Interest: The authors declare no conflict of interest.

References


42. Raucoules, D.; Bourgine, B.; De Michele, M. Validation and intercomparison of Persistent Scatterers Interferometry: PSIC4 project results. *J. Appl. Geophys.* 2009, 68, 335–347. [CrossRef]


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
Article
Application of InSAR and Gravimetry for Land Subsidence Hazard Zoning in Aguascalientes, Mexico

Jesús Pacheco-Martínez 1, *, Enrique Cabral-Cano 2, Shimon Wdowinski 3, Martín Hernández-Marín 1, José Ángel Ortiz-Lozano 1 and Mario Eduardo Zermeño-de-León 1

1 Centro de Ciencias del Diseño y de la Construcción de la Universidad Autónoma de Aguascalientes, Av. Universidad # 940, Ciudad Universitaria, C. P. 20131 Aguascalientes, Mexico; mhernandez@correo.uaa.mx (M.H.-M.); aortiz@correo.uaa.mx (J.A.O.-L.); mezerme@correo.uaa.mx (M.E.Z.-L.)
2 Departamento de Geomagnetismo y Exploración, Instituto de Geofísica, Universidad Nacional Autónoma de Mexico, Circuito de la investigación Científica, Ciudad Universitaria, D.F. 04510 Delegación Coyoacán, Mexico; ecabral@geofisica.unam.mx
3 Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, Miami, FL 33149, USA; shimonw@rsmas.miami.edu
* Correspondence: jesus.pacheco@edu.uaa.mx; Tel.: +52-449-910-8456; Fax: +52-449-910-8451

Academic Editors: Zhenhong Li, Roberto Tomas, Norman Kerle and Prasad S. Thenkabail

Received: 23 August 2015; Accepted: 7 December 2015; Published: 17 December 2015

Abstract: In this work we present an application of InSAR and gravimetric surveys for risk management related to land subsidence and surface ground faulting generation. A subsidence velocity map derived from the 2007–2011 ALOS SAR imagery and a sediment thicknesses map obtained from the inversion of gravimetric data were integrated with a surface fault map to produce a subsidence hazard zoning in the city of Aguascalientes, Mexico. The resulting zoning is presented together with specific recommendations about geotechnical studies needed for further evaluation of surface faulting in these hazard zones. The derived zoning map consists in four zones including null hazard (stable terrain without subsidence), low hazard (areas prone to subsidence), medium hazard (zones with subsidence) and high hazard (zones with surface faulting). InSAR results displayed subsidence LOS velocities up to 10 cm/year and two subsidence areas unknown before this study. Gravimetric results revealed that the thicker sediment sequence is located toward north of Aguascalientes City reaching up to 600 m in thickness, which correspond to a high subsidence LOS velocity zone (up to 6 cm/year).

Keywords: land subsidence; Aguascalientes; surface cracks; ground failure; subsidence zoning; subsidence hazard

1. Introduction

Land subsidence induced by groundwater extraction is a man-induced geological hazard affecting many cities in the word. One of the main hazards on ground subsiding areas is the development of subsidence-related surface faults and earth fissures, because they damage housing and other infrastructure, decreasing their real estate value. In Aguascalientes (see Section 3.1), surface faults can develop displacement across their escarpment (Figure 1), with a variable width of an active zone up to ten meters in which housing structures are easily damaged. Surface faults usually develop over the bounds of the subsidence zones, but they are not rare toward the central part of subsiding areas. Another subsidence-related problem is the increased flood likelihood due to the disruption of sewage utilities and changes in the surface drainage.
Figure 1. Surface faults due to differential subsidence in Aguascalientes. (a) Surface fault affecting man-made structures; (b) Surface faulting has developed escarpments up to 1.80 m high. The location of the subsidence-related surface faults is shown in Figure 2b.

In this work we use the terms “ground fault”, “surface fault”, “surface crack”, or “fissure” to refer to subsidence-related terrain discontinuities, and the terms “quaternary fault”, “pre-existing fault” or simply “fault” refer to pre-subsidence tectonic faults.

Land subsidence due to groundwater extraction is a slow and gradual process whose effects are usually observed long after subsidence has begun. Because there is no loss of life usually associated with its occurrence, subsidence may not be considered a major disaster. Hence, land subsidence is usually not considered in risk mitigation public policies in Mexico as opposed to other hazardous events such as earthquakes, volcanic eruptions, hurricanes and floods. Unlike subsidence, these events are characterized by short duration and high intensity with consequences and catastrophic effects that are immediately observable.

The Disaster Risk Index (DRI) defined by the United Nations Development Programme [1] provides a methodological tool to assess the impact of catastrophic events [2]. Currently, the DRI is solved for three natural hazards: earthquake, tropical cyclones and flooding, but risk assessment of other hazards including land subsidence due to groundwater extraction remains as scientific challenges and as works in progress.

Many studies have reported the occurrence of land subsidence and its effects [3–10]. Studies dealing with monitoring and detecting subsidence are numerous [11–17], as well as works addressed to land subsidence modeling for calculating the expected magnitudes and rates of subsidence for different scenarios [18–23]. Nevertheless, subsidence risk assessment are still scarce [24–29].

In general, risk assessment involves three basic elements: (a) characterization of hazardous event, which is the natural or anthropogenic event able to cause life-loss or property damage (b) quantification of physical exposure, which refers the number of lives or assets exposed to the hazardous event; (c) determination of vulnerability, which is the features that make the physical exposure able to absorb the impact of the hazardous event [1,2,30,31]. Once the elements of risk have been determined, risk is calculated through a model which relates these elements of risk.

Risk subsidence assessment has been addressed in different ways. For example some works propose methodologies for the determination of areas prone to develop subsidence-related faulting [23,26,27,32–34]. Other works proposed methodologies to determine the areas that are likely to develop subsidence [29,35]. In all these approaches, the physical exposure has been the terrain but not constructions.
In this work we propose an approach for producing subsidence hazard maps, where man-made structures are the assets exposed to the hazardous event, using the city of Aguascalientes, Mexico as a study case. We processed InSAR data to derive subsidence velocity maps which were combined with geophysical and geological information to produce a zoning map of subsidence hazard.

The presented work covers only the first part of the subsidence risk assessment process: the characterization of the hazardous event. Nevertheless, the resulting maps from this approach provide valuable information for damage prevention in cities subject to subsidence.

2. Methods

2.1. Identification of Factors Contributing to Subsidence Hazard

Hazard is usually characterized in terms of the probability that a hazardous event occurs with certain intensity during a given time period [30]. This approach works for phenomena such as earthquakes or hurricanes, which cannot be avoided and whose magnitude cannot be reduced. In those cases, historical data of the events and their magnitudes are processed with probabilistic techniques in order to estimate the magnitude of an event in a particular time period.

However, land subsidence due to groundwater extraction is a predictable phenomenon, whose occurrence can be avoided if groundwater is not extracted, and whose intensity can be reduced by changing the ground water extraction policies. The mechanism leading to the generation of subsidence due to groundwater extraction and associated faulting is well-known [7,36–40]. Nevertheless, the ability to calculate time, location, speed and magnitude of subsidence is limited due to the difficulty to determine the parameters that control the process [41,42].

Subsidence occurrence is a combination of two factors: (a) the existence of unconsolidated or poorly consolidated sediments deposits that comprise the aquifer system and (b) lowering of the groundwater level [37,38,43]. The existence of large thickness of sediments prone to consolidation in the subsoil, which contain water susceptible to be pumped, is by itself the geological environment potentially prone to subside, whilst water table lowering is the triggering factor of subsidence. However, if the aquifer system is not comprised of unconsolidated sediments, then subsidence will not develop, even if a groundwater level reduction takes place. Thus, the existence and spatial distribution of consolidation-prone basin-fill sediments is the first independent factor to be considered for analyzing land subsidence hazard zoning, which is the mapping of the areas prone to subsidence and the lowest level of hazardous for constructions.

In areas where subsidence began long ago, subsidence can evolve in two manners: (a) subsidence develops in a uniform way generating a vertical displacement field with low horizontal gradients; (b) subsidence develops differentially, then the horizontal gradients of the vertical displacement field are large enough to develop horizontal stresses leading to ground failure development including surface faults and cracks [7,23,44,45].

Uniform subsidence changes the surface slope and modifies the slope of the sewer utilities, developing new flooding-prone areas. Hence, zones where uniform subsidence occurs are the second hazard level for civil structures and urban infrastructure.

Surface faulting is the main concern for property owners and local governments in subsiding urban areas not only for their damage potential but also because ground failures can severely reduce real estate value. As a result, areas where surface faults have developed pose the highest hazard level to civil structures and other urban infrastructure.

In summary, the key elements to assess the hazard level for man-made structures due to land subsidence by groundwater extraction are: (1) determination of existence of deformable sediments containing groundwater; (2) detection of sediment zones with subsidence in progress; and (3) location of affected areas by surface faulting.
2.2. InSAR and Gravimetry for Determining Areas Having Subsidence Hazard Factors

Both gravimetry and InSAR are useful tools for determining two of the three subsidence hazard factors previously discussed. Gravimetric analysis can be used to determine the distribution of sediments prone to be consolidated, while InSAR detects the affected areas by subsidence. The presence of surface faults, which is the other subsidence hazard contributing factor, can be obtained through field based cartography.

Gravimetric measurements have been used successfully to determine the distribution and thickness of granular filling in sedimentary basins [23,34,46]. In central Mexico, areas undergoing land subsidence due to groundwater extraction are usually comprised of sequences of unconsolidated sediment overlying a more dense volcanic or sedimentary rock formation. Consequently, the density contrast between the underlying consolidated rock and the overlying sediments is significant, which is very favorable for a gravimetric study.

The result of a gravimetric survey is a gravimetric anomaly map, which is the difference between the measured and theoretical gravity field. This difference is attributed to the density heterogeneities in subsurface: low gravimetric anomalies suggest presence of low density material close to the surface, and high values of the anomaly indicate denser material strata close to the surface. Hence, a gravimetric analysis provides valuable information about the spatial distribution and thickness of sediments.

Furthermore, some studies have shown that in geological settings susceptible to developing subsidence by groundwater extraction, there is an inverse correlation between sediment thickness and the gravimetric anomaly [23,34,47]. This relationship can be used as a qualitative way to characterize areas with large sediment thickness.

Additionally, gravimetric anomaly data can be inverted or directly modeled in order to elaborate models of the sediment thickness distribution. The modeling may be enhanced by constraining the model with lithological and other direct observations, allowing the generation of detailed maps of sediment thickness distribution.

Because land subsidence due to groundwater extraction may affect large areas, its detection and quantification are quite suitable for satellite remote sensing techniques. Interferometric synthetic aperture radar (InSAR) techniques have been successfully used to characterize subsiding areas [11,48,49].

3. Case Study: Subsidence Hazard Zoning of Aguascalientes, Mexico

3.1. Study Area

The city of Aguascalientes is located within the Aguascalientes graben in central Mexico, 430 km NW of Mexico City. Close to one million inhabitants live in the city and suburban municipalities, 725,000 in the city of Aguascalientes, and the other 225,000 in the 7 surrounding municipalities [50]: Cosío, Jesús María, Rincón de Romos, Pabellón de Aretaga, San Francisco de los Romo, San Pedro Piedra Gorda and Luis Moya (Figure 2a). Intense groundwater extraction initiated in the early 1970’s due to an increase in agricultural and industrial activities, triggering land subsidence and development of surface faults [51–53] and even the reactivation of tectonic faults [7].

According to SIFAGG (Sistema de Información de Fallas y Grietas) [54], currently 208 surface faults and fractures have been mapped throughout the entire Aguascalientes valley with an accumulated length of 290 km, affecting 1865 buildings mainly housings, from which 1438 of those are located within the city of Aguascalientes. Figure 2b shows only those surface faults within Aguascalientes City.
3.2. Determination of Deformable Sediment Distribution

The total thickness of the unconsolidated sediments in the study area (Figure 2b) was determined through a gravimetric study, which consisted in the surveying and processing of 339 ground-based gravimetric measurements, according to Telford et al. [55] and using a Scintrex CG5 gravimeter.

Figure 3a shows the Bouguer’s anomaly of the study area. Gravity data were modeled according to Singh and Guptasarma [56] using PyGMI software developed by Cole [57] in order to determine the shape of the sedimentary package and underlying rocks. The software works out the gravimetric anomaly produced by an initial model of the rocky stratum and the sediments distribution. That is the calculated anomaly. The initial model is modified until the calculated anomaly matches the anomaly determined by processing field data (measured anomaly). The model is more realistic if there are data for constraining some points of depth to rocky stratum, or density information of the subsoil materials.
Figure 3. Gravimetric anomaly of the study area. (a) Measured anomaly; (b) calculated anomaly from model.

We used $1500 \times 1500$ m and 500 m high blocks for the gravimetric modeling. Figure 3b shows the calculated Bouguer’s anomaly from the obtained model of bedrock and sediment thicknesses distribution.

The geology of the Aguascalientes Valley was broadly described by Aranda-Gómez [51] and Loza-Aguirre et al. [58]. For the purpose of this work, we used a simplified stratigraphic column of the subsiding area elaborated with lithological logs wells information (Figure 4). The densities associated to each stratigraphic unit used for the gravimetric modeling were those reported by Pacheco-Martínez et al. [59]. The non consolidation-prone units are composed of a polycmite conglomerated, rhyolite and ignimbrite, and the consolidation-prone unit is a quaternary alluvial sequence with a diverse content of silt, sand and gravel.

Figure 4. Simplified stratigraphic column of the study area.

The top surface of the rocky stratum (Figure 5a) was constrained in several locations through information of logs wells lithology and rock outcrops on both sides of the graben. Finally, an isopach
map, which is a sediment thickness variations map, was calculated from the resulting elevation difference between the top bedrock surface of the conglomerate and volcanic rocks and the ground surface.

Figure 5. (a) 3D model of the surface topography and bedrock topography; and (b) isopach map of sediment thickness in the city of Aguascalientes.

In order to characterize the subsidence field within the Aguascalientes graben, a subsidence LOS (line of sight) velocity map was obtained from 34 ALOS scenes (ascending track 191, frames 420, 430) acquired between August 2007 and March 2011. We used the ROI_PAC software developed by NASA’s Jet Propulsion Laboratory [60]. Imagery was processed with the differential InSAR technique (D-InSAR) to obtain 28 interferograms to derive an InSAR subsidence LOS velocity map. The InSAR time series analysis was obtained using the Short Baseline Subset Analysis (SBAS) technique [61–63]. Topographic correction was applied according to Fattahi and Amelung [64]. A threshold of 0.7 temporal coherence was used for the resulting velocity map (Figure 6a). The high calculated temporal coherence of the study area indicates a minimal effect from unwrapping errors.

3.3. Determination of the Subsidence-Affected Area

Figure 6a shows three subsiding areas. The largest one corresponds to Aguascalientes Valley, where the rate of subsidence has reached 10 cm/year in two locations (Rincón de Romos and José Gómez Portugal). Both locations are agricultural areas with intense groundwater pumping. The other two subsidence areas located at north Loreto and east NISSAN II (Figure 6a), also correspond to agricultural zones. These two subsiding areas located outside of Aguascalientes Valley had not been documented before this study. Subsidence velocity in both areas is up 6 cm/year.

Figures 2a and 6a show that subsidence is developing in flat areas surrounded by topographic elevations of rocky outcrops. The velocity map of Figure 6a shows that during the period from 2007 to 2012 the area of rocky outcrops encircling at the Valley did not develop subsidence. This is consistent with other works which indicated that subsidence occurs only within the Valley where a significant thickness of sediments exists. Hence, the rocky outcrops located outside the valley (dark blue areas...
in Figure 6a and surrounding mountains in Figure 2a) can be considered stable zones and reference points for monitoring areas with subsidence.

Figure 6. Subsidence velocity maps for (a) Aguascalientes Valley; (b) Aguascalientes City including surface faults and (c) Subsidence graph from INEGI GPS station. White circled numbers indicate surface faults outside of subsidence area.

Subsidence in Aguascalientes has a nonlinear trend as shown by a subsidence time series from the INEGI GPS station (Figure 6c). This graph shows that subsidence velocity is decreasing more and more since the beginning of the records, but it has been at a steady pace since 2005.

3.4. Surface Faults Mapping

An updated subsidence-related surface fault cartography was generated using both information available in previous surface fault maps and field verification of recently generated surface faults not included in the previous maps. The resulting surface fault location was integrated into Figures 2b and 6b.

Figure 6b shows that subsidence-related surface faults preferentially develop within the subsiding area and along its limit where differential subsidence favors their formation. Nevertheless, some segments of them (1, 2 and 3 in Figure 6b) are located away from the limit of subsidence area, in locations where the subsidence velocity map shows null subsidence. The location of these segments of surface faults coincides with pre-subsidence tectonic faults [7] which are cutting the sediment unit. These segments of surface faults located outside of the subsiding area suggest that preexisting tectonic faults are reactivated by the influence of the subsidence stress field.
4. Results

Subsidence Hazard Zoning

The subsidence hazard zoning map shown in Figure 7c was calculated from the hazard matrix described in Table 1. The hazard level is directly dependant on the existence of those factors discussed in Section 2.1. For example, the null hazard level (last column in Table 1 and green tones in Figure 7c) corresponds to the area where there is not a single hazard factor: consolidation-prone sediments do not exist (Figure 7a), and subsidence and faults have not been observed (Figure 7b). Conversely, the high hazard zone (second column in Table 1 and red lines in the hazard map of Figure 7c) is the area affected by surface faults, and in which the three hazard factors are present, except for the surface faults segments 1, 2 and 3 discussed in Section 3.3.

![Figure 7](image_url)

*Figure 7*. (a) Total sediment thickness map; (b) subsidence LOS velocity map; (c) subsidence hazard zoning map of Aguascalientes City. Hazard zones are shown over a shaded relief for reference.

<table>
<thead>
<tr>
<th>Hazard Factors</th>
<th>Hazard Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
</tr>
<tr>
<td>Unconsolidated sediments existence</td>
<td>Yes</td>
</tr>
<tr>
<td>Subsidence detected by InSAR</td>
<td>Yes</td>
</tr>
<tr>
<td>Observed damages by subsidence-related surface faults</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 1. Hazard factor matrix of subsidence effects on constructions.

In each of the resulting hazard zones, the effect of ground subsidence on the existing civil structures and urban infrastructure is different, and the prevention and mitigation actions must be in accordance with the type of potential threat. We provide recommendations for geotechnical studies that are needed to assess subsoil conditions according to the hazard level, both for any future real

95
estate development and for those existing structures. We list the recommendations from lowest to highest hazard level.

Null hazard zone: This corresponds to the area in which land subsidence was not detected during the period of InSAR data acquisitions (subsidence LOS velocity = 0 in Figure 6), and surface faults have not been observed yet. Furthermore, in this zone the total thickness of unconsolidated sediments is negligible. Null hazard zone corresponds to the rocky outcrops area and those areas in which the bedrock is covered by a thin stratum of sediments. Standard geotechnical studies are sufficient for this zone. 

Low hazard zone: This is the zone of unconsolidated sediments forming the overexploited granular aquifer system (sediment thickness > zero in Figure 5), which are bounded by rocky outcrops, and in which subsidence LOS velocity was zero for the observation period (dark blue tones in Figure 6) and surface faults have not been observed.

In this zone, standard geotechnical studies are needed, plus a superficial geological study searching for evidence of tectonic faults and cracks in the sediment unit, which could be reactivated if a decline in groundwater table take place in this zone.

Medium hazard zone: This corresponds to the zone of unconsolidated sediments forming the overexploited granular aquifer system (sediment thickness > zero in Figure 5), and in which subsidence was detected through the LOS velocity map (subsidence LOS velocity > 0 in Figure 6). Although this zone is undergoing subsidence, it is presented uniformly, such that terrain ruptures have not developed.

Geotechnical studies similar to those suggested for low hazard zone are needed, plus a geophysical survey for detecting any incipient or blind surface faulting. Resistivity profiles have been used with relative success for detecting surface faults in their initial stage when they may not be visible on the surface yet [7]. In case that resistivity profiles show an anomaly that could be related to an incipient surface fault, direct excavation of a trench may be needed in order to confirm its existence. Another threat to buildings in this area is the modification of surface drainage as a result of subsidence, which could develop flooding areas and also renders sewage systems ineffective. As this is a regional effect of subsidence, the municipal agency responsible for operating the sewage system along with the agency responsible for urban planning should take this issue into consideration for developing timely mitigation measures.

High hazard zone: This zone corresponds to areas where subsidence-related surface faults and fissures have been detected due to the damages they cause to constructions and terrain surface along their trace. The zone includes a band of terrain in each side of the trace of the surface fault in which subsidence is developing in a differential manner.

Geotechnical studies similar to those suggested for the medium hazard zone are needed, plus a detailed analysis to determine the influence width of subsidence-related faults in which civil structures may be damaged due to uneven subsidence.

5. Discussion

The approach of this work was to elaborate a zoning map of the hazard to which constructions are exposed in subsidence areas. The resulting hazard zoning map includes areas prone to subsidence, those currently undergoing ground subsidence and existent surface fault traces. The hazard map does not include the zones prone to surface faults generation or future zones of surface faults growth.

Although the subsidence-induced faulting mechanism due to groundwater extraction is very well understood, examples of determination of parameters involved in surface fault generation are scarce, and their field and laboratory measurements are still not a common practice for many real estate developers. As a consequence, subsidence-related fault modeling customized for specific civil structures that may be derived in individualized construction design is still an unattainable goal.

Some authors have proposed solutions for determining the zones prone to develop subsidence-related ground faulting [15,23,27,32,33]. These methodologies are based on the determination of the horizontal gradients of a measured parameter. The proposed solutions correlate
ground faulting areas to high values of horizontal gradients of gravimetric anomalies [23], soil vibration frequency [32,33] and high values of horizontal gradients of subsidence magnitude [15,27]. Their results show that these methodologies can forecast shallow ground fault generation, but further research is still needed to determine critical gradient threshold values related to ground faulting developing in specific cases. Also, further work is still needed to improve the temporal and spatial accuracy for surface fault generation predictions.

The most hazardous zone for constructions is on ground failures. The zoning includes only the linear feature representing the trace of the terrain rupture. However, ground failures have an active zone defined by the width of the trace in which differential subsidence is significant to induce damage to the constructions. Studies to determine the width of the active zone or the influence zone of a ground failure are practically nonexistent. More research is needed to obtain a reliable and practical methodology to determine accurately enough this parameter.

The recommendations in this work for geotechnical studies in subsidence areas are to explore subsoil in order to prevent effects of subsidence, mainly those related to surface faults. The current practice in Aguascalientes City and other Mexican cities where surface faults have damaged constructions is the continuous repair of buildings as long as ground failure does not affect their structural stability. Otherwise, buildings are demolished and the terrain is used for parking, green area or other uses except for building construction. Performance investigations of constructions built over ground failures are incipient [65,66]. Preliminary results show that a combination of flexible materials and special structures, including a device for restoring the level of construction at certain intervals, could be a solution to prevent damage to constructions.

As subsidence is a progressing deformation and rupture subsoil process, hazard levels could change with time. If subsidence continues, then new surface faults and fissures will be generated, and the existing ones will enlarge, producing new zones of high hazard. Conversely, if subsidence stops, then terrain surface will not experience differential subsidence. Surface faults and fissures will become inactive, and the terrain surface will be stable for constructions. In any case, the subsidence hazard zoning map should be periodically updated to prevent the implementation of wrong measures in risk management.

Parameters to determine hazard factors were obtained directly by on-site measurements (gravimetry, mapping of subsidence-related surface faults and rocky outcrops) and by remote sensing (InSAR). Hence, accuracy of the zoning depends on the resolution of the methods used for processing of measured field data.

In the case of the determination of rocky stratum and sediment distribution (low hazard zone), we used $1500 \times 1500$ m and $50$ m high blocks for the gravimetric modeling. Hence, the resolution achieved for the map production was $1500$ m. Therefore, features with sizes lesser than $1500$ m are not represented in the model of rocky stratum (Figure 5a). Hence, the model of rocky stratum might not include detail sufficient to identify topographic features that are triggering differential subsidence and causing increasing subsidence in specific locations.

However, the resolution of sediment thickness map distribution does not affect the accuracy of hazard zoning, because the limit of zones of low and null hazard was defined from a GPS surveying on the bounds of rocky outcrops with an accuracy of $\pm 4$ m. Likewise, the limit of zones of low and medium hazard is not influenced by the resolution of sediment thickness map because this limit was determined by subsidence measured by InSAR.

The ALOS-InSAR images used to produce the subsidence velocity map have a resolution of $30$ m. Hence, as we used a GPS with sub-meter resolution for surveying of rocky outcrops and surface faults, the greatest error in zoning map limits could come from the subsidence velocity map. However, this error is negligible for scale maps 1:10,000 which is the scale more frequently used for urban planning management.

The zoning map does not include either the effect of lateral variations in the sediment thickness or the influence of subsidence magnitude developed in specific locations. The insertion of these factors...
would improve the zoning process, which would result in a map with more zones than the map derived from the factors analyzed in this work.

Each of such zones would indicate the hazard related to subsidence more accurately. Nevertheless, the simplification of the hazard zoning in four zones as was proposed in this work allows specific and practical recommendations for geotechnical exploration.

A map with many zones or a continuum zoning map could be an interesting scientific contribution. However, the resulting zoning map could lose its practicability and usefulness for urban planning and development. Additionally, zones with higher subsidence can be addressed as special and specific cases of medium hazard level (as determined in this work).

Risk assessment involves three stages: (1) hazard characterization; (2) quantification of exposed elements at the hazard and determination of their vulnerability; and (3) calculation of risk. Although the characterization of hazard factors is really relevant only in the context of risk assessment, our work represent a significant contribution to the development of a further complete methodology, mainly because stages 2 and 3 are still unresolved issues and lines of research in progress.

6. Conclusions

A hazard subsidence zoning map is a necessary element in risk management of subsidence effects on man-made structures. Gravimetric measurements and InSAR-derived subsidence velocity maps provide valuable information for this hazard map. Gravimetric surveys are quite suitable for determining sediment thickness which may be prone to consolidation and consequently develop ground subsidence, while InSAR techniques are most suitable for a precise characterisation of the subsidence field.

A combination of both techniques, along with reliable subsurface surface faulting information, was used to derive a subsidence hazard map for the city of Aguascalientes. This map will allow state and municipal government agencies to clearly specify specific geotechnical and geophysical studies according to the hazard level of the zone in which new constructions are planned.

Due to the dynamic nature of the subsidence process, hazard zoning maps, such as the one described in this work, need to be continuously updated. Future InSAR analysis using newly acquired data or enhanced processing techniques along with other geodetic infrastructure development such as continuously operating GNSS stations will provide updated ground subsidence velocity maps which, in turn, will allow periodic updates of the subsidence hazard maps.

Acknowledgments: Jesús Pacheco-Martínez thanks CONACYT and PRODEP for financial support “Investigación apoyada por el CONACYT”. Shimon Wdowinski and Enrique Cabral-Cano acknowledge support through NASA-ROSES grant NNX12AQ08G. Some of the geodetic infrastructure used in this study was developed with support from UNAM-PAPIIT projects IN104213-2 and IN109315-3 for Enrique Cabral-Cano. This material is partly based on data provided by the UNAVCO Facility with support from the National Science Foundation (NSF) under award EAR-1338091.

Author Contributions: Jesús Pacheco-Martínez developed the research approach. He processed both InSAR and gravimetric data, and wrote the manuscript. Enrique Cabral-Cano and Shimon Wdowinski participated in the advising of the InSAR data processing and they gave support improving the manuscript. Martín Hernández-Marín, José Ángel Ortiz-Lozano, Mario Eduardo Zermeño-de-León contributed sharing information on land subsidence, and damages caused by surface faults to the constructions in Aguascalientes City, they also collaborated in the reviewing the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References


© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
Article

Methodology for Detection and Interpretation of Ground Motion Areas with the A-DInSAR Time Series Analysis

Roberta Bonì *, Giorgio Pilla and Claudia Meisina *

Department of Earth and Environmental Sciences, University of Pavia, Via Ferrata 1, Pavia 27100, Italy; giorgio.pilla@unipv.it
* Correspondence: roberta.boni01@universitadipavia.it (R.B.); claudia.meisina@unipv.it (C.M.);
Tel.: +39-03-8298-5842 (R.B.)

Academic Editors: Zhenhong Li, Roberto Tomas, Zhong Lu, Richard Gloaguen and Prasad S. Thenkabail

Received: 30 June 2016; Accepted: 16 August 2016; Published: 22 August 2016

Abstract: Recent improvement to Advanced Differential Interferometric SAR (A-DInSAR) time series quality enhances the knowledge of various geohazards. Ground motion studies need an appropriate methodology to exploit the great potential contained in the A-DInSAR time series. Here, we propose a methodology to analyze multi-sensors and multi-temporal A-DInSAR data for the geological interpretation of areas affected by land subsidence/uptilt and seasonal movements. The methodology was applied in the plain area of the Oltrepo Pavese (Po Plain, Italy) using ERS-1/2 and Radarsat data, processed using the SqueeSAR™ algorithm, and covering time spans, respectively, from 1992 to 2000 and from 2003 to 2010. The test area is a representative site of the Po Plain, affected by various geohazards and characterized by moderate rates of motion, ranging from −10 to 4 mm/yr.

Different components of motion were recognized: linear, non-linear, and seasonal deformational behaviors. Natural and man-induced processes were identified such as swelling/shrinkage of clayey soils, land subsidence due to load of new buildings, moderate tectonic uplift, and seasonal ground motion due to seasonal groundwater level variations.

Keywords: Advanced Differential Interferometric SAR (A-DInSAR); land subsidence; expansive soils; time-series analysis; principal component analysis; Oltrepo Pavese; Po Plain

1. Introduction

Land surface dynamics can be the evidence of different processes (i.e., swelling/shrinkage of expansive soils, aquifer-system compaction, tectonic movements, and consolidation due to the load of new buildings) of natural and anthropic origins. In many cases, the movements are due to the complex interactions of multi-driving forces, which act at various spatial and temporal scales [1–3]. The interference and/or overlap makes the effective decomposition of the components of motion difficult. For this reason, only a few authors have performed the decomposition of multi-driving force processes that trigger surface movements [4,5].

The areas affected by displacements need effective identification methods, semi-continuous spatio-temporal surface deformation monitoring and mechanism recognition analysis, in order to adopt the appropriate risk mitigation strategies, and to develop a sustainable management of natural resources.

Advanced Differential Interferometric SAR (A-DInSAR) is a powerful remote sensing tool, capable of mapping displacements over wide areas at very high spatial resolutions. The technique is based on the processing of multiple interferograms derived from a large set of SAR images in order to obtain a displacement time series, along the line of sight (LOS) of the satellite, of radar targets on the Earth’s
To date, several studies have documented the application of A-DInSAR data to monitor and to improve knowledge regarding different types of land surface dynamics [9–11].

Up to now, geohazard-mapping methodologies are mainly based on the visual inspection of average LOS velocity, such as in the case of the ground stability areas in the framework of the PanGeo project [12], or in hotspot analysis and hotspot and cluster analysis of average LOS velocity to detect slow-moving landslides [13,14].

Most of the studies aimed at identifying the areas affected by ground motion, at a regional and local scale, were focused on the average rates of displacement detected by the A-DInSAR technique [15–17]. Meisina et al. (2008) [16] implemented a systematic methodology to identify anomalous areas selected on the basis of the average LOS velocity, and Peduto et al. (2015) [17] introduced the average LOS velocity by weighting the radar target values on their coherence value. Both methodologies were mainly based on the average LOS velocity analysis, constrained by the fact that previous A-DInSAR data allowed the obtaining of noisy time series that are difficult to interpret, showing some limitations in the distinguishing of ground deformation due to different processes, or in detecting shallow deformations caused by seasonal processes.

Taking into account the recent improvements of A-DInSAR data acquired by the COSMO-SkyMed satellites and the current ESA Sentinel missions, which act at higher spatial and temporal resolution, it is necessary to develop an appropriate methodology to analyze extremely large datasets that consist of a large number of measuring points.

The goal of this study is to present a novel methodology to analyze multi-sensors and multi-temporal A-DInSAR data for the geological interpretation of areas affected by land subsidence/uplift and seasonal movements. The approach is aimed at (i) improving the detection of ground motion areas; (ii) understanding the spatio-temporal evolution of ground motion areas; and (iii) supporting the interpretation of driving-force mechanisms. The proposed methodology is addressed to overcome limitations, such as the analysis of large data sets, by allowing exploitation of the great potential contained in A-DInSAR time series. In this paper, for the first time, ground motion areas are identified on the basis of their peculiar ground deformation behaviors.

The procedure was applied in the Oltrepo Pavese (Po Plain, Italy) by use of ERS-1/2 and Radarsat data processed using the SqueeSAR™ algorithm, and covering the time spans from 1992 to 2000 and from 2003 to 2010, respectively. The test area is a representative site of the Po Plain, characterized by various geohazards with moderate rates of motion. The obtained results are helpful for scientists and authorities in charge of land use planning in order to improve public safety and to assess the management of groundwater resources.

2. Study Area

The study area is located in the central area of the Po Plain, within the Oltrepo Pavese (Italy), covering an area of about 440 km² (Figure 1a). The plain of the Oltrepo Pavese is constituted of alluvial quaternary deposits, originating from the combined action of Apennine streams that form coalescent fans and of the Po River. These deposits overly the Miocene-Pliocene marine substratum, constituted by sandy-marls, sandstones, conglomerates, gyspy-marls, and calcareous-marls [18,19].

Three main geomorphological units composed of quaternary deposits were previously distinguished: post-Würmian alluvial deposits, Würmian–Holocene alluvial deposits, and pre-Würmian alluvial deposits [20,21].

The post-Würmian alluvial deposits were sedimented by the most recent depositional activity of the Po River, and by the recent and present deposition from flooding events of the main Apennine streams. These deposits are localized in proximity to the Po River, and they are composed of sand, sandy silt, and silt. The deposits from the Apennine streams are characterized by gravel and sand. In general, post-Würmian alluvial deposits contain a shallow phreatic aquifer with a water level near ground surface.
The Würmian–Holocene alluvial unit is the most prevalent geomorphologic unit of the study area. These deposits are made up of alternating sand and gravel, with interbedded clays or argillaceous silt. A shallow phreatic aquifer and deeper aquifers, of both phreatic and confined types, were distinguished in this unit. Deeply buried structures, consisting of a series of folds and fold-faults, shaped in the tertiary marine basement, have a direct control over aquifer geometry [22]. These alluvial deposits show the constant presence of a tabular clayey-silty cover, which acts as a seal, and limits water infiltration [21]. Groundwater exploitation of the Oltrepo Pavese is mainly derived from the aquifer contained in this unit.

Conversely, the pre-Würmian unit, located in the southern part of the study area, consists of older fluvial terraces, and is constituted of gravel and sand with a silty matrix, characterized by lower permeability [19]. These deposits are strongly weathered and covered with clay soils. Locally, loess deposits overlap this unit. The depth of the water level in this unit reaches values up to 20 m in depth.

In the study area, quaternary sediment thickness decreases from west to east, with the minimum thickness and the outcrop of the marine substratum in correspondence with the Stradella thrust, evidence of the neotectonic activity that affects the area [23]. The distensive regime of the neotectonic activity supported higher alluvial sediments accumulation in the northwestern sector of the plain with respect to the southeastern sector [24]. The structural setting of the Oltrepo Pavese plain is conditioned by the presence of an important NE–SW tectonic discontinuity, known in the literature as the Vogherese Fault (Figure 1b) [25]. This tectonic lineament is a portion of the composite seismogenetic source called Rivanazzano-Stradella, belonging to the Northern Apennines outer thrust front [26], which triggered an earthquake of Mw 3.94 ± 0.34 in 1971 [27].

From the climatic point of view, the plain area of the Oltrepo Pavese is characterized by a humid continental climate, with average annual rainfall of around 700 mm. Two rainy seasons, with maxima in May and in October–November, have been detected by previous authors [28]. In the last two decades, the region has experienced many drought periods, such as from March 1989 to August 1993, and from May 1998 to September 2000 [28].

Here, the analyses of the drought periods, performed by Meisina et al. (2003) [28] were extended from 1960 to 2010, using available measurements from 1960 and to match satellite observation time intervals. Therefore, climatic measurements acquired at the Voghera weather station were exploited, and the annual rainfall deficit was computed as the difference between the monthly water balance (difference between precipitation and potential evapotranspiration) and the average monthly water balance obtained for the period from 1960 to 2010. A significant drought period was identified from 2003 to 2007.

It is worth noting that the plain of the Oltrepo Pavese is characterized by discontinuous urban fabric, and 68% of the monitored area is covered by non-irrigated arable land [29]. The study area is a representative site of similar geological contexts in the Po Plain, where geohazards, due to natural and anthropic factors, were previously recognized. Volume changes of clayey soils; shrinkage during drying periods, and swelling during wet ones, resulted in severe damage to overlying structures [30].

**Engineering Geological Units**

Recent investigations have been performed to investigate the geotechnical proprieties of the quaternary deposits in the first 15 m of depth of the Oltrepo Pavese plain. On the basis of the analyses of penetrometer results (176 Dynamic Cone Penetration Test, 20 Standard Penetration Test and 266 Cone Penetration Test), calibrated against soil profile logs, six geotechnical classes of non-cohesive soils (from I1 to I6) and four geotechnical classes of cohesive soils (from C1 to C4), have been previously distinguished [31].
Figure 1. (a) Location of the study area. Sources: Esri, DigitalGlobe, Earthstar Geographics, CNES/Airbus DS, GeoEye, USDA FSA, USGS, Getmapping, Aerogrid, IGN, IGP, and the GIS User Community. (b) Distribution of the engineering geological units in the plain area of the Oltrepo Pavese. Damaged buildings, detected by Meisina et al. (2006) [31], are also reported. (c) Thickness percentage of the geotechnical classes in different depth intervals for each engineering geological unit (from Meisina et al., 2006) [31].

Taking into account the geotechnical classification of these soils, eight engineering geological units were introduced in order to reduce the lithologic variability and to map the geotechnical behavior of representative and homogeneous geotechnical profiles (Figure 1b,c). Unit 1 represents the post-Würmian alluvial deposit of the Po River, and is mainly constituted of non-cohesive soils (I3). Conversely, Units 2, 3, 4 and 7 are Würmian–Holocene alluvial deposits, characterized by cohesive soils. Unit 5 is composed of the alluvial fan of the Scuropasso River, and Units 6 and 8 of the pre-Würmian terraced deposits.
Figure 1b highlights the prevalent reported occurrence of damaged buildings [31], located over Units 8 (34%) and 6 (29%), which correspond with pre-Würmian deposits, and over Units 5 (20%) and 4 (9%), which correspond, respectively, with the alluvial fan of the Scuropasso River and Würmian-Holocene deposits, which are characterized by a high content of clay deposits that are susceptible to swelling/shrinkage processes due to the moisture changes in the dry/wet periods [31].

The majority of problems are experienced by single-story family residences, founded on conventional shallow strip, concrete footings, which generally extend to a depth of between 1 m and 2 m below ground level [30]. In most of the cases, the remedial measures adopted were pile underpinnings, which reduces the value of the house by 20% [30].

3. Advanced DInSAR Data

The Persistent Scatterer Interferometry (PSI) technique represents a class of the A-DInSAR techniques that uses radar targets on the Earth’s surface to produce, starting from a set of images, the displacement time series along the line of sight of the satellite (LOS) of individual persistent scatterers (PS) [6]. The technique allows the obtaining of data regarding land deformation over wide areas with millimeter precision [6]. However, it has some limitations in rural areas due to low density of PS and phase ambiguities. A new algorithm, named “SqueeSAR™”, has been developed in order to overcome this problem and to extract the deformation of distributed scatterers (DS), increasing the density of measuring points; for this reason, it was exploited to process the SAR scenes used in this work [32].

The input satellite data for this analysis consist of SAR images acquired by sensors operating in the C-band (wavelength, 5.6 cm; frequency, 5.3 GHz) onboard the ERS-1 and ERS-2 and Radarsat (RSAT) satellites. The scenes were acquired in the ascending mode, covering the time intervals from 9 July 1992 to 2 August 2000, and from 24 March 2003 to 5 May 2010. The dataset acquired in the descending mode covers the time intervals from 3 April 1992 to 7 January 2001, and from 28 April 2003 to 5 January 2009. The ERS-1/2 images were acquired with a nominal repeat cycle of 35 days, while the Radarsat images were acquired with a nominal repeat cycle of 24 days. These scenes were processed with the SqueeSAR™ technique by Tele-Rilevamento Europa (TRE). The algorithm allows the extraction of movement measurements, not only from traditional persistent scatterers (PS), such as anthropic structures or rocks, but also from distributed scatterers (DS), such as sparse vegetated areas [32]. This permitted a high density of A-DInSAR data over non-urban areas. The algorithm improved the quality of the displacement time series; atmospheric effects can be better estimated and removed, and so the resulting time series are characterized by less noise. Available SAR images were processed with the SqueeSAR™ technique to identify acceleration and slowing of scatterers, in order to better observe the kinematic evolution of seasonal and non-linear processes.

The processing results have precision along the LOS, usually better than 1 mm/year, depending on the amount of available data, the local PS density (a key element in the estimation of spurious atmospheric phase components), and the distance from the reference point.

The geocoding accuracy of the PS locations is around +7 m in the eastern direction, and +2 m in the north–south direction, while the estimated accuracy of the elevation values was better than 1 m.

Additional details of the SAR datasets are summarized in Table 1. It has been possible to detect more than 20,000 PS-DS over an area of around 430 km² (Figure 2). It worth noting that, as explained in Section 2, the area is mainly characterized by non-irrigated arable land, which was classified by Plank et al. (2009) [33] as being not at all suitable for the A-DInSAR technique using the C-, X-, and L-band sensors. To provide a quantitative assessment of the PS-DS density for this area, given the land use, the Corine Land Cover (CLC), which was implemented by the Lombardia region using color and infrared orthophotos from IT2007 (made by Blom CGR; 50-cm pixels), and named “DUSAF 2007” [29] was exploited. More precisely, such analysis was carried out by selecting, respectively, PS and DS in the non-irrigable arable land (CLC code 211). The results reveal that the ascending scenes show a PS density of 3.29 and 18.69 PS/km², respectively, for ERS-1/2 and Radarsat sensors. Conversely, the PS-DS densities for the same dataset are 20.74 and 39.42 PS-DS/km². In the case of the descending
Remote Sens. 2016, 8, 686

scenes, PS densities are 8.83 and 13.56 PS/km², respectively, for ERS-1/2 and Radarsat sensors, and the PS-DS densities for the same dataset are 37.63 and 29.10 PS-DS/km². By applying the SqueeSAR™ technique, an improvement of measuring-point density was obtained for the non-irrigable land. However, the PS-DS density for this land use coverage still shows low densities, which are comparable to values obtained using the PSI technique over the non-irrigable arable land in Britain [34]. Higher PS-DS density was detected in the continuous and discontinuous urban fabric (CLC code 11), and the road and rail networks and associated land (CLC code 122), which reach an average of 334.27 ± 43.57 and 137.35 ± 31.96 PS-DS/km², respectively.

Figure 2. Average line of sight (LOS) velocity measured by the use of ERS-1/2 ascending (a) and descending (b) data (time interval from 1992 to 2000), and Radarsat ascending (c) and descending (d) data (time interval from 2003 to 2010). The reference point for each dataset is also reported.

Table 1. Details on the SAR datasets, their processing, and results.

<table>
<thead>
<tr>
<th>Satellite</th>
<th>Orbit</th>
<th>Incidence Angle</th>
<th>N° of Scenes</th>
<th>N° of PS-DS</th>
<th>PS-DS Density in Study Area (PS-DS/km²)</th>
<th>PS-DS Density for CLC Code 11 (PS-DS/km²)</th>
<th>PS-DS Density for CLC Code 122 (PS-DS/km²)</th>
<th>PS-DS Density for CLC Code 211 (PS-DS/km²)</th>
<th>Average LOS Velocity (mm/yr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERS-1/2</td>
<td>Asc.</td>
<td>22.1°</td>
<td>27</td>
<td>21,308</td>
<td>227.46</td>
<td>96.66</td>
<td>20.74</td>
<td>-0.2</td>
<td></td>
</tr>
<tr>
<td>ERS-1/2</td>
<td>Desc.</td>
<td>21.2°</td>
<td>79</td>
<td>32,917</td>
<td>367.77</td>
<td>149.40</td>
<td>37.63</td>
<td>-0.4</td>
<td></td>
</tr>
<tr>
<td>RSAT</td>
<td>Asc.</td>
<td>35.3°</td>
<td>89</td>
<td>34,503</td>
<td>369.09</td>
<td>172.32</td>
<td>39.42</td>
<td>-0.3</td>
<td></td>
</tr>
<tr>
<td>RSAT</td>
<td>Desc.</td>
<td>35°</td>
<td>84</td>
<td>27,273</td>
<td>322.76</td>
<td>131.03</td>
<td>29.10</td>
<td>-0.2</td>
<td></td>
</tr>
</tbody>
</table>


4. Methodology

Taking into account the quality improvement of the A-DInSAR time series, it is necessary to develop an appropriate methodology to exploit the great potential contained in the A-DInSAR data for the geological interpretation of ground-motion areas. A systematic methodology to analyze the temporal evolution of ground motion areas is proposed here (Figure 3).
Figure 3. Flowchart of the methodological approach to detect ground motion areas and to recognize the triggering mechanisms.

The procedure consists of three main phases. In the first phase, the vertical and E-W components of motion are disentangled, and a displacement time series (TS) accuracy assessment is performed. In the second phase, different statistical tests are applied in order to find the spatio-temporal pattern of the principal components of movement, and the kinematic model of the targets. The result of this step is the identification of areas with significant movement, so-called “ground motion areas”. Ground motion areas correspond to a cluster of a minimum 3 of PS, with a maximum distance of 50 meters, characterized by the same trends (linear, non-linear, seasonal). Ground motion areas do not have a geological significance, but the systematic procedure represents a useful and fast approach to detect sectors where A-DInSAR analysis detects spatio-temporal ground deformation, and where the attention of scientists has to be focused. Finally, the third step consists of a data fusion of the A-DInSAR data and the geological data to determine the causes of ground motion processes. Multidisciplinary data, such as geotechnical, hydrogeological, and land use data, are fundamental to recognize ground motion mechanisms.

Figure 3 illustrates the proposed methodological approach, which consists of the following steps:

**Step 1.** Decomposition of the vertical and horizontal components of motion (Phase I). A-DInSAR techniques allow the obtaining of the average LOS velocities of targets. Smaller values of the satellite incidence angle (usually 23°–35°) make the measurements more susceptible to the vertical component of motion with respect to the horizontal; however, ignoring the horizontal components may implicate an over-/underestimation of the deformation, leading to misinterpretation of the results [35]. By the use of the datasets acquired in the ascending and descending modes, the vertical and horizontal components of target motion can be extracted using the following equations:

\[
V_{\text{EW}} = \frac{V_{\text{desc}} / h_{\text{desc}} - V_{\text{asc}} / h_{\text{asc}}}{e_{\text{asc}} / h_{\text{asc}}} - \left( e_{\text{desc}} / h_{\text{desc}} \right) - \left( e_{\text{asc}} / h_{\text{asc}} \right) \tag{1}
\]

\[
V_{\text{VERT}} = \frac{V_{\text{desc}} / e_{\text{desc}} - V_{\text{asc}} / e_{\text{asc}}}{h_{\text{desc}} / e_{\text{desc}}} - \left( h_{\text{asc}} / e_{\text{asc}} \right) \tag{2}
\]

\(V_{\text{asc}}\) is the average LOS velocity acquired in the ascending mode, \(V_{\text{desc}}\) is the average LOS velocity acquired in the descending mode, \(V_{\text{EW}}\) is the vertical component of motion, and \(V_{\text{VERT}}\) is the E–W component of motion. The \(e_{\text{asc}}, h_{\text{asc}}\) and \(e_{\text{desc}}, h_{\text{desc}}\) are the LOS directional cosines, respectively, for ascending and descending orbits. First, the average LOS velocity of each dataset was interpolated by an inverse distance weighted (IDW) approach, and, then, the component of motion was decomposed using these interpolated maps. The method was previously used by several authors [36,37]. If the horizontal component of motion is absent, the vertical component of motion can be extracted using
the incidence angle ($\theta$) of the dataset acquired in one mode (ascending or descending) using the following equation:

$$V_{\text{VERT}} = \frac{V_{\text{asc,desc}}}{\cos \theta}$$ (3)

**Step 2.** A-DInSAR displacement time series accuracy assessment (Phase I). The displacement time series detected by each target may be affected by local error or regional trends, which can be observed in the whole dataset. Therefore, post-processing analyses of the A-DInSAR data are fundamental in order to avoid misinterpretation of unreal ground movements. The check of the displacement time series is performed by selecting the most coherent (>0.9) targets with an average LOS velocity in the range ±0.5 mm/yr, where no significant movements are expected. More precisely, the approach proposed by Notti et al. (2015) [38] was applied in order to correct problems due to phase unwrapping, regional unreal trends, and anomalous displacement detected on certain dates (i.e., unreal movements at the same time as meteorological events, such as snowfall).

**Step 3.** Principal component analysis (PCA) of the A-DInSAR displacement time series (Phase II). A-DInSAR measurements represent the cumulative ground movements (natural versus anthropic, superficial versus deep displacements, multi-year, and seasonal components); the overlapping of several ground motion components may make the accurate interpretation of ground deformation processes difficult. Recent studies have applied PCA to satellite-based time series analysis [39,40] to detect spatio-temporal deformation patterns.

Here, principal component analysis (PCA), implementing a matrix organization of location versus time (T-mode), was performed in order to decompose the long-term (multi-year) and the seasonal components of ground motion. A matrix, in which each column contains the LOS displacements for each SAR image, and where each row contains the displacement time series of the targets (PS-DS), was formed. The main outcomes are the correlation and covariance matrices, the eigenvalues and eigenvectors, the percent variance that each eigenvalue captures, and the principal component (PC) score maps. In interpreting the principal components, PC scores related to each observation (PS-DS) are useful for knowing the correlations with the principal components in the whole dataset; the higher values correspond to higher correlations with the analyzed PC. To detect the displacement time series of the PC, the displacement time series can be multiplied for the PC eigenvectors of each SAR image date.

**Step 4.** Identification of the kinematic model of the targets (Phase II). A-DInSAR data proves to be efficient for the estimation of linear and non-linear ground motion movements [8]. Hence, automated classification of displacement time series (TS) in large datasets are needed for TS interpretation at a regional scale. Here, an automatic sequential procedure, based on statistical tests, the PS-time program implemented by Berti et al. (2013) [41], was used to classify the TS into one of three predefined target trends: uncorrelated, linear, and non-linear. The program permits the detection of the date (break) where abrupt changes in slope in the non-linear TS are recorded. This tool also assigns an additional parameter to the non-linear TS, index “BICW”, related to the bi-linearity of the TS (i.e., TS that records BICW higher than 1.2 indicates a strong bi-linearity [38]).

**Step 5.** Identification of “ground motion areas” (Phase III). This task aims at detecting ground motion areas in order to focus the interpretation on significant sectors. Physical processes can be characterized as linear, non-linear, and seasonal ground movements. While the average velocity of targets with a linear trend is useful to separate movement and non-movement areas, the same parameter is not efficient for recognizing ground motion areas affected by non-linear and seasonal movements. Even targets with an average LOS velocity in the stable range ($\pm 1.5$ mm/yr) can be affected by significant seasonal and/or non-linear movements. Taking into account that non-linear and seasonal movements may cause considerable damage to buildings and infrastructure [42], PCA results were exploited to identify ground motion areas with linear, non-linear, and seasonal trends. The ground motion areas were identified using a spatial cluster of significant PC scores. First, we define the threshold of the PC scores by applying a statistical inspection of the PC score frequency distributions.
If the histogram of the PC scores shows a normal distribution (skewness = 0), we consider the threshold equal to twice the standard deviation, while, if the PC scores shows asymmetrical distribution (skewness \( \neq 0 \)), the threshold is defined as being equal to the interquartile range (IQR). Then, PS-DS with PC scores higher than the threshold are selected, and a buffer zone of 50 m around the detected measuring points can be derived.

Therefore, ground motion areas consist of clusters of a minimum of 3 PS-DS with a maximum distance of 50 meters. The approach to extracting ground motion areas is based on the procedure proposed by Meisina et al. (2008) [16], but, herein, the delineated areas are derived from the PCA analysis (Figure 4). The ground motion areas are not found on the basis of unstable velocity; hence, even seasonal and non-linear processes can be identified and taken into account for the interpretation of displacement time series. Following this, the kinematic model of the targets is combined with the ground motion areas in order to distinguish linear and non-linear processes. More precisely, the TS with BICW higher than 1.2 are selected to distinguish non-linear TS from the linear ground motion areas.

Figure 4. Flowchart of the phase II in the methodological approach to identify ground motion areas.
Step 6. Interpretation of “ground motion areas” (Phase III). Once the ground motion areas have been identified, the processes are interpreted by integrating them into a Geographical Information System (GIS) with ancillary data (such as geological, geotechnical hydrogeological, compressible thickness map, digital orthophoto, and damaged buildings distribution). Mechanism recognition is based on a cross-comparison of the representative subsoil geological profiles, and the relative displacement time series with multidisciplinary information. Analysis of the breaks and the detection of the deceleration and acceleration periods are crucial to correlate the processes with the triggering mechanisms. The final outcome is the attribution of driving-force systems and triggering mechanisms to ground motion areas.

5. Results

5.1. Post-Processing Elaborations (Phase I)

The decomposition of the horizontal and vertical components of motion highlights that the motion is mainly vertical, both in the period of 1992–2000 and 2003–2010 in the Oltrepo Pavese. The results are consistent with literature data over the study area [43]. Therefore, the east–west component of motion was considered to be negligible. Following this, the procedure was implemented on the datasets that shows the higher PS-DS density and a higher number of scenes (Table 1) in order to analyze the datasets with higher spatial and temporal resolutions. Hence, the ERS-1/2 descending and Radarsat ascending data were exploited to monitor the time intervals, from 1992 to 2000, and from 2003 to 2010.

A-DInSAR displacement time series accuracy assessment has provided the detection of anomalous displacements recorded on 9 March 1997 and 16 July 2000 by the ERS-1/2 descending datasets, and on 10 December 2008, for the Radarsat ascending dataset. The anomalous displacement identified in the Radarsat dataset might be related to the snowfall occurred on the day of SAR acquisition. Consequently, these scenes were not included in the following analyses.

5.2. Identification of the Ground Motion Areas (Phase II)

Once the post-processing elaborations have been performed on the A-DInSAR data, the improved displacement time series have been exploited in order to delineate ground motion areas.

In order to achieve the spatio-temporal extraction of the principal components of motion, a matrix of the ERS-1/2 desc. and Radarsat asc. dataset was formed.

The results show that the PC1 explains 79% of the variance, and PC2 and PC3 explain, respectively, 4.7% and 1.8% of the variance in the ERS-1/2 dataset. In the Radarsat, 69% of the variance is explained by PC1, while 5.7% and 1.7% are explained by PC2 and PC3.

Figure 5a,b shows the spatial distribution of the principal components score units. The spatial pattern of the PC of the deformation highlights that the first and the third principal components of motion mainly affect the southwestern sector of the study area, while the second component is mainly localized in the northeastern zone. It is worth noting that the distribution of the components of motion is heterogeneous across the Oltrepo Pavese, and that the comparison between the ERS-1/2 and Radarsat data reveals that the spatial distribution of the components of motion detected during the period 1992–2000 is correlated with that of the period 2003–2010 (Figure 5a,b).

A visual inspection of the principal components eigenvectors (Figure 5c,d) highlights that in the ERS-1/2 and Radarsat datasets, the first component (PC1) corresponds to the long-term lowering of the Earth’s surface. The second corresponds to the long-term uplift ground motion, and the third PC highlights seasonal deformations.
Ground motion areas were delineated using the PC scores, applying the procedure described in Section 4, in order to focus the interpretation of the processes on significant sectors. The results reveal that 30%–40% of the ground motion areas, delineated using the PC1 of the ERS-1/2 and Radarsat measurements, is located in proximity of the town of Voghera and along the railway of Voghera–Pavia (Figure 6), and it records an average LOS velocity in the range from $-2$ to $-3.7$ mm/yr, in the period 1992–2000, and in the range of $-2$ to $-4.8$ mm/yr in the period 2003–2010. A distinct concentration of ground motion detected by PC1 is visible only in the time span of 1992–2000, along Emilia Road, from the town of Broni to the town of Redavalle, and along the railway from Broni to the town of Arena Po. The average LOS velocity ranges from $-6.4$ to $2$ mm/yr.

Figure 5. Principal component score maps for the period 1992–2000 (a) and 2003–2010 (b). Eigenvectors of the principal components (PC) of the ERS-1/2 desc. (c) and Radarsat asc. (d) datasets.
Figure 6. Ground motion areas detected using PC1, PC2, and PC3, and non-linear PS-DS of the ERS-1/2 desc. and Radarsat asc. data overlapped on the engineering geological units. TS of the deformational behavior detected at Voghera, Stradella–Monteacuto, and Casei Gerola are also reported. In the case of Voghera, it is represented by the TS linear regression line (dotted red line). Conversely, for Stradella–Monteacuto a 5° order polynomial regression line (dotted red line) is reported. For the TS detected at Casei Gerola, the red line is the moving average over a period of four months.

Ground motion areas mapped via PC2, of both ERS-1/2 and Radarsat data, are centered in the area between the town of Stradella and the town of Monteacuto (Figure 6). The average LOS velocity records an uplift maximum of 3.40 mm/yr, in the period 1992–2000, and 2.58 mm/yr, in the period 2003–2010.

Ground motion areas observed using PC3 are mainly located in the southern sector of Voghera and in proximity to Lungavilla, Codevilla, Casei Gerola, and Broni (Figure 6). In the second period, ground motion areas via PC3 are not present in the southern sector of Voghera, and the movements are mainly localized along the Coppa River (see location in Figure 1) and in proximity to Verrua Po. Seventy-four percent and 97% of the ground motion areas detected using PC3, respectively, for the ERS-1/2 and Radarsat data show an average LOS velocity in the range of ±1.5 mm/yr.

As was previously explained, an automated time series classification tool was used to distinguish uncorrelated, linear, and non-linear trends. The outcome of this step reveals that around 40% of the targets are classified as uncorrelated, both in the periods 1992–2000 and 2003–2010. Thirty-five percent and 22% of the targets are classified, respectively, as non-linear and linear during the time interval 1992–2000. In the following period (2003–2010), 28% and 27% of the targets are classified as linear and non-linear, respectively.

Furthermore, the results of the kinematic model analysis were combined with the ground motion areas detected using the PCA approach. Figure 6 shows the results of the overlapping of the ground motion areas, detected using PC1, PC2, and PC3, and the non-linear PS-DS characterized by a strong non-linearity (BICW > 1.2). Mainly, changes in trends in displacement time series were identified at
the end of 1999 and 2008 in the areas of Broni, Voghera, Lungavilla, and Casei Gerola, and, in most of the cases, the non-linear targets are superimposed on seasonal ground motion areas detected via PC3, which represent the seasonal components of motion.

Overall, the ground motion area detected at Voghera records a linear displacement time series that is active during both periods, while, in Broni, the ground motion occurs only in the period from 1992 to 2000.

The ground motion area centered in the north sectors of Voghera could be related to the presence of over 6–7 m of shallow thick clay deposits in the subsoil, and to groundwater exploitation for industrial supply. Unfortunately, the lack of historical data regarding the piezometric level variations did not allow us to analyze the effects of its variations on the land subsidence recognized in this area. The phenomenon is characterized by a linear trend for both 1992–2000 and 2003–2010, and no evident acceleration has been recognized (Figure 6).

The ground motion areas along the railway Voghera–Pavia and along the railway from Broni to Arena Po show a linear trend and rates of average LOS velocities in the range of −5 to 1 mm/yr for both periods. Given the close association with the local transport network, these ground motion areas could be due to the consolidation of the railroad embankment, and anthropogenic activity could have an influence on the surface settlement of these areas.

Areas of moderate uplift were observed via PC2 of the ERS-1/2 and Radarsat data, in the sector from Stradella to Monteacuto, where geomorphic evidence of an active emergent thrust was previously observed [43]. The terraced surfaces derived from topographic profiles, combined with inferences on the ages of the terraces of the examined area, were used to carry out estimations about the late quaternary uplift rates across the escarpment. The agreement between the distribution and trend of deformation with the uplift trend, found by previous authors in the same area, suggests that the deformation could be due to deep-ground motion related to tectonic movements.

5.3. Mechanism Recognition (Phase III)

A spatio-temporal analysis for mechanism recognition was performed, considering some factors and evidence, which might have relevance in explaining the patterns of ground motion, as observed in others ground motion areas by previous authors [44–46]: (i) the geotechnical properties of the deposits; (ii) the thickness of the superficial clayey soils and the hydrogeological setting; (iii) the distribution of damaged buildings; and (iv) land use change effects.

Different natural and man-induced processes were recognized such as swelling/shrinkage of clayey soils located over engineering geological Unit 8, and seasonal ground motions due to seasonal groundwater level variations, which are located over engineering geological Units 3 and 4.

Land subsidence due to the load of new buildings was observed over engineering geological Units 1, 2 and 3, and moderate tectonic uplift was recognized in the sector from Stradella to Monteacuto, where geomorphic evidence of an active emergent thrust was previously observed [43].

The main conditioning and triggering factors on ground motion are examined in detail in Section 6.

6. Discussion

The approach of this work was to identify areas with significant movements where the attention of scientists has to be focused by the use of A-DInSAR time series analysis. Ground motion areas were mapped on the basis of their peculiar ground deformation behaviors using a systematic and reproducible procedure.

Although various methodologies were implemented to map the ground motion areas using the average velocity detected by A-DInSAR data [12–17], ground motion areas mapping on the basis of A-DInSAR time series is still not a common practice in the scientific community.

While the average velocity may be useful to investigate physical processes characterized by linear trends, the same parameter is not efficient to detect and interpret ground motion areas affected by
non-linear and seasonal movements. In these cases, the use of the average velocity may result in misleading interpretations of the process.

Our intent is to propose a novel approach to take into account different deformational behaviors. Hence, the reliability of the results obviously depends on the quality of the dataset. However, the first phase of the procedure is focused on the time series check in order to correct problems due to phase unwrapping, regional unreal trends, and anomalous displacement detected on certain dates. In addition, if the whole dataset is affected by a relevant error, it can be easily detected by applying the PCA, being clearly visible as principal component of motion, and could be subtracted from the entire dataset.

This approach allowed us to detect ground motion areas due to different processes, even at low average LOS velocities in the range of ±1.5 mm/yr, which are affected by geohazards that may impact public safety.

In this section, the main conditioning and triggering factors on ground motion were discussed.

6.1. Comparison between Ground Motion and Engineering Geological Units

As regards geotechnical interpretation, the relationship between the engineering geological units and the ground motion was analyzed by computing the maximum and minimum of the displacement measured for each unit (Figure 7). Furthermore, the PS-DS density for each unit was extracted in order to take into account the distribution of measures. Figure 7a,b shows the results of the spatial analysis for the periods 1992–2000 and 2003–2010. It is worth noting that the extreme values of cumulated displacement observed in the first period are higher than those of the second one. A decrease of deformation was observed, and, in general, the cumulated displacement decreases from Unit 1 to 8. The PS-DS density reaches an average for the engineering geological units of 108 and 113 PS-DS/km², respectively, for the ERS-1/2 and Radarsat data.

Figure 7. Cumulated displacement (mm) and PS-DS density (PS-DS/km²) for the periods 1992–2000 (a) and 2003–2010 (b) for each engineering geological unit, which describes the geotechnical properties of the first 15 m of depth. See Figure 1 for the engineering geological unit map. Areas (km²) of ground motion detected by PC1, PC2, and PC3 for each engineering geological unit, for the time intervals 1992–2000 (c) and 2003–2010 (d).
Engineering geological Unit 1, which is composed of recent and actual alluvial deposits of the Po River, shows the lower PS-DS density (~40 PS-DS/km²), and, for this reason, the maximum cumulated land subsidence may be due to local movements.

Units 2, 3, and 4 are characterized by cohesive soils (50%–60% of the thickness) in the first five meters, and show an average PS-DS density of ~100 PS-DS/km². These units record a maximum cumulated land subsidence that varies from 100 mm to 200 mm in the first period, and an average value of 150 mm in the second one.

Unit 5 is composed of the alluvial fan of the Scuropasso River, with heterogeneous soils in the first 15 m of depth. As shown in Figure 7, this unit shows the highest PS-DS density (~300 PS-DS/km²) and the maximum cumulated land subsidence reaches ~51 mm for both periods.

Units 6, 7, and 8 are characterized by cohesive soils, predominant only in the first five meters of depth, and gravel and sand in the sub-soil. The PS-DS detected over these units is around 100 PS-DS/km² and the maximum cumulated land subsidence reaches values in the ranges of 100–180 mm and 50–100 mm, respectively, for the periods 1992–2000 and 2003–2010. Note that the maximum cumulated uplift of 159 mm was detected for Unit 6, in the period 1992–2000.

In addition, comparisons between the engineering geological units and the areas of ground motion detected using the principal component analyses were carried out for the monitored periods (Figure 7c,d). The results of the analyses show that the area of ground motion reaches around 23 and 17 km², respectively, in the periods 1992–2000 and 2003–2010. Therefore, the ground motion area decreases from the first period to the second one. It is worth noting that the study area is mainly characterized by PC3, which corresponds to seasonal deformation. Seasonal deformations may be due to seasonal aquifer compaction and/or swelling/shrinkage of clayey soils.

The PC1 ground motion that corresponds to the long-term lowering trend is mainly localized on Units 2 and 3 for both periods, while PC2 ground motion that corresponds to the uplift long-term trend is mainly localized on Units 1 and 6 for the entire monitored period.

6.2. Comparison between Ground Motion and Hydrogeological Settings

Another important factor that controls ground motion is the presence, and the thickness, of clayey soils, which may be susceptible to the consolidation process. A thickness map of the upper clayey-silty deposits was elaborated by Pilla et al. (2007) [21] by interpolating the thickness data available from around 490 water wells, performed in the Oltrepo Pavese (Figure 8a). This map allows us to analyze the relationship between the average LOS velocity and the thickness of the upper compressible deposits. The Oltrepo Pavese is characterized by the presence of a tabular clayey-silty cover, which increases from the northeast to the southwest. The thickness of the clayey-silty deposits ranges from 0 to 40 m, and the maximum thickness is observed in the southeastern sector of Voghera, and in the proximity of the towns Broni and Stradella (Figure 8b,c). Figure 8d,e shows the comparison between the average LOS velocity, detected by ERS-1/2 and Radarsat data, and the thickness of the upper clayey-silty deposits along the N–S and E–W sections (Figure 8a). The results reveal that the peaks of average LOS velocity are consistent with the maximum thickness of clayey-silty deposits. As a consequence, the thickness of the upper clayey-silty deposits seems to be a controlling factor of the rate of ground motion, as observed in others case histories, such as Murcia, Vega Media of the Segura River Basin, and Alto Guadalentín Basin in Spain [46–48], and along the Venice coast and Sibari plain in Italy [49,50].

Furthermore, the constant presence of a tabular clayey-silty coverage, which acts as a seal, is a crucial factor in determining the hydraulic condition of aquifer state. Indeed, the clayey-silty cover limits infiltration, and influences aquifer recharge, which occurs in correspondence with the coalescent fans originating from the deposition of the Apennine streams [21]. Analysis of the hydraulic condition role on ground motion was performed using the unsaturated zone thickness, measured in October 2005 (Figure 9). The unsaturated zone thickness detected in October 2005, was established by Pilla et al. (2007) [21] using a piezometric level obtained from around 60 water wells across the study area. For the computation of the unsaturated zone, the piezometric levels were subtracted to the base of
the clayey-silty cover. When the values are negative, the aquifer is confined and the flow is mainly horizontal. Conversely, when the values are positive, the aquifer is phreatic. Figure 9 shows the overlapping of the ground motion areas mapped using PC3 of the Radarsat data with the hydraulic condition of the aquifer in October 2005. The results reveal that the seasonal ground motions (PC3), which are located over engineering geological Units 3 and 4, occur where phreatic condition of the aquifer was observed. The role of the seasonal fluctuations of the groundwater level on the ground motion areas is clear. Seasonal ground motions related to fluctuations in the groundwater level were previously recognized by other authors in Santa Clara Valley and Los Angeles Basin in the United States [51,52]. Meanwhile, the ground motion areas located over engineering geological Unit 8 are due to other processes, which are explained in Section 6.3.

**Figure 8.** (a) Thickness of the upper clayey-silty deposits (modified from Pilla et al., 2007 [21]). Hydrogeological sections 1 (b) and 2 (c), simplified from Cavanna et al. (1998) [17]. Relationship between the average LOS velocity detected by ERS-1/2 asc. and Radarsat (RSAT) desc. sensors, and the thickness of the upper clayey-silty deposits along sections 1 (d) and 2 (e).
6.3. Comparison between Ground Motion Areas and Damaged Buildings

Most locations in the Oltrepo Pavese experienced damage to buildings due to swelling and shrinkage of clayey soils, respectively, during wet and dry periods. In most cases, the damaged buildings are single-story family residences, built on conventional shallow strip concrete footings. The damage was mainly detected over engineering geological Unit 8, where clayey soils susceptible to swelling/shrinkage were previously identified [28]. Meisina et al. (2006) [30] implemented a damaged building database.

Here, comparisons between the damaged building distribution and ground motion areas were performed. The results of the analyses highlight that the occurrence of damaged buildings is mainly localized over the ground motion areas detected using PC3. Therefore, this is consistent with the hypothesis that the seasonal deformational behavior detected by the third component of ground motion in engineering geological Unit 8 may be explained by the swelling/shrinkage of clayey soils.

In Codevilla (see location in Figure 6), located over engineering geological Unit 8, the higher density of damaged buildings (12 damaged buildings in an area of around 0.6 km²) corresponds to ground motion areas delineated using the PC3 of ERS-1/2 and Radarsat data (Figure 10a). In this area, the shallow alluvial deposits of the first 8 m of depth exhibit swelling potentials from medium to very high [30].

The swelling potential appears to be in good agreement with the displacement time series. Average LOS displacement time series was computed in the buffer zone of 100 m from the damaged building with code 425 in order to consider a significant number of measurements (around 20 PS-DS). The results reveal that the seasonal component of motion detected in the TS is directly correlated with the effective rainfall detected at the nearest pluviometric station, located in Voghera (Figure 10a,b). In particular, significant damage, which consists of cracks, appeared in this building in August 1999 (crack width 5–10 mm). The pre-damage period from 1998 to 1999 was recognized as a drought period [30], and the average effective rainfall from May to November of 49.45 mm corresponds to around −6.14 mm of ground motion. In February and March 2001, partial pile underpinning was carried out as a remedial measure for this building. In the following drought period, 2003–2007, the variability of the effective rainfall corresponds to there being no clear cyclic displacement time series.
However, the average effective rainfall from May to November, of 62.95 mm, corresponds to around \(-1.27\) mm of ground motion.

![Figure 10](image)

**Figure 10.** (a) Ground motion areas in Codevilla, detected by the ERS-1/2 and Radarsat PC3, and the distribution of damaged buildings detected by Meisina et al. (2006) [30]. (b) Average LOS displacement time series in the buffer zone of 100 m from the damaged building (code 425) versus effective rainfall measured at the Voghera weather station. The dotted red line is the polynomial regression trend (5th order) of the displacement time series. The equation for polynomial regression has been also indicated.

The results confirm that the ground motion areas detected at Codevilla via PC3 are due to swelling/shrinkage phenomena triggered by the variation of water content in these soils.

Overall, Codevilla seems to be characterized by an average LOS velocity range that is commonly defined as stable \(\pm 1.5\) mm/yr.) during the periods 1992–2000 and 2003–2010; however, in this case, the seasonal deformation is significant, and provides evidence of damage, such as cracks in buildings, dislocations, and the breaking of walls, which were recognized by Meisina et al. (2006) [30,53]. In particular, the damaged buildings required the installation of micropile construction up to 8 m in depth, in order to reinforce structures.

### 6.4. Comparison between Ground Motion and Land Use Change Effects

In order to define the influence of land use change on ground motion, a visual inspection was performed by comparing orthophoto 1988 and 2000 [54] and orthophoto 2003 and 2007 provided by the Lombardia Region, with PC1 ground motion areas. In addition, Corine Land Cover changes
Remote Sens. 2016, 8, 686

(CLIC) from 1990–2000 and 2000–2006, implemented by the Institute for Environmental Protection and Research (ISPRA) [55,56] that match our satellite data, were exploited. These layers were created by using a working scale of 100,000; therefore, the land cover cannot be mapped in all its complexity, and the definitions of the units were established by the combination of surfaces, which are homogeneous to a greater or lesser degree, whatever the scale used. From the total of the ground motion areas delineated by PC1 of the ERS-1/2 and Radarsat data, only two and three areas seem to be influenced by land use changes, respectively. More precisely, in the period 1992–2000, land use change from non-irrigated arable land to discontinuous urban fabric, which occurred over engineering geological Units 1 and 2, matches with PC1 ground motion areas detected by ERS-1/2 data. In the period 2000–2006, land use change from non-irrigated arable land to construction sites occurred over engineering geological Unit 3, which matches the PC1 ground motion area detected using Radarsat data.

It worth noting that these ground motion areas, detected during the period 1992–2000, are not present in the period 2003–2010. Therefore, the stress increment produced by a building’s load seems to have dissipated in the first period. Figure 11 shows ground motion detected by the PC1 of ERS-1/2 data in Voghera. The average displacement time series of the PS-DS in the ground motion areas highlight the decrease of the average LOS velocity from $-3.15 \text{ mm/yr}$ during 1992–2000 to $-0.80$ for 2003–2010 (Figure 11c).

![Figure 11](image)

**Figure 11.** Location of the representative ground motion area detected using PC1 of the ERS-1/2 sensors in Voghera, superimposed on the land use change from 1990–2000, and on the Ortophotos acquired from 1988 (a) and 2000 (b). (c) Average displacement time series detected in the ground motion area. The dotted red line is the polynomial regression trend (2° order) of the displacement time series. The equation for polynomial regression has been also indicated.
7. Conclusions

A challenge to ground motion studies is the disentanglement of different phenomena and to assess their relative contributions. Indeed, specific causes of deformation phenomena in lowlands are presently unknown due to the interaction of either deep (“geological”) or superficial causes (withdrawal of water and gas, soil consolidation, load-induced compaction), which take place at different spatio-temporal scales and of which the mixed effects must be decorrelated. This paper demonstrated the ability of a novel methodology to detect and disentangle the contributions of different geohazards, by analyzing multi-sensors and multi-temporal A-DInSAR data. The principal novelty of the proposed approach is the exploitation of a full time series, which enables the detection of ground motion areas affected by linear, non-linear, and seasonal movements. The proposed methodology consists of three phases:

1. Post-processing elaborations; disentanglement of the vertical and horizontal components of motion and displacement time series accuracy assessments, performed by selecting the most coherent (>0.9) targets with an average LOS velocity in the range ±0.5 mm/yr, where no significant movements are expected.

2. Identification of ground motion areas; application of two approaches for automatic A-DInSAR time series analyses. First, principal components analysis of the time series is performed, implementing a matrix organization of location versus time (T-mode). Then, PS-time software is used to find the kinematic behavior of the time series (i.e., uncorrelated, linear, and non-linear trend). Finally, ground-motion areas were identified using a spatial cluster with a buffer zone of 50 m around the targets, characterized by a significant correlation with the principal components of motion. The results of both approaches are overlapped to distinguish linear and non-linear ground motion areas.

3. Mechanism recognition: Comparison between ground motion areas and ancillary data (i.e., geological, geotechnical, hydrogeological, and land use information) in order to interpret the driving forces of the subsidence mechanisms.

The methodology was tested in Oltrepo Pavese (Italy), which is a representative site of moderate subsidence of the Po Plain, using ERS-1/2 and Radarsat data covering a time span of around 20 years. Results confirm its suitability for the detection of the temporal evolution of deformation patterns due to different geohazards, not recognizable by conventional analyses of the average LOS velocity alone (i.e., seasonal components of motion). Three deformational behaviors were detected: linear, non-linear (accelerations and decelerations of the movements), and seasonal components of motion. The comparison between a kinematic model of the ground motion areas and hydrogeological and engineering geological data allow for an understanding of the causes of motions that are often superimposed. The main natural and man-induced processes are swelling/shrinkage of clayey soils, land subsidence due to the load of new buildings, moderate tectonic uplift, and seasonal ground motion due to seasonal groundwater level variations. Overall, the cumulated displacements observed in the 1992–2000 period are higher than those detected in the period 2003–2010, and a decrease of deformation was observed. The findings in this work confirmed that the combination of A-DInSAR time series analysis and hydrogeological data is capable of supporting knowledge of the relationship between groundwater level fluctuations and deformations, and of providing fundamental information about the aquifer state conditions.

The proposed methodology can be easily applied to other geological contexts, using different A-DInSAR data. The reliability of the results depends on the quality of the dataset. New opportunities will be provided by the improvement of the time series acquired by new sensors of the COSMO-SkyMed and Sentinel missions.

The procedure will also allow for specifying the priority area of where to address prevention activities, in order to optimize the costs and benefits and to draw a management plan of land use and groundwater resources, at national and regional scales. The procedure may also be used in land
subsidence studies to discern the contributions of the seasonal components of motion from others processes, such as soil consolidation and solid and fluid extraction or injection. The approach will be tested, in the future, in geological contexts characterized by other geohazards, such as landslides.

Acknowledgments: The A-DInSAR data we used were provided by Lombardia Region (Italy) in the framework of the project Report IRER n 2010B009: “Interpretazione geologica e idrogeologica dei dati radar derivanti dal Piano Straordinario di Telerilevamento nazionale (PST) e da dati Radarsat con particolare riferimento alle province di Lecco e Pavia”. The data were processed by Tele-Rilevamento (TRE). The authors would like to thank the reviewers for their valuable comments and suggestions to improve the quality of the paper.

Author Contributions: Roberta Bonì analyzed the A-DinSAR data and prepared the manuscript. Giorgio Pilla provided hydrogeological data and supported their interpretation. Claudia Meisina provided the engineering geodetical data and provided guidance and support throughout the research process. All authors co-wrote and reviewed the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References
2. Cigna, F.; Jordan, H.; Bateson, L.; McCormack, H.; Roberts, C. Natural and anthropogenic geohazards in greater London observed from geological and ERS-1/2 and ENVISAT persistent scatterers ground motion data: Results from the EC FP7-SPACE PanGeo Project. Pure Appl. Geophys. 2014. [CrossRef]


36. Raspini, F.; Cigna, F.; Moretti, S. Multi-temporal mapping of land subsidence at basin scale exploiting Persistent Scatterer Interferometry: Case study of Gioia Tauro plain (Italy). *J. Maps* **2012**, *8*, 514–524. [CrossRef]


Complex Deformation Monitoring over the Linfen–Yuncheng Basin (China) with Time Series InSAR Technology

Cheng-sheng Yang1,2, Qin Zhang1,*, Qiang Xu2, Chao-ying Zhao1, Jian-bing Peng1 and Ling-yun Ji3

1 College of Geology Engineering and Geomatics, Chang'an University, Xi'an 710054, China; ycsgps@163.com (C.Y.); zhaochaoaying@163.com (C.Z.); dicexy_1@chd.edu.cn (J.P.)
2 State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu 610059, China; xq@cdut.edu.cn
3 Second Crust Monitoring and Application Center, CEA, Xi'an 710054, China; dinsar010@163.com

* Correspondence: zhangqinle@263.net.cn; Tel.: +86-029-8233-9261

Academic Editors: Zhenhong Li, Roberto Tomás, Magaly Koch and Prasad S. Thenkabail

Received: 16 January 2016; Accepted: 21 March 2016; Published: 28 March 2016

Abstract: The Linfen–Yuncheng basin is an area prone to geological disasters, such as surface subsidence, ground fissuring, fault activity, and earthquakes. For the purpose of disaster prevention and mitigation, Interferometric Synthetic Aperture Radar (InSAR) was used to map ground deformation in this area. After the ground deformation characteristics over the Linfen–Yuncheng basin were obtained, the cross-correlations among regional ground subsidence, fault activity, and underground water level were analyzed in detail. Additionally, an area of abnormal deformation was found and examined. Through time series deformation monitoring and mechanism inversion, we found that the abnormal deformation was related mainly to excessive groundwater exploitation.

Keywords: Complex deformation; time series InSAR; inversion of mechanism; underground water; fault activity

1. Introduction

The Shanxi fault belt basin is a Cenozoic fault basin that has strong earthquake activity. The Linfen–Yuncheng basin is located in the south-central region of the Shanxi fault belt basin, where the crustal structure is especially complex and seismic activity is very strong [1,2] (Figure 1). In total, 10 earthquakes with $M_s \geq 5.0$ have been recorded since A.D. 1177 ($M_s$ stands for surface wave magnitude): one event with $M_s \geq 8.0$, one event with $7.0 \leq M_s < 8.0$, two events with $6.0 \leq M_s < 7.0$, and six events with $5.0 \leq M_s < 6.0$ [3]. According to modern seismic network records, moderate–small earthquakes have been more frequent in this area in the recent past [3,4] (Figure 1). At the same time, ground subsidence, ground fissuring, fault activity, and abnormal deformation have been observed and have received substantial attention from the local government [5,6]. For the purpose of disaster prevention and mitigation, there is an urgent need to characterize the ground deformation characteristics within the basin.

Conventional ground-based geodetic techniques, such as GPS and leveling, have difficulties in detecting more detailed and comprehensive ground deformation due to their low spatial resolution. Although terrestrial laser scanning can present very high spatial resolutions, it hard to be carried out ground deformation monitoring in a wide range. Interferometric Synthetic Aperture Radar (InSAR) has demonstrated potential for high-density spatial mapping of ground displacement associated with earthquake, volcanic, and other geologic processes [7–11]. However, the application of conventional InSAR is limited by several inherent factors, such as spatial-temporal decorrelation, DEM inaccuracy,
phase unwrapping error and atmospheric delay. Nevertheless, all of these issues can be addressed with the InSAR time series technique of the Small BAseline Subset (SBAS) [13–17], which combines interferograms by using a threshold for the temporal and spatial baseline criterions. The SBAS InSAR technique can also eliminate or mitigate the possibility of phase unwrapping errors and atmospheric delay errors through temporal high-pass and spatial low-pass filtering of interferogram. Finally, the time series deformation information of the coherent points can be obtained.

Thus, in this study, we examined basin ground deformation using Envisat ASAR images. The Stanford Method for Persistent Scatterers (StaMPS) SBAS InSAR technique was used to obtain the time series surface deformation [17,18]. The ground deformation characteristics over the Linfen–Yuncheng basin were obtained, and the cross-correlation among the regional ground subsidence, fault activity, and underground water level were analyzed in detail. At the same time, an abnormal deformation was identified and analyzed.

The StaMPS InSAR technique is presented in Section 2 of this paper. The data collection and processing procedures are described in Section 3. The results of the study and some discussion are provided in Section 4, and the abnormal deformation is discussed in Section 5.

2. StaMPS Method

StaMPS is one of the most mature InSAR time-series analysis software packages. StaMPS uses a statistical relationship between amplitude stability and phase stability that makes consideration of amplitude useful for reducing the initial number of pixels for phase analysis. The amplitude dispersion index, defined by [19], can be written as

$$D_A = \frac{\sigma_A}{\mu_A}$$  \hspace{1cm} (1)
where $\sigma_A$ is the standard deviation of the amplitude values and $\mu_A$ is the mean of a series of amplitude values. The amplitude dispersion index threshold value that is used is commonly higher, typically on the order of 0.4, than the value of 0.12 recommended by [19]. As a result, a large number of pixels are selected as Persistent Scatterers (PS) initial candidates (PSCs). The PSCs are tested for phase stability with an indicator $\gamma_x$, which is defined based on the temporal coherence and can be used to evaluate whether the pixel is a PS:

$$
\gamma_x = \frac{1}{N} \left| \sum_{i=1}^{N} \exp \left( j(\phi_{\text{int},x,i} - \overline{\phi_{\text{int},x,i}}) - \Delta \phi_{h,x,i} \right) \right|
$$

where $N$ is the number of interferograms and $\overline{\phi_{\text{int},x,i}}$ is an estimate of the wrapped phase $\phi_{\text{int},x,i}$ of the $x$th pixel in the $i$th flattened and topographically corrected interferogram $\Delta \phi_{h,x,i}$ is an estimate of the DEM error. After every iteration, the root-mean-square change in coherence $\gamma_x$, determined as in Equation (2), is calculated. When the change is less than a specified threshold, the solution has converged and the algorithm stops iterating. Then, the selected pixels are considered as PS pixels, considering their amplitude dispersion, as well as $\gamma_x$ (see [18] for details).

Once the PSs have been selected, their phase are corrected for spatially uncorrelated look angle (SULA) error by subtracting the estimated values, that is the DEM error [20]. As long as the density of PS is such that the absolute phase difference between neighboring PSs, after correction for estimated SULA error, is generally less than $\pi$, the corrected phase values can be unwrapped. Optionally, after unwrapping, a high-pass filter can be applied to the unwrapped data in time followed by a low-pass filter in space in order to remove the remaining spatially correlated errors (atmosphere and orbit errors). Finally, subtracting this signal leaves essentially deformation and spatially uncorrelated errors that can be modeled as noise.

The method is able to characterize the temporal model of deformation, rather than using an assumed model. StaMPS is successful at finding PS pixels in both urban and nonurban areas, which makes it applicable in areas covered by forest and vegetation.

3. Data Collection and Processing

In total, eight C-band Envisat ASAR images with track 2347 in stripe mode, acquired from February 2009 to October 2010, were used to study ground deformation in the Linfen–Yuncheng basin (red frame in Figure 1). Underground water level data also were collected for this study. ASTER GDEM (Version 2) with resolution of 1 arc-second (30 m) was used as an external DEM to remove the topographic phase from the interferograms, and ESA DORIS precision orbits data for Envisat satellites were employed to remove the reference phase and orbital bias from the differential interferograms.

The StaMPS SBAS InSAR technique was used in this experiment [17, 21]. First, to ensure the reliability of the deformation measurements, SAR interferometric combinations with a temporal baseline of less than 300 days and a spatial baseline of less than 300 m were outlined. Accordingly, 15 interferograms were generated (Figure 2) to study the ground deformation within the basin. After spatially filtering each interferogram, the spatially uncorrelated DEM error was iteratively optimized by setting a maximum error of 8 m. The standard deviation of each coherent pixel was estimated for all of the interferograms. Any pixels with a standard deviation larger than 1.0 radian were eliminated to refine the coherent points [22]. Considering the deformation gradient within the basin, the unwrapped grid size was resampled to 50 m. Atmospheric phases were separated from each differential interferogram by filtering the spatial and temporal domains according to their spatial and temporal correlations.
4. Results and Discussion

Ground deformation results over the Linfen–Yuncheng Basin were assessed from 2009 to 2010. The annual average deformation map is shown as Figure 3, in which positive and negative signs represent displacement toward and away from the satellite line-of-sight (LOS), respectively. According to the geological survey, the western mountainous area of Linfen is with hard rock bottom and close to the stable Ordos block tectonic units. So we selected an area about 10 km × 10 km as the reference region, which was represented in the Figure 3 as a five-pointed star marked area. The maximum displacement is from −62 to 20 mm/a. According to the results, the internal and marginal areas of the basin have different deformation characteristics. The internal basin showed land subsidence, but both sides of the basin showed relative uplift. Especially, at the eastern margin of the basin, there is an obvious difference of deformation on the two sides of the ZhongTiaoShan piedmont fault. From the local perspective, the tectonic units have different deformation rates, which show a gradually increasing trend from north to south.

4.1. Deformation Characteristics Analysis

As is well known, the stratification of the lower troposphere may lead to topographically correlated propagation delays in interferometric phase (hereafter called stratified delay). Hence, the stratified delay can be approximately considered as a linear function of the terrain, as has been proven by earlier works [23–28]. After the temporal and spatial filtering in the SBAS–InSAR processing, the influence of tropospheric effects on the results comes mainly from the stratified delay. Because the deformation shares an obvious boundary with the basin topography (see Figure 3), we needed to analyze whether there was a topographically correlated stratified delay in the InSAR results; so we chose two regions, one in the eastern side and one in western side of the basin (shown as A and B in Figure 3, respectively), where analyzing the correlation between deformation rate and topography. The results showed that there was no obvious correlation between deformation rate and topography (Figure 4A,B). Thus, we concluded that the stratified delay had little effect on the monitoring results in this study.
According to the results shown in Figure 3, there are two obvious centers of subsidence and an uplift zone located near the towns of Xiaxian, Xinjiang and Hejin, respectively. Their locations are depicted in Figure 3. In order to analyze the deformation characteristics of the points that are marked...
with triangle in Figure 3, time series deformation results were extracted (shown in Figure 5). Based on Figure 5, the cumulative subsidence values around Xiaxian and Xinjiang through nearly one and one-half years were 100 mm and 70 mm, respectively, whereas the surface uplift in the Hejin region reached 40 mm. The uplift deformation in Hejin is discussed further in Section 5.

4.2. Fault Activity Analysis

According to the survey, there are more than 10 active faults in the study area. Among these, the ZhongTiaoShan piedmont fault, Emei–Zijinshan fault, and Mingtiaogang southeast fault are the three main active faults. In order to analyze the impacts of the faults on the regional deformation, two profiles were extracted along P1–P1’ and P2–P2’, as shown in Figure 3, and the results are shown in Figure 6. The terrain profiles are also plotted in Figure 6.

![Figure 5. Cumulated subsidence curves. "RefP" is the reference point, "XinJ" is the town of Xinjiang, "XiaX" is the town of Xiaxian, and "HeJ" is the town of Hejin.](image)

![Figure 6. Profiles across the faults: (A,B) terrain profiles; and (C,D) surface deformation rate profiles. The dotted lines show the locations of faults.](image)
Remote Sens. 2016, 8, 284

Profile P1–P1’, crossing the FI-2 and FI-3 faults, is shown as Figure 6C. The ground subsidence shows an obvious difference between the two sides of the faults. The locations of the faults are showed by the dotted lines in the figure. The middle region of the two faults shows subsidence, but the lateral sides show uplift. There is about 15 mm/year of differential subsidence. We also extracted profile P2–P2’ perpendicular to faults FI-1 and FI-2 in Figure 3 (Figure 6D). The deformation difference on both sides of the faults is not as obvious, which shows that the influence of faults on ground subsidence varies in different areas.

As shown above, fault activity and the surrounding ground subsidence have a certain spatial correlation in the Linfen–Yuncheng basin. Deformation occurs within the fault-controlled basin, but no deformation occurs in the surrounding mountainous region without faults.

4.3. Relationship Analysis between Underground Water and Ground Subsidence

The Linfen–Yuncheng basin is surrounded by mountains. Its annual average rainfall is 572.5 mm, and the annual average evaporation is 1148.0 mm, twice the former. As a result, underground water is the main source of drinking water in the region.

Research has demonstrated that groundwater exploitation is an important factor in ground subsidence [11,29]. In order to study the relationship between ground subsidence and groundwater exploitation in the area, groundwater level data were collected by pressure-type water level meters, and the spatial distribution of the sensors was shown as in Figure 7. Two profiles of groundwater level, along lines P3–P3’ and P4–P4’ in Figure 7, were extracted (Figure 8A,B), along with the corresponding deformation values (Figure 8C,D), respectively.

![Figure 7. Groundwater level contours and groundwater level meter locations.](image)

It is worth noting that the two profiles are located in the groundwater level upheaval area, which is conducive to investigating the relationship between underground water level and ground subsidence. As shown in Figure 8, although the ground subsidence along P3–P3’ has small fluctuations, it has a trend that is generally similar with the trend of groundwater level. This similarity shows that the ground subsidence in the Xinjiang area is affected to a certain extent by the underground water level. However, this effect is not obvious in the Linfen region (Figure 8). Due to differences of soil physical...
and mechanical properties, the ground subsidence also shows some differences under the case of
groundwater overexploitation. The Xinjiang–Xiaxian area is part of the Fen River floodplain, which
has obvious plastic characteristics, and the degree of preconsolidation is relatively low. With the
overexploitation of groundwater, pore water pressure was reduced, which inevitably led to ground
subsidence. However, the Linfen region is part of the loess tableland, and the groundwater content
varies widely, leading to the weak correlation between groundwater and land subsidence.

Figure 8. Profiles of groundwater level and ground deformation: (A,B) groundwater level profiles
along P3–P3’ and P4–P4’ (Figure 7); and (C,D) respective deformation profiles along the same lines.

5. Hejin Abnormal Deformation Analysis

The Linfen–Yuncheng basin, where moderate–small earthquakes have occurred frequently in
the recent past, is the only area in Shanxi Province that has incurred an $M_s$ 8.0 earthquake. Hence,
the abnormal deformation that occurred in this area particularly attracted our attention. The InSAR
results showed that the abnormal deformation is located in Hejin, south of the LuoYunShan fault
(Figure 3). Exploring the process and the cause of this abnormal uplift is very important to an analysis
of the seismic risk of the region.

5.1. Interferogram Analysis

In order to obtain the process of the uplift deformation, we collected 18 scenes of ENVISAT ASAR
standard images with track 2347 (blue frame in Figure 1), covering the Hejin area and spanning the
years of 2003–2010 (there are no archived images for 2007 and 2008). Using these data, we began SBAS
InSAR processing. ASTER GDEM Version 2 with resolution of 1 arc-second (30 m) was used as an
external DEM to remove the topographic phase from the interferograms. Due to the notable temporal
gap (no images for 2007 and 2008), we divided the total images into two subsets and carried out the
processing separately. After careful data processing, 18 interferograms with good coherence were
generated (Figure 9). We show eight typical interferograms, spanning several months to several years,
in Figure 10. From these interferogram patterns, we developed the following observations.

1. Fringes that persist in time (e.g., Figure 10A–H) show deformation anomalies. These fringes are
unlikely to be atmospheric artifacts because the interferograms were produced from independent
SAR images. Additionally, the signals cannot be attributed to DEM errors because the baselines
of these interferograms are short, making them insensitive to any plausible errors in the DEM.
2. The deformation patterns are different during the periods spanning 2003–2006 (Figure 10A–D)
and 2009–2010 (Figure 10E–H). The color changes from blue–red–yellow–blue along the direction
of the arrow during 2003–2006 (Figure 10A–D). However, the color change shows the opposite
trend along the direction of the arrow (yellow–red–blue–yellow) during 2009–2010 (Figure 10E–H). Another phenomenon of note is that the deformation center moved to the west during 2009–2010 (Figure 10E–H).

Figure 9. Configurations of InSAR pairs used for the abnormal deformation monitoring.

Figure 10. (A–H) Interferograms labeled with period (date format is yyyyymmdd). The satellite flight direction and radar look direction are labeled in (A). Each fringe (full color cycle) represents 28 mm of range change between ground and satellite. Areas that lack interferometric coherence are uncolored.

5.2. Deformation and Mechanism Inversion

In order to understand the uplift deformation process, we obtained the time series results of the deformation center (Figure 11) using the SBAS method for the data processing. As determined from the results, the deformation is divided into two stages, a subsidence stage from December 2003 to February 2006, with maximum subsidence of 57 mm, and an uplift stage from February 2009 to November 2010, with maximum uplift of 38 mm. Because of the absence of images from 2007 to 2008, we were unable to obtain the deformation process during this time period.
Figure 11. InSAR Line-Of-Sight (LOS) deformation time series for the HeJin area.

We assumed that the deformation was caused by a volume change beneath the northern Hejin area. In order to explain the InSAR-derived deformation field, we tried a uniformly opening sill (i.e., crack) embedded in an elastic half-space to model all the interferograms with reasonably good coherence individually [30]. Eight parameters defined the sill: length, width, depth, strike, dip, opening, and location (two parameters). In the model, we introduced linear terms to account for any possible phase ramp due to uncertainties in satellite positions [31]. We used the downhill simplex method and Monte Carlo simulations to estimate the optimal parameters and their uncertainties [32] and the root mean square error (RMSE) between the observed and modeled interferograms as the prediction-fit criterion. Figure 12 shows two examples with observed (Figure 12A,D), modeled (Figure 12B,E), and residual (Figure 12C,F) interferograms for the sill models. The models fit each of the observed interferograms reasonably well. Because the subsidence and uplift may have been caused by different mechanisms, we separately calculated the parameters of the sill models for seven deflationary and four inflationary interferograms. In other words, we achieved the parameters of the two sill models by averaging all the results for seven deflationary and four inflationary interferograms separately. This strategy provided a first-order approach to reducing errors in the modeling. Table 1 shows the averaged parameters of the sill sources for the seven deflationary and four inflationary interferograms. All of the model parameters are well constrained according to the uncertainties shown in Table 1. We attribute the goodness of fit to the averaging of the many interferograms. The best-fit model sill for the deflationary interferograms is 5.2 km long, 2.5 km wide, 1.5 km deep, and has a strike of N104°E, and the best-fit model sill for the inflationary interferograms is 5.4 km long, 1.6 km wide, 1.4 km deep, and has a strike of N78°E.

Figure 12. (A–F) Observed interferograms; synthetic interferograms for sill sources that best fit the individual interferograms; and residual interferograms, which are the differences between the observed (left) and the modeled (middle) interferograms.
Table 1. Parameters for the best-fitting model sill (see text for details). Uncertainties correspond to 95% confidence.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Deflation</th>
<th>Inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length (km)</td>
<td>5.2 ± 0.9</td>
<td>5.4 ± 0.5</td>
</tr>
<tr>
<td>Width (km)</td>
<td>2.5 ± 0.4</td>
<td>1.6 ± 0.2</td>
</tr>
<tr>
<td>Depth (km)</td>
<td>1.5 ± 0.5</td>
<td>1.4 ± 0.3</td>
</tr>
<tr>
<td>Strike (°)</td>
<td>104.1 ± 1.4</td>
<td>78.3 ± 8.7</td>
</tr>
<tr>
<td>X (km)</td>
<td>7.7 ± 0.3</td>
<td>7.0 ± 1.0</td>
</tr>
<tr>
<td>Y (km)</td>
<td>3.7 ± 0.3</td>
<td>4.7 ± 0.1</td>
</tr>
<tr>
<td>Open (mm)</td>
<td>−34.5 ± 14.5</td>
<td>38.6 ± 9.8</td>
</tr>
</tbody>
</table>

5.3. Discussion of Abnormal Deformation

The differences of location and depth of the sill sources beneath the northern Hejin area between the results of the subsidence and uplift interferograms are not significant at the 95% confidence level (Table 1), so we attribute both processes to the same source. The dimensions of the two sill models are also similar.

We speculate that the deformation was caused by underground water withdrawal and influx. First, no active faults or volcanoes are located in the northern Hejin area. Second, according to the patterns shown in the interferograms (Figure 10), it does not seem that the deformation was caused by tectonic movements. Third, the ground surface, as shown by the high-resolution remote sensing images, is covered mainly by crops, and there are some chemical plant distribution here. Therefore, we speculate that underground water was overextracted during 2003–2009. According to our field investigation, more than 50% of the small- and medium-sized enterprises were closed during 2009. After 2009, the local enterprises and residents stopped extracting underground water excessively, which caused the ground to rebound, i.e., subsidence to uplift.

6. Conclusions

Linfen–Yuncheng basin has suffered from severe geo-hazards, including large scale land subsidence and small scale fault activity, which has caused serious infrastructure damages and property losses. High accuracy InSAR results can show the relationships among the different scale deformation. In this study, ground deformation, fault activity, and the mechanism of abnormal deformation over the Linfen–Yuncheng basin were analyzed using InSAR deformation results. The influences of underground water on ground subsidence were analyzed, as were the controlling effects of the fault distribution on the deformation trend. The following conclusions were drawn.

First, there are two obvious centers of subsidence and one uplift zone within the Linfen–Yuncheng basin located in the towns of Xiaxian, Xinjiang, and Hejin, respectively. Second, deformation occurs within the fault-controlled basin, and the distribution of faults controls the ground subsidence trend. Third, excessive exploitation of underground water is one of the main causes of land subsidence in the Linfen–Yuncheng basin, and the effects of underground water level on ground subsidence show differences in different regions, which are influenced mainly by formation conditions. Fourth, an area of abnormal deformation was found and this deformation presented a reversal of trend, from ground subsidence to uplift, in about 2009. The abnormal deformation was mainly related to excessive groundwater exploitation.

Acknowledgments: This research was funded jointly by the National Natural Science Foundation of China (NSFC) (Nos. 41304016, 41372375, 41274004, and 41274005), the Special Earthquake Research Project of the China Earthquake Administration (No. 201508009), the National Program on Key Basic Research Project (973 Program) (Grant No. 2014CB744703), the Fundamental Research Funds for the Central Universities (310826161017), and the State Key Laboratory of Geohazard Prevention and Geoenvironment Protection (SKLP2013K013). The SAR data used in this study were supplied under the Category 1 project (C1F28413, C1F28193). We also thank JPL/Caltech for ROI PAC, TU-Delft for DORIS, and Andy Hooper for StaMPS. We are grateful for the helpful corrections and suggestions made by four anonymous reviewers.
Author Contributions: Qin Zhang and Chao-ying Zhao conceived and designed the experiments; Cheng-sheng Yang performed experiments and drafted the manuscripts; Qiang Xu and Jian-bing Peng contributed to InSAR data analysis; Ling-yun Ji contributed to model calculations.

Conflicts of Interest: The authors declare no conflict of interest.

References
5. Chang’an University. Monitoring Land Subsidence and Ground Fissures over Fenwei Basin by using InSAR and GPS Techniques; Chang’an University: Xi’an, China, 2014.


InSAR Time Series Analysis of Natural and Anthropogenic Coastal Plain Subsidence: The Case of Sibari (Southern Italy)

Giuseppe Cianflone 1,*, Cristiano Tolomei 2, Carlo Alberto Brunori 2 and Rocco Dominici 1

1 Department of Biology, Ecology and Earth Sciences, University of Calabria, Via P. Bucci, cubo 15B, 87036 Arcavacata di Rende (Cs), Italy; rocco.dominici@unical.it
2 Istituto Nazionale di Geofisica e Vulcanologia (INGV), Via di Vigna Murata, 605, 00143 Rome, Italy; cristiano.tolomei@ingv.it (C.T.); carloalberto.brunori@ingv.it (C.A.B.)
* correspondence: giuseppe.cianflone@unical.it; Tel.: +39-338-615-5605

Abstract: We applied the Small Baseline Subset multi-temporal InSAR technique (SBAS) to two SAR datasets acquired from 2003 up to 2013 by Envisat (ESA, European Space Agency) and COSMO-SkyMed (ASI, Italian Space Agency) satellites to investigate spatial and temporal patterns of land subsidence in the Sibari Plain (Southern Italy). Subsidence processes (up to ~20 mm/yr) were investigated comparing geological, hydrogeological, and land use information with interferometric results. We suppose a correlation between subsidence and thickness of the Plio-Quaternary succession suggesting an active role of the isostatic compensation. Furthermore, the active back thrusting in the Corigliano Gulf could trigger a flexural subsidence mechanism even if fault activity and earthquakes do not seem play a role in the present subsidence. In this context, the compaction of Holocene deposits contributes to ground deformation. Despite the rapid urbanization of the area in the last 50 years, we do not consider the intensive groundwater pumping and related water table drop as the main triggering cause of subsidence phenomena, in disagreement with some previous publications. Our interpretation for the deformation fields related to natural and anthropogenic factors would be a comprehensive and exhaustive justification to the complexity of subsidence processes in the Sibari Plain.

Keywords: SBAS-InSAR; Sibari Plain; subsidence; geology; anthropic processes

1. Introduction

Many coastal and delta plains worldwide are affected by land subsidence phenomena [1,2], which often involve inhabited areas causing conspicuous economic costs.

Examples come from different geodynamic, geological, climatic and social contexts. In Italy, subsidence due to a combination of natural causes and anthropogenic activities is observed in the Venice coastland and the Po delta plain [3–6], while in the United States in correspondence of the coastal areas of Southern Louisiana and Mississippi [7–11]. Other examples are observed in the Yellow River and Yangtze China Deltas [12,13], along the Taiwan coastline [14–16], in the Indonesian Semerang city [17], in Thessaloniki coastal plain and municipality region (Northern Greece) [18,19].

Over the last decades, the land subsidence monitoring has been significantly improved thanks to the earth observation techniques based on Interferometric Synthetic Aperture Radar (InSAR). Furthermore, the recent advances in radar satellite capabilities and techniques based on the interferometric analysis of large datasets [20–29] have allowed even better spatial and temporal resolutions.
We applied multitemporal differential SAR Interferometry technique and specifically the Small Baselines Subset (SBAS) approach [20] to investigate the ground deformations of the Sibari Plain (SP), located in the northeastern sector of the Calabria Region (Southern Italy) covering an area of ~500 km². This Plain involves an important economic weight due to the agricultural production and a cultural appeal related to the presence of the ancient Sybaris, a powerful Greek colony in Magna Grecia founded in 720 BC. The area is also susceptible to flooding risk, reduced by means of a drainage channels network.

The used SAR datasets cover the temporal interval spanning from May 2003 and September 2013 for both ascending and descending orbits acquired by Envisat and COSMO-SkyMed satellites (for details see Table 1). Envisat is a European Space Agency (ESA) satellite operating in the C-Band (5.6 cm of wavelength) launched in the 2002 and operating up to 2012, with a revisiting period of 35 days, covering an area of about 100 × 100 km. The COSMO-SkyMed mission consists of a constellation from the Agenzia Spaziale Italiana (ASI) in the X-Band frequency (3.2 cm of wavelength). The first satellite of the constellation has been launched in June 2007 and the constellation (four satellites) has been fully operative by the end of 2010. The used acquisitions (Himage mode) are characterized by a swath of about 40 × 40 km² with a revisiting time period of 16 days.

Moreover, the availability of ascending and descending datasets allowed us to discriminate the vertical and east-west displacement components.

Finally, the retrieved InSAR results were validated by means of GPS measurements then analyzed and interpreted considering the available geological and hydrogeological information as well as new data collected during our field surveys.

2. Geological Setting

The SP is located along the boundary between Calabrian Arc and Southern Apennines (Figure 1), with the Pollino massif to the north and Sila massif to the south. The Calabrian Arc represents a fault-bounded continental fragment within the western Mediterranean orogeny. The tectonic evolution is related to the subduction and rollback of Ionian oceanic lithosphere and the slow convergence between the Eurasian and African-Adriatic continental plates [30–32]. The Southern Apennines are a NW-SE oriented segment of the Apennines thrust belt. It is characterized by a duplex structure, which consists of two thrusts belts overlapping the Apulian platform [33–36].

In the SP, the chain is prevalently composed by Apenninic terranes overlapped by thin Calabrian Arc units [37]. The first ones, outcropping along the northwestern side of the SP, are made by Mesozoic-Tertiary carbonatic succession and “flyschoid” deposits overlapped by Plio-Pleistocene siliciclastic succession [38–40]. Instead, the Calabrian Arc terranes, cropping out along the southern margin, consist of igneous and metamorphic rocks [41] (Figure 1).

In the central sector of the SP, the basement is overlapped by a thick Miocene-Holocene succession; the mio-pliocenic deposits represent the infilling of the Sibari-Corigliano Basin, which evolution is strictly connected with the Sangineto Line activity [42,43], and they are deformed by fault activity
with formation of morphological/structural highs and lows [44]. In [45], the authors suggest the existence of WNW-ESE trending shallow-crustal folds, developed within a recent and still active transpressional field.

Figure 1. Simplified lithological map of the study area with projection of the faults of ITaly HAzard from CApable faults database [46] and of the Sybaris fault zone (f.z.) [47].

The boundaries between the SP and Pollino and Sila massifs are marked by dislocations produced along plio-holocenic high angle faults. The latter is represented in the NW sector by the NW-SE normal fault of the Sangineto Line [48,49], while in correspondence of the Sila Massif it consists of the WNW-ESE Rossano-Corigliano fault (Figure 1). The recent activity of these faults is still debated [49–53].

The existence of NE-SW trending normal fault in the area of the Crati Delta is suggested by [48]; recent coseismic evidences related to a NE-SW structural lineament, maybe active during the period between the II and the VII-IX century A.D., are observed in Parco del Cavallo and Casa Bianca archaeological sites [47].

Five to eleven different order of marine terraces are recognized along the outer limit of the plain and uplift rates greater than 3.5 mm/yr are estimated [50,54,55].

Holocene deposits are the result of the post Last Glacial Maximum (LGM) transgression and the following (started about 6 ka B.P.) normal regression related to the Crati Delta progradation [56,57]. The Holocene evolution occurred in a tectonic-controlled setting, which drives the thickness and the facies association lateral variability of the deposits [57]. The post LGM succession consists of marine clayey-silty deposits, overlapping a late Pleistocene coarse-grained unit, passing upward to costal, deltaic and continental ones [57–59].
In the Northern sector, the main geomorphological elements are the alluvial fans of Raganello River, Satanasso Fiumara and Saraceno Fiumara.

**Historical Subsidence**

Subsidence is a well-known phenomenon in the Sybaris archaeological area (Figure 2), testified by the presence of three overlapping ancient towns, the Greek Sybaris (720-510 BC), the Hellenistic Thurii (444-203 BC) and the Roman Copiae (193 BC), presently buried between 7 and 3.5 m [60].

In [61], the author calculates subsidence rates ≥0.57 mm/yr (2330–860 ka B.P.) and ≥4.31 mm/yr (860 ka B.P.—present) for Casa Bianca, ≥0.35 mm/yr (2660–1600 ka B.P.) and ≥2.05 mm/yr (1600 ka B.P.—present) for Parco del Cavallo, ≥0.4 mm/yr for Stombi.

The subsidence in the archaeological area starts in the late Pleistocene up to the Holocene and is due by a combination of causes as neotectonic activity, glacio-eustatic variations and sediments compaction [62–64]. In the archaeological area the subsidence variability is correlated to the lateral variation of facies [65] and the most of geotechnical subsidence can be ascribed to a very compressible clay layer, laterally discontinuous, between 35 and 40 m of depth [62,65]. Since the 1950s, ~20 cm of subsidence have been recorded and ascribed to groundwater exploitation and related primary consolidation [65,66]. The subsidence in the Casa Bianca and Parco del Cavallo sites is confirmed by [67], which instead calculates a mean uplift rate of ~0.5 mm/yr in the last 11.2 ka for the Stombi site and hypothesizes that subsidence started 4ka B.P. and is ascribed to the deposition of fine compressible sediments, so excluding the tectonic contribution. The author of [68] highlights a temporal variability for the mean subsidence rates showing a peak of ~5–6 mm/yr in the Early Holocene whilst stability or small uplift in historical age (~3–1.3 ka B.P.) for the Casa Bianca and Parco del Cavallo sites, no subsidence in the Stombi site and uplift (~1.5 mm/yr) 6 km toward SE from archaeological area; this differential subsidence is mainly controlled by local tectonic structures.

![Figure 2. The archaeological sites of the ancient Sybaris (see inset in Figure 1 for its location).](image)

### 3. InSAR Adopted Technique

The classical differential SAR Interferometry (InSAR) is a technique that allows us to estimate the ground surface movements occurred between two different passes of the satellite over the same area using Synthetic Aperture Radar (SAR) data. The phase difference at each pixel is calculated after proper image co-registration resulting in a new image, called interferogram, an interference pattern
made up of fringes [69]. Each fringe represents a ground movement along the sensor line of sight (LOS) equal to $\lambda/2$ (where $\lambda$ is the adopted radar wavelength).

There are many physical phenomena contributing to phase measurements: phase variations within a pixel, the contribution of orbital variations, topography, atmosphere, and displacement. These contributions are estimated and then removed so to obtain only the effective ground displacement.

Our goal is to investigate the temporal evolution of the detected deformations. To this aim several different approaches have been proposed in the last decade to generate time series of ground displacement, capable of measuring the ground velocities rates with accuracies of ~1 mm/yr [70]. In this work, we applied the Small Baseline Subset InSAR technique (SBAS; [20]). The used SAR images were coupled based on constraints of small temporal and spatial baselines. The SBAS algorithm combines these acquisitions, also included in different interferometric subsets, using a minimum norm criteria combination of the velocity deformation, based on the Singular Value Decomposition (SVD) method. During the SBAS processing the estimation and removal of temporally decorrelated atmospheric artefacts are performed using double-pass filtering in the temporal and spatial domains, as explained in [20]. Moreover, we used the SRTM DEM (90-m resolution) for the subtraction of topography phase (http://www2.jpl.nasa.gov/srtm). Once the displacement time series is retrieved, the mean ground velocity in the time period covered by the data was calculated.

In detail, we applied the SBAS technique to four image datasets, two acquired from the European Space Agency (ESA) Envisat sensor and two from the Italian Space Agency (ASI) COSMO-SkyMed satellites. The first group contains 32 Envisat images acquired on the descending orbit (track 222, frame 2803), in the period 2004–2010 imposing 700 days as maximum temporal baseline and 350 meters for the maximum perpendicular one and considering 100 pairs. Instead, the ascending dataset (track 86, frame 789) is formed by 38 ASAR images acquired in the period 2003–2010 setting 900 days for the maximum temporal baseline and 350 meters for the maximum spatial one resulting in 135 interferograms. The second group (COSMO-SkyMed data) was composed of 28 images from the ascending orbit considering 180 days for the maximum temporal baseline and 700 meters as maximum spatial baseline resulting in 137 interferograms, spanning the interval 2011–2013 and 12 images from the descending orbit, acquired in the interval 2011–2012 setting 100 days for the temporal and 1000 meters for the spatial maximum baselines respectively forming 47 interferograms. We used the Sarscape software (Sarmap, CH) applying a looking factors equal to 20 and 4 for the azimuth and range direction, respectively for the ascending and descending Envisat processing obtaining a ground resolution of 80 meters and 11, 12 for the COSMO-SkyMed case with a final ground posting of 25 meters; later downgraded to 100 meters for the post-processing analysis. The initial performed multi-looking operation reduced the image radar noise (speckle) and increased the signal to noise ratio.

During the processing the precise orbital files (Envisat case) were used to estimate and remove orbital errors and Ground Control Points (GCPs) were selected, especially along the frame borders (in slant range geometry) and possibly over stable areas, to remove residual ramps. A stable point (under the geological point of view) was chosen as reference point (red star) for the retrieved mean ground velocity maps (Figure 3).

Initially, we validated our results comparing the ascending and descending velocity map after shifting the former with respect to the reference point of the latter preventing the subsidence areas. We also validated the Envisat results comparing with the available GPS in the whole SAR frames (Figure 3). Firstly, we projected the GPS velocities onto the ascending and descending SAR LOS respectively then we averaged the InSAR velocities in correspondence of the GPS benchmarks considering a buffer area of 200 meters. The comparison between GPS and InSAR velocities is shown in Figure 3a,b for the ascending and descending tracks. Except for very few GPS benchmarks, the velocities differences are within 1.5 mm/yr, which is the associated error to the multi-temporal SAR outputs. Unfortunately, it has not been possible to perform the same validation method for the CSK results because not enough CGPS are available. However, considering the common similar patterns and trends for both Envisat
and CSK ascending and descending retrieved maps, we convinced the good quality of the obtained CSK results. Note that for the CSK processing we focused our study on the Sibari Plain.

To this aim, we calculated the East ($E$) and Vertical ($V$) component solving the following system of equations:

$$E = ((u_2d/det) \times (Asc - (n_2a \times N))) - ((u_2a/det) \times (Dsc - (n_2d \times N)))$$

$$V = ((-e_2a/det) \times (Asc - (n_2a \times N))) + ((e_2a/det) \times (Dsc - (n_2d \times N)))$$

where $det = ((e_2d \times u_2d) - (u_2a \times e_2d)$

where $u_2d, e_2d, n_2d, u_2a, e_2a$ and $n_2a$ are the descending and ascending cosine directors respectively, and Asc and Dsc are the ascending and descending mean velocity maps.

**Figure 3.** (a) Ascending and (b) descending line of sight (LOS) ground velocity maps. Red star represents the stable point (stable means that the mean velocity is 0 mm/yr) used as reference point. In the maps are also plotted the GPS benchmarks used during the validation of the Envisat results. The diagrams show the comparison between the LOS Synthetic Aperture Radar (SAR) velocities and GPS. The correlation coefficient is respectively 0.51 and 0.65 for the ascending and descending cases, while the mean squared error is 1.1 mm and 1.2 mm.

Concerning the third condition, we imposed the North equal to the interpolated CGPS one. The final Vertical and East component (Figure 4) result quite sparse, this mainly depends on the difficulty to obtain a large coverage for the ascending and descending velocity maps due to the land cover
and agricultural use of the investigated area leading to a fast coherence loss. Moreover, about the East component some overestimated values are present in the North Eastern part probably due to unwrapping errors due to the presence of the mountains.

We performed the previous post-processing steps for the full SAR frames and only at a later stage; we focused our analysis and interpretations referring to the East and Vertical component about the Sibari Plain area.

Figure 4. Mean deformation velocity components computed from (a,b) Envisat and (b,c) COSMO-SkyMed dataset. (a,c) Vertical velocity component; positive values indicate uplift and negative values subsidence. (b,d) East component; positive values indicate eastward displacement and negative values indicate westward displacement. The area covered by COSMO-SkyMed data is represented by inset (1) in the panel a.
InSAR Results

In the following, we report the deformation pattern of SP at a large scale retrieved by InSAR processing and discuss the results in detail using available tectonic, stratigraphic, geomorphological and hydrogeological information. We investigate natural and/or human causes of the vertical and horizontal surface displacements.

The distribution of Vertical component shows a coastal area (from Villapiana Lido to Marina di Schiavonea) characterized by subsidence with velocity up to ~20 mm/yr (Envisat and COSMO-SkyMed sensors); on the contrary, boundary areas of SP show ~0 mm/yr on the vertical velocity component (Figure 4).

The horizontal East-West component shows low displacement values that are inside the errors range and are not significant.

In the archaeological area, the subsidence rate is ~2 mm/yr (Envisat) and ~3 mm/yr (COSMO-SkyMed) for the Parco del Cavallo site and ~2 mm/yr (Envisat) for the Casa Bianca site. These values are comparable to the velocities calculated by previous authors [61,67].

4. Discussion

We analyzed different causes contributing to subsidence phenomena. To this aim, we compared the InSAR results in time and space with the geodynamic and structural settings, the seismicity distribution, the spatial variation of the Plio-Quaternary and Holocene deposits successions thickness and compaction, depth variation of the water table and urban sprawl.

4.1. Geodynamic Setting

The SP is located at the junction of the Calabrian Arc and the Southern Apennine. Moving few kilometers towards SE from the SP, the geodynamic setting is complicated by the presence of the junction (Apulia Escarpment) (Figure 1) between the thin Ionian basin crust (southward) and the thick crust of the Apulian platform (northward). In the Ionian Sea, in front of the SP an active oblique-contractional belt (the Amendolara Ridge) is recognized. The latter is due to the combined effects of subduction retreat of the Ionian slab underneath Calabria and stalling of Adriatic slab retreat underneath the Apennines [71]. The belt consists of the alignment of three anticlines (Amendolara, Rossano and Cariati) bounded by a main SW-verging back-thrust (Figure 5a).

A flexural subsidence, probably triggered by a back-thrusting [72], is considered one of the mechanisms to explain the ground subsidence detected by InSAR results along coastal sector of the SP (Figure 5b).

4.2. Structural Setting and Earthquakes

Possible relationship between SP subsidence and recent tectonic activity relies on the analysis of structural lineaments reported into the Italy Hazard from Capable faults database [46]. According to the main faults bounding the SP, we subdivided it in three sectors (Figure 5a) and with foreseen different deformation patterns for the three sectors from InSAR ascending and descending time series (Figure 6). The analysis of the averaged Envisat displacement time series show very similar trends for the three investigated areas, although with different absolute values. On the other hand, COSMO-SkyMed time series show a different deformation pattern for each sector.
Figure 5. (a) The Amendolara Ridge located in the Ionian Sea in front of the Sibari Plain. The bathymetric map (modified from [73]) shows the presence of three morphological highs (Amendolara, Rossano and Cariati highs). The main structural elements (thrusts and anticlines) are extracted from [71]. In the map is showed the hypocenters location of the earthquakes occurred between 2003 and 2010 (ISIDe [74]) with the relative magnitude and hypocentral depth. Onshore capable faults from [46] and focal mechanisms from [45]. Yellow dashed lines are the boundaries of three selected sectors (1. Corigliano; 2. Crati Delta; 3. Sibari-Villapiana) for Envisat and COSMO-SkyMed time series analysis (Figure 6). (b) Sketch (based on the interpretation of the seismic profile F75-89 of [45]) representing the trigger of the flexural subsidence due to the overlap of the Amendolara Ridge back thrusts on Calabrid and Southern Apennines units.

In detail, the ascending time series reveal a general trend characterized by an increase of the negative displacements but with greater absolute values in the southern sector. The descending time series highlight an increase of negative displacements in the Corigliano sector with values between 10 mm and −5 mm in the Villapiana-Sibari sector, and an expected increasing positive trend in the Crati Delta sector. This trend can be related to the presence of a horst bounded by Crati and Timparelle faults [48].

A comparison between ground deformation pattern and seismicity was performed taking into account the earthquakes occurrences and magnitude [74] during the time covered by SAR data (Figure 5a). No significant earthquakes struck the study area excepting for one shallow offshore M 4.5 event on 27 June 2006 [74] with none ground deformation effects (Figure 6).

Therefore, we can exclude the presence of seismically-induced displacement and we hypothesize that the structural settings could indirectly influence the ground subsidence phenomena.
4.3. Role of the Plio-Quaternary Succession Load

Possible relationships between subsidence and spatial variation of the Plio-Quaternary succession thickness in the Crati Delta area (Figure 7a) as reconstructed from multichannel seismic profiles and well data [75] was investigated. The thickness observed in the seismic profile has been converted from Two Way Travel time to meter using an average velocity of 2000 m/s for the Pliocene and Pleistocene deposits based on the sonic logs, available for Thurio and Ogliastrello wells, analyzed by [44].

The Envisat Vertical component superimposed to the Plio-Quaternary succession isopach map (Figure 7b) shows a prevalent distribution of vertical displacement rates toward the coastline, where the deposits are thicker.

We compared ground velocity profiles, for the Envisat Vertical component, with the corresponding profiles of the Plio-Quaternary deposits thickness. The comparison (Figure 7c) shows a good correlation between the general subsidence trend and the Plio-Quaternary succession thickness. We observed higher subsidence values in the E and NE sectors of Crati Delta, where the Plio-Quaternary succession is thicker. In the SW sector, the reduced thickness corresponds with lower subsidence values. Lower values are also detected in the S-SE sector (Corigliano harbor area), where the presence of shallow igneous-metamorphic bedrock drives the Plio-Quaternary succession thickness.

Furthermore, we considered the ascending and descending ground velocity values in correspondence of the seismic profiles and wells locations, correlating the ground velocity with the thickness of Plio-Quaternary deposits (Figure 8). Both ascending and descending velocities increase with the growth of the succession thickness.

Based on the above analysis results, we found a direct correlation between subsidence spatial trends and distribution of the thickness of Plio-Quaternary deposits, which amplify the subsidence phenomenon as well documented during the evolution of the sedimentary basins [76–81].

4.4. Holocene Deposits

A possible correlation between subsidence and Holocene sediments compaction was investigated focusing on the lithological features, thickness and stratigraphic architecture of the Holocene deposits.
We analyzed 200 boreholes (Demanio Idrico Prov. Cosenza, [82,83]) and literature data [57,61,65] to define lithology, vertical-lateral variations and thickness of the post-LGM deposits. An isopach map of the Holocene succession superimposed with the Envisat Vertical component was generated (Figure 9a). The greater vertical displacement values are located in the area of the Holocene deposits accumulation but the larger subsidence rates do not correspond to the major depozones.

To well investigate the second topic, we compared each Vertical displacement value of with the corresponding thickness of the Holocene deposits (Figure 9b). We also performed a comparison between thickness and vertical component in correspondence of each borehole, if SAR data are available (Figure 9c).

The analysis reveals that the Holocene sediments compaction does not play a dominant role in the subsidence of the SP area, on the contrary as described for other delta plains [5,7].

Subsidence rates of 3–5 mm/yr induced by to the Holocene deposits consolidation due to phreatic and confined water tables drop are calculated by [66] and are not sufficient to justify the vertical displacements inferred by SAR data.

Although the minor role of Holocene deposits compaction, the boreholes analysis highlights a clear influence of the lateral variations of the sediments lithology for the subsidence phenomena. In detail, we took into account the ground deformation patterns from ascending and descending Envisat data, and the boreholes into the Saraceno and Satanasso alluvial fans areas (Figure 10a), considering the thickness of fine-grained (silt and clay) deposits (Figure 10b,c). Greater ground velocities are usually retrieved close to the boreholes recording a major thickness of fine-grained sediments while the lower values are in correspondence of the coarse-grained (gravel and sand) deposits which made up the alluvial fan bodies. Thus, the lateral variation of the fine-grained materials thickness creates differential land subsidence.

Figure 7. (a) Location of the used wells and seismic profiles. (b) Isopach map of the Plio-Quaternary succession with the distribution of the Envisat Vertical component. (c) Comparison between the thickness of the Plio-Quaternary succession and ground velocity, based on the Envisat Vertical component, along one of the analyzed profiles (its trace in Figure 7a). The discontinuous lines for the InSAR profile are due to missing data in the ascending and/or descending original velocity maps due to coherence lack.
Figure 8. Correlation between (a) ascending and (b) descending velocities with the thickness of the Plio-Quaternary succession.

Figure 9. (a) Envisat Vertical component overlapped to the Holocene deposits isopach map. (b) Scatter plot Envisat Vertical component vs. Holocene deposits thickness inferred from isopach map. (c) Comparison between Holocene sediments thickness and Envisat Vertical component in correspondence of each borehole.
Figure 10. (a) Analyzed boreholes between Saraceno and Satanasso alluvial fans (see Figure 1 for their location). (b) Correlation among Envisat ascending and (c) descending velocity and the thickness of fine-grained deposits (silt and clay).

4.5. Land Use and Historical Evolution

SP represents a geomorphological system with a rapid evolution in historical and recent times controlled by geological processes and anthropogenic activities. The present morphological setting results from the works of a 1960s–1990s land reclamation project aimed to convert a marshy area to an agricultural zone. Several widespread agricultural areas (e.g., orchards, arable soils) with localized urban settlements [84] are present in SP.

In order to define land use variations, we analyzed the 1954 and 1998 aerial photos; for a detailed analysis of urban area growth effects. We considered three areas (Marina di Sibari; Laghi di Sibari and C.da Ricota; see Figure 4), characterized by the same subsoil stratigraphy, and compared the Envisat ascending and descending displacement time series concerning urban and agricultural zones respectively (Figure 11).

In Laghi di Sibari and C.da Ricota sites, we observe a progressive temporal differentiation in the subsidence values that increase faster in urban areas reaching a maximum difference compared with agricultural ones of ~35 mm in August 2010 (for the descending track) and of ~25 mm in September 2010 (for the ascending one). The subsidence values are very similar for the two considered sectors in the Marina di Sibari zone, where the descending time series show a little increase (~5 mm) of ground deformation values into the urban area.

We observed a diffuse subsidence irrespective of the land use, with some increments of deformation velocity in correspondence of Laghi di Sibari and C.da Ricota settlements.

Therefore, we suggest that the urban-induced loading represents an incremental factor of subsidence.
4.6. Water Table Variations

SP is characterized by two well-defined aquifers: the shallower one (from soil surface to \(-20/\sim 30\) m above sea level) separated by clayey and silty-clayey layer and the deeper one (from \(-50/\sim 60\) m a.s.l.) \[85\]. Both the aquifers are characterized by intense water exploitation. The piezometric level variations of the shallow one shows 5 meters of drop between 1930s and 2002 \[86,87\].

We investigated to find possible correlation between groundwater exploitation and ground deformation comparing the spatial distribution of phreatic water table variations in the time interval 1930s–2002 \[86\] and the Vertical velocity distribution from Envisat and COSMO-SkyMed data in the following 2003–2013 time interval (Figure 12). In the most recent period, we have observed that subsidence is present in areas showing rise, drop and stability of the water table.

Considering the effect of water table variation between 2002 (data from \[86\]) and 2013 (this work), we analyzed the distribution of the Vertical velocity (from Envisat and COSMO-SkyMed data, Figure 12). The subsidence location does not show a strict correlation with the water table drop.

We selected 6 wells (Figure 1), with piezometric records in 2002 and 2013, and we analyzed the Envisat time series for both ascending and descending orbit (Figure 13) considering a circular buffer of 500 m radius. For the two wells No. 2 and 3, showing a groundwater drawdown, we observe a displacement trend similar to the No. 1, 4, 5 and 6 ones characterized by water table rise. Moreover, the time series show smaller displacements rates in correspondence of the well No. 2. There is no
evidence of correlation between water table variation and displacement time series, probably because the progressive groundwater level stabilization from 2002 to 2013. In contrast, the groundwater level drop between 1930s and 2002 \cite{86,87} could have been caused by a more recent consolidation process of soft materials as observed by \cite{88} in the Alto Guadalentin Basin.

Figure 13. Envisat (both ascending and descending orbit) time series of the wells No. 2 and No. 4 (see Figure 1 for their location) recording respectively a water table drop and rise. In table are reported the water table variation between 2002 and 2013 for each well.

For the confined aquifer, the lack of historical data did not allow us to analyze the effects of its variations on subsidence.

Therefore, we conclude that water table variations are only an incremental factor of subsidence in some areas as the Sybaris archaeological area (see Figure 2 to localize Sybaris site), where a well-points system (80 l/s) works every time to maintain the excavations above the water table. Moreover, in these areas, a water depletion in compressible soils could trigger an amplified and long-term subsidence as shown in other Quaternary basins \cite{88,89}.

5. Conclusions

InSAR results from data acquired by Envisat and COSMO-SkyMed satellites, and spanning from 2003 to 2013, are able to detect ground displacements in the SP area (Calabria, Southern Italy). The ground deformation field is dominated by a widespread-subsidence with rate up to ~20 mm/yr.

We can summarize that the location of the SP in an area marked by geodynamic and geological complexity precluding the attribution of the ground deformation triggering factor to a unique subsidence mechanism.

We suppose that the active oblique-contractional belt in the Ionian Sea in front of SP \cite{45} can trigger a flexural subsidence mechanism due to the back thrusting load above the underlying Calabrid and Southern Apennines Units. Furthermore, the observed correlation between Plio-Quaternary sediments thickness and ground deformation suggests that the isostatic compensation plays a major role in the SP present subsidence. We highlight the absence of correlation between the occurring earthquakes close to the main capable faults, so excluding the contribution of faulting activity and earthquakes for the present SP subsidence phenomena, according to \cite{67} (based on geomorphological data); we hypothesize that the structural settings could indirectly influence the ground subsidence phenomena.

Another minor triggering mechanism of subsidence is represented by the compaction of the compressible fine-grained sediments related to the Holocene post-LGM transgression and following Crati Delta progradation. We show the effect of the lateral transition between fine- and coarse-grained Holocene sediments on the ground deformation spatial pattern.

Anthropogenic activities can be considered as an incremental factor of the SP subsidence. In fact, we observe higher subsidence velocities in the more anthropized areas, where soil compaction is amplified by urban-induced loading. In some area showing excessive water withdrawals (e.g., Sybaris archaeological site), water table variations may increase the deformation velocity because can be the origin of delayed consolidation process in soft materials.
The presence of subsidence phenomenon in the whole SP and the complicated identification of a main triggering mechanism, due to the peculiar geological setting, is also in agreement with [90].

The results of this study show as the deformation field of complex basins may be due to coexisting regional and local natural processes and anthropic activities. To better assess and monitor the hazard(s) related to the detected ground deformation phenomena, a multiparametric monitoring of SP area would be recommended.

Acknowledgments: This paper is sponsored by PON (Operational National Plan) 2007–2013 from MIUR (Italian Research Ministry of Research) Project AMICUS (Study for the environmental protection and the mitigation of Anthropogenic Pollution In the Coastal Environment of selected areas of Calabria; ID: PON01_02818) and resulted from the Ph.D. work of Giuseppe Cianflone. Results carried out using COSMO-SkyMed® PRODUCTS, © ASI (Italian Space Agency)—provided under license of ASI in the framework of the S3 Project “Short term earthquake prediction and preparation” (DPC-INGV, 2013). The Envisat images are provided by ESA (European Space Agency) under the CAT.1P 5605. The InSAR processing are carried out using the Sarscape© software (Sarmap, CH). We thank Paolo Riccardi and Alessio Cantone (Sarmap, CH) for the fruitful collaboration and useful suggestions during the SAR data processing. We also thank Enrico Serpelloni (INGV-Bo) for providing the GPS solutions. We are grateful to G. Ventura for his careful and critical reading of the manuscript. We thank editors and four anonymous reviewers, which improved the paper quality with their suggestions.

Author Contributions: All authors participated in the writing of the manuscript, and have read and approved the final manuscript. Giuseppe Cianflone developed the project and conducted the geological survey. All authors analyzed and interpreted the data. Cristiano Tolomei processed the SAR data.

Conflicts of Interest: The authors declare no conflict of interest.

References

16. Tung, H.; Hu, J.C. Assessments of serious anthropogenic land subsidence in Yunlin County of central Taiwan from 1996 to 1999 by Persistent Scatterers InSAR. Tectonophysics 2012, 578, 126–135. [CrossRef]
44. Spina, V.; Tondi, E.; Mazzoli, S. Complex basin development in a wrench-dominated back-arc area: Tectonic evolution of the Crati Basin, Calabria, Italy. *J. Geodyn.* **2011**, *51*, 90–109. [CrossRef]


72. Carapa, B.; Garcia-Castellanos, D. Western Alpine back-thrusting as subsidence mechanism in the Tertiary Piedmont Basin (Western Po Plain, NW Italy). Tectonophysics 2005, 406, 197–212. [CrossRef]


76. Watts, A.B.; Ryan, W.B.F. Flexure of the lithosphere and continental margin basins. Tectonophysics 1976, 36, 25–44. [CrossRef]


© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
DInSAR-Based Detection of Land Subsidence and Correlation with Groundwater Depletion in Konya Plain, Turkey

Fabiana Caló 1,*, Davide Notti 2, Jorge Pedro Galve 2, Saygin Abdikan 3, Tolga Görüm 4, Antonio Pepe 1 and Füsun Balık Şanlı 5

1 National Research Council (CNR) of Italy—Institute for the Electromagnetic Sensing of the Environment (IREA), via Diocleziano 328, 80124 Napoli, Italy; pepe.a@irea.cnr.it
2 Department of Geodynamics, University of Granada, 18071 Granada, Spain; davidenotti@gmail.com (D.N.); jpregalve@gmail.com (J.P.G.)
3 Department of Geomatics Engineering, Engineering Faculty, Bulent Ecevit University, 67100 Zonguldak, Turkey; sabdikan@beun.edu.tr
4 Geography Department, Istanbul University, Ordu Cad. No. 6, 34459 Laleli, Istanbul, Turkey; tgorum@gmail.com
5 Department of Geomatic Engineering, Civil Engineering Faculty, Yildiz Technical University, 34220 Esenler, Istanbul, Turkey; fbaliksanli@gmail.com

* Correspondence: calo.f@irea.cnr.it; Tel.: +39-81-7620-633

Academic Editors: Zhenhong Li, Roberto Tomas, Richard Gloaguen and Prasad S. Thenkabail
Received: 30 June 2016; Accepted: 10 January 2017; Published: 17 January 2017

Abstract: In areas where groundwater overexploitation occurs, land subsidence triggered by aquifer compaction is observed, resulting in high socio-economic impacts for the affected communities. In this paper, we focus on the Konya region, one of the leading economic centers in the agricultural and industrial sectors in Turkey. We present a multi-source data approach aimed at investigating the complex and fragile environment of this area which is heavily affected by groundwater drawdown and ground subsidence. In particular, in order to analyze the spatial and temporal pattern of the subsidence process we use the Small Baseline Subset DInSAR technique to process two datasets of ENVISAT SAR images spanning the 2002–2010 period. The produced ground deformation maps and associated time-series allow us to detect a wide land subsidence extending for about 1200 km² and measure vertical displacements reaching up to 10 cm in the observed time interval. DInSAR results, complemented with climatic, stratigraphic and piezometric data as well as with land-cover changes information, allow us to give more insights on the impact of climate changes and human activities on groundwater resources depletion and land subsidence.

Keywords: remote sensing; Differential SAR Interferometry (DInSAR); Small Baseline Subset (SBAS); GRACE; subsidence; sinkholes; groundwater; climate change; Turkey

1. Introduction

Protection of water resources and mitigation of the risk associated with their overexploitation represent challenging issues worldwide. In Turkey, one of the Mediterranean countries more severely affected by water scarcity [1], the Konya basin, in central Anatolia, is facing with serious water-related problems, due to climate changes [2] and lack of efficient water management policies especially in the agriculture sector [3].

Moreover, growing irrigation needs are often satisfied by uncontrolled practices of groundwater extraction: among 92,000 wells located in the Konya basin, it is estimated that 66,000 are illegal [4]. Such an aquifer overexploitation has important environmental implications for the region.
Groundwater overuse and depletion not only directly impact on the hydrological cycle [5,6], but have severe negative side-effects on biodiversity, ecosystem health, soil deterioration and drying up of lakes and wetlands [7,8]. Moreover, excessive groundwater withdrawal makes the region extremely vulnerable to natural and man-made hazards. As a result of piezometric level decrease, the Konya area is extensively affected by two different processes of land subsidence: (1) the progressive settlement induced by the compression of unconsolidated alluvial sediments; (2) sudden collapses of subterranean cavities formed by the combination of rock dissolution and suffosion processes. These phenomena cause damage to buildings and infrastructures, leading to high socio-economic costs for the concerned communities. To our knowledge, most of the studies conducted on the Konya area have focused on sinkholes investigations [9–12] while only a few studies on the analysis of the land subsidence caused by consolidation of alluvial fans [13,14].

In such a context, it is clear that obtaining information on the spatial and temporal evolution of land subsidence plays an important role on hazard assessment and may significantly contribute to risk prevention and mitigation. However, the continuous monitoring of surface deformations at regional/basin scale is a resource-intensive task to be accomplished by using traditional ground-based techniques. The use of satellite based information in Earth’s surface studies represents an important breakthrough, fulfilling the requirements of high spatial and temporal coverage. In particular, remote sensed data acquired by satellite Synthetic Aperture Radar (SAR) sensors and processed through Differential SAR Interferometry (DInSAR) techniques [15] greatly contribute to surface deformation analyses. DInSAR firstly relied on the use of a pair of pre- and post- event SAR acquisitions in order to retrieve ground displacements caused by single trigger events, i.e., earthquakes and volcanic unrest [16,17]. Since such early applications, relevant algorithmic advances have been achieved and several DInSAR approaches fully exploiting large SAR datasets to generate time-series of deformations have been developed. These techniques can be grouped in two main categories: the Persistent Scatterers (PS) [18,19] and the Small Baseline (SB) [20,21]. In particular, the PS techniques focus on point-like targets that are not significantly affected by decorrelation noise [22], and are suitable to monitor deformations affecting point-wise structures. Conversely, the SB methods rely on selecting SAR data pairs with short temporal (i.e., time interval between two acquisitions) and perpendicular (i.e., the distance between two satellite passes) baselines in order to reduce the noise, and allow the investigation of displacements over distributed scatterers (DS) on the ground.

Subsidence studies have greatly benefited from the application of both PS and SB DInSAR techniques which have been used for the analysis of the temporal evolution of surface deformations induced by natural and anthropogenic factors, from intense human activities in urban areas [23,24] to rock dissolution processes (i.e., karst subsidence) [25–27]. In particular, several works have focused on the application of such DInSAR techniques to the detection and monitoring of land subsidence induced by groundwater extraction in many regions of the world. Chaussard et al. [28] and Castellazzi et al. [29] used DInSAR deformation time-series of ALOS and RADARSAT-2 data, respectively, to study the land subsidence caused by massive groundwater over pumping in the entire central Mexico. Dehghani et al. [30] exploited both descending and ascending ENVISAT datasets to measure the magnitude of groundwater related subsidence in the Tehran basin, Iran. The effectiveness of DInSAR techniques in mapping and monitoring surface deformations has been proved in many applications from Europe to Asia [31–33] showing that ground subsidence triggered and/or accelerated by aquifer overexploitation is a common hazard impacting extensive areas worldwide.

In this paper, we apply the SB approach referred to as Small BAaseline Subset (SBAS) [20], which is capable of detecting and measuring displacements affecting extended rural areas, thus resulting particularly suitable for the regional scale analysis of the land subsidence affecting the Konya plain. Through the generation of displacement time-series and maps of deformations rate we analyze the temporal and spatial pattern of the land subsidence caused by groundwater resource overexploitation. Furthermore, to investigate the correlation between land subsidence and its conditioning/triggering factors, we adopt a multi-technique approach, based on the use of several data, from traditional
ground-based measurements (piezometric data) to remote-sensed data which recently opened new perspectives in water resources studies (GRACE data), from meteorological data to Landsat images and CORINE Land Cover maps used for extracting land cover information. Integration of such a variety of data, embedding local- and regional-scale information, allows us to get a holistic insight into the climatic, hydrogeological and human dynamics of the whole area.

2. Study Area

2.1. Geographic and Climatic Setting

The study area is located within Konya basin, central Anatolia, which is the biggest endorheic basin in Turkey extending over an area of about 54,000 km², with elevations ranging from 900 m to 3500 m above sea level (a.s.l.) (Figure 1). The basin lies on the northern side of the Central Taurus Mountains which have an important role on surface and sub-surface water supply. The Konya basin has a typical arid to semi-arid climate with a long-term average yearly precipitation of 380 mm, ranging from 250 mm in the plain areas to more than 1000 mm in the mountainous areas in the south [5], and an evapotranspiration of about 550–600 mm per year. Winters are generally cold and wet whereas summers are hot and dry. Despite having such a dry climate with low precipitation input, there are extensive and intensive agricultural activities in the region, whose main source of water is the groundwater which is extracted via more than 92,000 groundwater wells distributed around the basin [34]. In the last years the increase of irrigated land (about 3700 km²) related to some cultivations like sugar beet and corn caused an increase of groundwater consumption. The estimated groundwater consumption in 2010 was about 2.7 billion m³ [35]. Numerous smaller fresh/brackish wetlands have dried out in the last decades as consequence of groundwater depletion [5].

![Figure 1](image.png)

Figure 1. Study area: location and topography. Frames show the footprints of the ascending (red) and descending (blue) ENVISAT SAR data used in this study. Location of meteorological stations, groundwater and stratigraphic wells used in this study is also shown.
2.2. Geological and Geomorphological Setting

In the study area, three main geological sectors can be identified: the Konya-Karapinar plain, the volcanic field and the Obruk plateau (Figure 2). The Konya-Karapinar plain is characterized by the presence of Late Miocene to Quaternary formations, the latter mainly made up of lacustrine deposits. This formation, deriving from the fill of Pleistocene lakes, is mainly made of unconsolidated clay, silt and sandstone. In the plain it is also possible to find aeolian and alluvial fan deposits [9,36,37]. The Late Miocene to Pliocene stratigraphy starts with a basal conglomerate and continues upward with lacustrine limestone and weakly cemented marl and limestone alternation [12]. At S-E of Karapinar, a volcanic field with age variable from Miocene to Quaternary, whose activity is related to the extensional phase that originated the Konya basin, is present; cones and other volcanic structures are visible in the study area. The Obruk plateau is characterized by Mesozoic to Oligo-Miocene formations. The Oligo-Miocene formations are composed by continental clastics, tuff, limestone and evaporites at the bottom and gypsum-shale alternation at the top [12]. Mesozoic formations are sparsely distributed in the western part of the Obruk plateau and along the southern boundary of the Konya plain, and are made up of carbonates and metamorphic rocks [38].

From the geomorphological point of view, the landscape of the study area is conditioned by several active processes: volcanism, water and wind erosion, and rock dissolution (Figure 2). Among these, water erosion and dissolution-induced subsidence are the most widespread phenomena in the study area, while aeolian processes are effective in relatively local sectors of the Konya basin.

The dissolution-induced subsidence is represented by the karst depressions called ‘sinkholes’. Sinkholes are common geomorphic features in the study area (Figure 2). These landforms are locally called “Obruks”. This term is also accepted at international level for describing the large sub-circular closed depressions observed in the Central Anatolia [39]. There are two groups of obruks according to [12]: the “old” obruks, characterized by large dimensions (up to 640 m wide), for which the time of formation is unknown (Figure 2B); the “young” obruks that have been formed during the last decades and have diameters of less than 50 m (Figure 2A). The “old” obruks are concentrated in the Obruk Plateau, a horst slightly higher than the surrounding grabens formed by the Konya plain and the Tuz Lake basin. These “old” obruks can be classified as caprock and bedrock collapse sinkholes [40]. The analysis of geological, geophysical, hydrogeological data and the groundwater chemical and isotopic compositions indicates a hypogenic origin for these large obruks [12]. There are not evidences of subsidence activity in these sinkholes so far and they may be defined as relict karst landforms. On the other hand, most of the “young” obruks can be classified as cover collapse sinkholes [40] and their formation is associated to the human-induced groundwater withdrawal [9]. In the last decade, these “young” obruks have caused social alarm in the Karapinar plain. Indeed, 13 collapse sinkholes up to 25 m in diameter occurred from 2006 to 2009. They are especially concentrated in and around the Seyithaci settlement (Figure 2A). According to the occurrence of these sinkholes, we can expect future sudden collapses if the water lowering trend will continue. Future sinkholes could damage buildings and infrastructures in the study area and represent a hazard for the local population.

Another dynamic process in the region is the sand dunes (Figure 2). The long term displacements of the sand dunes have caused the formation of sand hills and dunes [37,41]. The movement direction of the very active sand dunes, that are located at the south of Karapinar settlement (Figure 2C), changes depending on the dominant direction of the seasonal wind. However, it is observed to be mainly from south to north [37].

In the study area it is also possible to observe some active faults (Figure 2). In particular, the Seyithaci fault system is a dynamic process delimiting the Obruck plateau from the plain at north of Karapinar. The fault has a vertical component character and follows a highly linear form between the floors of plateau and plain [42]. The fault system is defined as active and is known to establish a surface rupture during Holocene [43], and has an important effect in the evolvement of Karapinar plain and in the formation of sinkholes because the karstification of the bedrock may be developed through fault structures.
Remote Sens. 2017, 9, 83

Figure 2. Geological and geomorphological map of the study area (adapted from [37]). Dynamic geomorphological processes in the study area: (A) Active and (B) relict sinkholes near the Seyithaci and Kizoren settlements, respectively; (C) Karapinar sand dune field (photo courtesy of Atlas Dergisi).

3. Data and Methods

Different data and information have been integrated to carry out a comprehensive analysis of the phenomena under investigation, from satellite images (Table 1) to data acquired through in situ monitoring techniques (Table 2).

Table 1. Satellite-based data.

<table>
<thead>
<tr>
<th>Satellite</th>
<th>Band/Parameters</th>
<th>Final Product</th>
<th>Time Interval</th>
<th>Revisit Time</th>
<th>Spatial Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENVISAT—ASAR</td>
<td>C-band</td>
<td>Ground displacements</td>
<td>2002–2010</td>
<td>35 days</td>
<td>20 × 4 m²</td>
</tr>
<tr>
<td>Landsat-5/-8</td>
<td>Multispectral</td>
<td>Qualitative land cover changes/mNDVI</td>
<td>May 2000; May 2015</td>
<td>16 days</td>
<td>30 m</td>
</tr>
<tr>
<td>GRACE</td>
<td>Gravity field</td>
<td>Groundwater storage</td>
<td>2002–2010</td>
<td>30 days</td>
<td>350 km</td>
</tr>
<tr>
<td>Shuttle Radar Topography Mission (SRTM) [44]</td>
<td>C-band</td>
<td>DEM 3-arc sec</td>
<td>2000</td>
<td>Single mission</td>
<td>90 m</td>
</tr>
</tbody>
</table>
3.1. SAR Data and DInSAR Processing

The study area has been investigated by exploiting the European Space Agency (ESA) archives of raw data collected by the ENVISAT-ASAR satellite (Table 1). These data were acquired with a 35-day revisit period at a look angle of 23° and with a C-band frequency (5.65 cm wavelength). In particular, we processed two datasets spanning overall nearly eight years: 38 SAR images taken from 5 December 2002 to 9 September 2010 on descending orbits (Track 436, Frame 2838), and 27 SAR images acquired from 14 December 2002 to 5 June 2010 along ascending orbits (Track 71, Frame 746). We applied the SBAS—DInSAR technique [20] by selecting SAR data pairs characterized by a perpendicular baseline <400 m and a temporal baseline <1500 days, and generated 111 and 73 interferograms from the ENVISAT descending (Figure 3A) and ascending (Figure 3B) datasets, respectively. As a next processing step, we computed differential interferograms (embedding only the information relevant to the deformation phase component). The phase component related to the topography of the area was estimated and removed using a Digital Elevation Model (DEM) produced from the Shuttle Radar Topography Mission (SRTM) with about 90 × 90 m² spacing (Table 1) [44,45].

Then, the differential interferograms were unwrapped by making use of the space-time Extended Minimum Cost Flow (EMCF) technique proposed in [46] which takes into account both the spatial and temporal relationships among the sequence of interferograms. After the phase unwrapping operation, differential interferograms were inverted by solving a Least Squares (LS) minimization problem [20]. SBAS allows retrieving the surface deformation evolution and is also capable to mitigate possible topographic artifacts and atmospheric phase disturbances [20]. In particular, the filtering operation for the atmospheric phase component is based on the observation that the atmospheric signal is highly correlated in space and poorly in time [18]. Moreover, the overall analysis was carried out on differential interferograms subject to complex multi-look operation with 4 looks in the range direction and 20 in the azimuth direction in order to significantly reduce the phase noise [45]. Accordingly, in the ENVISAT case the ground resolution of the DInSAR deformation rate map turns to be about 80 × 80 m² which is particularly suitable to investigate land subsidence occurring at regional level as within the Konya plain. According to the experience gathered in the Ebro Valley evaporite karst [26], this resolution could also be enough to identify progressive subsidence due to karst processes if the area affected by settlements is over ~50,000 m². Furthermore, the availability of both ascending and descending geometries allowed us to investigate the direction of the ground movement and compute the E-W and vertical components of displacements as follows [47]:

\[
\begin{align*}
d_{\text{vert}} &= \frac{d_{\text{desc}} + d_{\text{asc}}}{2\cos \theta} \\
d_{E-W} &= \frac{d_{\text{desc}} - d_{\text{asc}}}{2\sin \theta}
\end{align*}
\]  

(1)

where \(d_{\text{desc}}\) and \(d_{\text{asc}}\) represent the Line of Sight (LOS)—projected deformation computed from descending and ascending datasets, respectively, and \(\theta\) the look-angle of the satellite. We remark that the retrieval of the N-S displacement component is very critical because of the near-polar orbit of the satellite (i.e., the flying direction of the satellite is nearly parallel to N-S) [48].

The applied decomposition method is performed on the LOS time-series [49], and is based on the consideration that over our study area the SAR data were acquired quite regularly from both orbits. Accordingly, the combination has been carried out on ascending and descending data that are temporally as close as possible each other. This operation is based on the assumption that no significant deformation has occurred between the two combined acquisitions, and this is reasonable for our case study.
Figure 3. Distribution of ENVISAT acquisitions (red points), used in the SBAS processing, in the temporal (x-axis)—perpendicular (y-axis) baseline plane for the descending (A); and ascending (B) dataset. Black arcs show the computed interferograms.

3.2. Groundwater Level and Stratigraphic Data

Groundwater level data have been provided by the General Directorate of State Hydraulic Works of Turkey, which is in charge of monitoring the piezometric wells widely distributed along the Konya plain and the Obruk Plateau (Figure 1, Table 2). We used data recorded in nine wells (Table 3) which, overall, provide information on the characteristics of the whole Konya plain aquifer and on the areal extension of the groundwater depletion.

For the analysis of the correlation between groundwater level changes and surface movements we focused on the piezometric measurements carried out monthly in the period 2003–2010, thus roughly matching the period of SBAS-derived deformation time-series.
As regards the stratigraphy of the area, we used data reported in [9] distributed within the Konya plain and the Obruk Plateau (Figure 1, Table 3). The thickness of compressible layers was derived by considering the thickness of both clay and silt layers, located within the alluvial sediments of the plain, which usually influence the processes of subsidence related to groundwater extraction [31,32].

Table 3. Piezometric wells and stratigraphic cores available for this study (for location see Figure 1).

<table>
<thead>
<tr>
<th>Well/Core ID</th>
<th>Latitude (°)</th>
<th>Longitude (°)</th>
<th>Altitude (m a.s.l.)</th>
<th>Well/Core Depth (m)</th>
<th>Lithological Unit</th>
<th>Compressible Layers Thickness (m)</th>
<th>Groundwater Level Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>181</td>
<td>37.620</td>
<td>32.751</td>
<td>1014 n/a</td>
<td>alluvial deposits</td>
<td>n/a</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>9434</td>
<td>37.770</td>
<td>32.450</td>
<td>1048 n/a</td>
<td>alluvial deposits</td>
<td>n/a</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>10472</td>
<td>38.078</td>
<td>32.644</td>
<td>1000 n/a</td>
<td>alluvial deposits</td>
<td>n/a</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>20826</td>
<td>37.787</td>
<td>32.995</td>
<td>1004 n/a</td>
<td>alluvial deposits</td>
<td>n/a</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>42778</td>
<td>37.948</td>
<td>33.841</td>
<td>1075 n/a</td>
<td>alluvial deposits</td>
<td>n/a</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>52267</td>
<td>37.486</td>
<td>32.911</td>
<td>1025 n/a</td>
<td>alluvial deposits</td>
<td>n/a</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>52268</td>
<td>37.552</td>
<td>33.185</td>
<td>1013 n/a</td>
<td>alluvial deposits</td>
<td>n/a</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>2889</td>
<td>37.694</td>
<td>33.254</td>
<td>999</td>
<td>alluvial deposits</td>
<td>120</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>4566</td>
<td>37.769</td>
<td>33.840</td>
<td>996</td>
<td>alluvial deposits</td>
<td>180</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>4685</td>
<td>37.672</td>
<td>33.381</td>
<td>1001</td>
<td>alluvial deposits</td>
<td>150</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>35530</td>
<td>37.710</td>
<td>33.387</td>
<td>1006</td>
<td>alluvial deposits</td>
<td>100</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>38306</td>
<td>37.983</td>
<td>33.334</td>
<td>1046</td>
<td>Limestone/marl</td>
<td>100</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>42289</td>
<td>37.691</td>
<td>33.400</td>
<td>1006</td>
<td>alluvial deposits</td>
<td>80</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>43405</td>
<td>37.790</td>
<td>33.286</td>
<td>1062</td>
<td>Limestone/marl</td>
<td>135</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>43666</td>
<td>37.713</td>
<td>33.232</td>
<td>1007</td>
<td>alluvial deposits</td>
<td>170</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>47768</td>
<td>37.801</td>
<td>33.433</td>
<td>1040</td>
<td>Limestone/marl</td>
<td>120</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>47782</td>
<td>37.850</td>
<td>33.344</td>
<td>1070</td>
<td>Limestone/marl</td>
<td>130</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>48083</td>
<td>37.739</td>
<td>33.474</td>
<td>1047</td>
<td>Limestone/marl</td>
<td>95</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>49703</td>
<td>37.752</td>
<td>33.248</td>
<td>1025</td>
<td>alluvial deposits</td>
<td>140</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>51903</td>
<td>37.768</td>
<td>33.311</td>
<td>1055</td>
<td>Limestone/marl</td>
<td>140</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>47761</td>
<td>37.583</td>
<td>33.514</td>
<td>1002</td>
<td>alluvial deposits</td>
<td>100</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>52258</td>
<td>37.717</td>
<td>33.471</td>
<td>1014</td>
<td>limestone/marl</td>
<td>175</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

3.3. Meteorological Data

We used precipitation (P) and temperature (T) data acquired monthly by five meteorological stations of the Turkish State Meteorological Service (Table 4) [50,51]. In particular, for the analysis of climatic conditions during the 2002–2010 period covered by ENVISAT acquisitions we averaged data recorded in all five stations; for a more long-term analysis of the climate trend in the area we exploited the rainfall and temperature measurements collected by the Konya station since 1935 and 1950, respectively.

Table 4. Meteorological stations used for the analysis (for location see Figure 1).

<table>
<thead>
<tr>
<th>Station ID</th>
<th>Location</th>
<th>Latitude (°)</th>
<th>Longitude (°)</th>
<th>Altitude (m a.s.l.)</th>
<th>Measurement Time (Years)</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>17902</td>
<td>Karapinar</td>
<td>37.7136</td>
<td>33.5277</td>
<td>997</td>
<td>2000–2014</td>
<td>P-T</td>
</tr>
<tr>
<td>17244</td>
<td>Konya</td>
<td>37.8691</td>
<td>32.4721</td>
<td>1032</td>
<td>1935–2014</td>
<td>P</td>
</tr>
<tr>
<td>17248</td>
<td>Ereğli</td>
<td>37.5265</td>
<td>34.0482</td>
<td>1044</td>
<td>1990–2014</td>
<td>T</td>
</tr>
<tr>
<td>17900</td>
<td>Çumra</td>
<td>37.5662</td>
<td>32.7932</td>
<td>1016</td>
<td>2000–2014</td>
<td>P-T</td>
</tr>
<tr>
<td>17264</td>
<td>Karaman</td>
<td>37.1927</td>
<td>33.2212</td>
<td>1021</td>
<td>2000–2014</td>
<td>P-T</td>
</tr>
</tbody>
</table>

3.4. Landsat Data and Land Cover Maps

For the analysis of the land cover (LC) changes occurred in the study area, we used CORINE Land Cover (CLC) maps relevant to the years 2000 and 2012 [52,53]. For our purposes, the legend of the CLC maps was simplified to 8 classes as shown in Table 5; the agricultural classes (5)–(7) were further classified on the basis of irrigated/non-irrigated lands.
Table 5. Simplified CORINE land cover classes.

<table>
<thead>
<tr>
<th>Simplified Classes</th>
<th>CORINE Classes Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Artificial surfaces</td>
<td>Urban fabric, industrial, transport and commercial units, mine, dump and construction sites</td>
</tr>
<tr>
<td>(2) Natural Vegetation</td>
<td>Natural grasslands, moors and heathland, Sclerophyllous vegetation, transitional woodland-shrub, inland marshes, peat bogs, salines, intertidal flats, broad-leaved forest, coniferous forest, mixed forest</td>
</tr>
<tr>
<td>(3) Bare rocks/soils</td>
<td>Bare rocks, beaches, dunes, sands, salt marshes</td>
</tr>
<tr>
<td>(4) Water bodies/Wetlands</td>
<td>Water courses, water bodies, coastal lagoons, estuaries, sea and ocean, inland marshes, peat bogs, salt marshes, salines, intertidal flats</td>
</tr>
<tr>
<td>(5) Non-irrigated arable land</td>
<td>Non-irrigated arable land</td>
</tr>
<tr>
<td>(6) Unknown or mixed irrigated/non-irrigated land</td>
<td>Complex cultivation patterns, land principally occupied by agricultural with natural space, vineyards and fruit tree</td>
</tr>
<tr>
<td>(7) Irrigated land</td>
<td>Irrigated land/rice field</td>
</tr>
<tr>
<td>(8) Pastures</td>
<td>Pastures</td>
</tr>
</tbody>
</table>

Moreover, we exploited Landsat 5 TM and Landsat 8 scenes acquired, respectively, in May 2000 and May 2015 in order to analyze the irrigated land changes. Landsat images were available from the USGS Glovis Archive [54] as Standard Terrain Correction products (Level 1T—precision and terrain correction). For our analysis we used surface reflectance in the blue, near-infrared and shortwave-infrared wavelength domains (Table 6). We generated false-color images by calibrating the brightness of each band to obtain similar colors for areas without any change between 2000 and 2015. Moreover, we produced classified Landsat images by applying a supervised classification based on the maximum likelihood method.

We have also calculated and calibrated a modified Normalized Difference Vegetation Index (mNDVI) for the two Landsat images in order to assess the changes of vegetation status in the main CORINE land cover types.

Table 6. Landsat band combination used for the analysis.

<table>
<thead>
<tr>
<th>Color</th>
<th>Domain</th>
<th>Landsat 5 TM</th>
<th>Landsat 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Band</td>
<td>Wavelength (μm)</td>
<td>Band</td>
</tr>
<tr>
<td>Blue</td>
<td>1</td>
<td>0.45–0.52</td>
<td>2</td>
</tr>
<tr>
<td>Green</td>
<td>NIR</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Red</td>
<td>SWIR</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

3.5. GRACE Water Storage Data

The GRACE satellite provides an estimate of the water storage anomaly by measuring the variation of the gravity field which is related to the variation of groundwater, soil moisture, snow coverage and water reservoirs.

In order to assess the groundwater storage variation, we used Level-2 data acquired from 2002 to 2010, available from the University of Colorado GRACE Data Analysis Website [55]. The used processed data represent the terrestrial water storage (TWS) anomaly (expressed in cm) calculated on a GRACE cell of about 100 km radius with a 25 km smooth filter. As reported by [56], GRACE is sensitive to TWS variation over areas smaller than its original footprint (400 km × 400 km) if such variation is significant. For instance, GRACE observations allow to detect 1 cm of TWS change within a basin of 200,000 km² (i.e., 2 km³ TWS volume change). Therefore, its detectability is of the order of 2 m water level change per 1000 km².

The Konya plain extends for about 6000 km² and the average groundwater variation is about 10 m, which results to be detectable by the GRACE cell. Moreover, we have corrected the GRACE
data by using data available from the Global Land Data Assimilation System (GLDAS) [57] in order to reduce the effect of soil moisture and snow water equivalent (SWE) masses. The residual variation of TWS can be associated to groundwater and surface water masses.

We remark that GRACE data are used to integrate the information from the available piezometric wells, and to show that the water resources depletion is an issue to be addressed not only in our study area but at regional scale.

4. Results

4.1. SBAS-DInSAR Deformation Maps and Time-Series

Maps of the average deformation rate and, for each pixel, displacement time-series spanning the December 2002–July 2010 time interval have been produced (Figures 4 and 5), allowing to detect a wide land subsidence extending for about 1200 km² (see the black dashed line in Figure 4) and measure its magnitude.

![Figure 4](image-url)

Figure 4. SBAS-DInSAR deformation rate maps of the study area, superimposed on hill shade image derived from DEM. Maps have been produced by processing ENVISAT data acquired in the December 2002–July 2010 time interval. (A,B) show displacements measured along the satellite LOS direction during ascending and descending passes, respectively; (C,D) show displacement components computed along the vertical and (E-W) horizontal directions, respectively. Black dashed line represents the contour of the DInSAR derived subsidence area.

Maps showing rates and spatial pattern of surface deformations observed from both satellite orbits (i.e., descending and ascending) are reported in Figure 4A,B. Displacements are measured along the LOS of the satellite: negative values (yellow to red color) represent movements away from satellite (i.e., land subsidence) whereas positive values (cyan to blue) represent movements towards the satellite (i.e., land uplift). Green-color pixels are characterized by velocities within the range of ±0.25 cm/year and are considered stable. DInSAR measurements are related to a reference point.
located in an area assumed to be geologically stable, thus not susceptible to subsidence processes (for location see black star in Figure 4). The achieved results show an accuracy of 0.8 cm for the displacements and 0.15 cm/year for the deformation rate.

In order to ensure reliable results, we selected SAR pixels characterized by a temporal coherence greater than 0.8. However, despite the high coherence threshold applied, the SBAS processing allowed to achieve a density of 17 and 21 measure points/km² for descending and ascending datasets, respectively. Such values are considerably large if we consider that it is an area lacking the radar targets provided by man-made features, highlighting the capability of the SBAS-DInSAR approach to effectively investigate displacements in both rural and urban settings.

Maps of the vertical and (E-W) horizontal components of the surface displacements are shown in Figure 4C,D. The horizontal movement component can be considered negligible confirming that the regional deformation process mainly occurs along the vertical direction (Figure 4C). This is also clearly visible from the inspection of the achieved time-series revealing that significant vertical displacements reaching up to 10 cm have occurred, with an average deformation rate of about 1.2 cm/year for the period 2002–2010 (Figure 5).

![Figure 5](image)

**Figure 5.** Average deformation time-series covering the December 2002–July 2010 time interval. (A,B) show LOS displacements measured during ascending and descending satellite passes, respectively; (C,D) show the computed vertical and (E-W) horizontal components of the displacements, respectively. Note that time-series have been obtained by averaging deformation time-series related to SAR measure points located within the detected subsidence area.

### 4.2. Land Cover Change Analysis

The analysis of the land cover (LC) changes occurred in the area during the observation period covered by the SBAS-DInSAR measurements was carried out by using two CLC maps from 2000 and 2012 [53]. Figure 6 shows the spatial distribution of the CLC classes within the whole Konya plain, in 2000 and 2012, and the increase of irrigated land during this period (see blue polygons in Figure 6).
Figure 6. CORINE Land Cover simplified maps of the Konya plain for years 2000 and 2012. Blue polygons in 2012 map show areas changed from non-irrigated to irrigated.

Table 7 reports the LC changes affecting the area during the 12-year analyzed period. It is interesting to note that irrigated lands, pastures and urban areas (i.e., artificial surfaces) increase while natural vegetation, unknown/not irrigated land and water bodies decrease. Such detected land cover changes are usually associated to an increase of demand of water supply which main source, in this area, is represented by groundwater [5]. Indeed, the increasing of crops (e.g., sugar beet and corn) translates into more irrigation needs as well as the growth of urban areas (i.e., of population and industry) and pastures (i.e., of breeding) leads to higher water consumption. Land changes that may be correlated to groundwater overexploitation are also analyzed by using a Landsat 5 TM image acquired on May 2000 and a Landsat 8 image taken on May 2015 (Figure 7A,B). Visual inspection of the two images reveals a clear pattern of crops in 2015 which is not visible in 2000, pointing out an increase of land covered by intensive agriculture and irrigated fields over the 15-year period of analysis. Such an observation is not straightforward for the area around Hotamis, probably due to a change in cultivation type that is not visible in the 2015 image. Moreover, a supervised classification of SAGA GIS based on the maximum likelihood method (Figure 7E,F) was applied. Landsat images were classified into four land cover types: (1) water bodies; (2) irrigated land; (3) not irrigated land/natural vegetation/pastures; (4) rock/bare soil/salt. Classification results, within the sample area, are in agreement with CORINE data, showing that irrigated lands increased from 7% to 14% while not-irrigated lands decreased from 26% to 11% (Table 8).

Table 7. Land cover changes computed by using the 2000 and 2012 CLC maps (see Figure 6).

<table>
<thead>
<tr>
<th>Land Cover Classes</th>
<th>Area (km²)</th>
<th>Change (km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2000</td>
<td>2012</td>
</tr>
<tr>
<td>Irrigated land</td>
<td>3723.4</td>
<td>3843.9</td>
</tr>
<tr>
<td>Unknown or mixed irrigated/non-irrigated land</td>
<td>509.1</td>
<td>441.0</td>
</tr>
<tr>
<td>Non-irrigated arable land</td>
<td>1049.4</td>
<td>981.0</td>
</tr>
<tr>
<td>Pastures</td>
<td>872.1</td>
<td>970.1</td>
</tr>
<tr>
<td>Natural vegetation</td>
<td>1351.2</td>
<td>1227.3</td>
</tr>
<tr>
<td>Bare soil/Rock</td>
<td>47.5</td>
<td>156.3</td>
</tr>
<tr>
<td>Artificial surfaces</td>
<td>400.5</td>
<td>407.7</td>
</tr>
<tr>
<td>Water bodies/Wetland</td>
<td>149.7</td>
<td>75.5</td>
</tr>
<tr>
<td>Total Area</td>
<td>8103</td>
<td>8103</td>
</tr>
</tbody>
</table>
We also computed the mNDVI for May 2000 and May 2015 which shows a general increase of vegetation activity in the cultivated land (Figure 7C,D). Table 9 reports the relative variation of mNDVI for some categories of CLC. Irrigated land shows the highest mNDVI. New irrigated land shows the highest increase of mNDVI along with the non-irrigated land, this latter probably partially irrigated. The other categories of land cover do not show significant changes of mNDVI, pointing out that the vegetation activity increase is probably more related to irrigation than to climate conditions, as also reported in [5] using MODIS data.

Figure 7. False color images of Landsat 5 TM data (R = Band 5; G = Band 4; B = Band 1) acquired on May 2000 (A) and of Landsat 8 data (R = Band 6; G = Band 5; B = Band 2) captured on May 2015 (B). (C,D) mNDVI images associated to (A,B). Blue polygons show areas changed from non-irrigated to irrigated (see Figure 6). Classified Landsat 5 TM (E) and Landsat 8 (F) images.
Table 8. Land cover changes computed by using classified Landsat images (see Figure 7).

<table>
<thead>
<tr>
<th>Land Cover Classes</th>
<th>Area (km²)</th>
<th>2000</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irrigated land</td>
<td></td>
<td>760</td>
<td>1634</td>
</tr>
<tr>
<td>Non-irrigated land/Natural vegetation/Pasture</td>
<td></td>
<td>3003</td>
<td>1271</td>
</tr>
<tr>
<td>Rock/Bare soils/Salt</td>
<td></td>
<td>7538</td>
<td>8389</td>
</tr>
<tr>
<td>Water bodies</td>
<td></td>
<td>20</td>
<td>6</td>
</tr>
<tr>
<td>Not classified</td>
<td></td>
<td>59</td>
<td>81</td>
</tr>
<tr>
<td>Total Area</td>
<td></td>
<td>11,380</td>
<td>11,380</td>
</tr>
</tbody>
</table>

Table 9. Average mNDVI in 2000 and 2015 for the main CLC categories.

<table>
<thead>
<tr>
<th>CLC Category</th>
<th>mNDVI</th>
<th>2000</th>
<th>2015</th>
<th>Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural vegetation</td>
<td>0.45</td>
<td>0.45</td>
<td>–0.001</td>
<td></td>
</tr>
<tr>
<td>Bare soil Rock</td>
<td>0.28</td>
<td>0.22</td>
<td>–0.068</td>
<td></td>
</tr>
<tr>
<td>Pastures</td>
<td>0.47</td>
<td>0.48</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>Already irrigated land</td>
<td>0.89</td>
<td>0.96</td>
<td>0.068</td>
<td></td>
</tr>
<tr>
<td>New irrigated land (2012)</td>
<td>0.66</td>
<td>0.98</td>
<td>0.325</td>
<td></td>
</tr>
<tr>
<td>Non-irrigated arable land</td>
<td>0.52</td>
<td>0.74</td>
<td>0.216</td>
<td></td>
</tr>
<tr>
<td>Artificial surfaces</td>
<td>0.57</td>
<td>0.58</td>
<td>0.011</td>
<td></td>
</tr>
</tbody>
</table>

4.3. Meteorological Analysis

The analysis of data recorded by the available meteorological stations (Figure 1, Table 4) allowed the identification of precipitation (P) and temperature (T) variations occurred in the area.

Analysis of temperature (Figure 8A), based on computing the average annual temperature data acquired by Konya station from 1950 to 2014, shows an increase of 0.75 °C during the observation period, with an higher rate starting from 1990. This increase is in agreement with the general global warming and with findings from climate-related studies [2,58,59]. Such a temperature increase may contribute to the rise of water needs for irrigation.

Analysis of precipitation, carried out by computing 1-year as well as 3-year cumulative rainfall data collected by Konya station from 1935 to 2014, did not reveal any clear trend (Figure 8B). However, it is possible to observe cycles of dry and wet periods, probably related to NAO index [60]. In particular, by analyzing the 3-year cumulative rainfall graph, it is possible to observe two minimum rainfall periods, in the early 1990s and in the 2004–2008 period, which partially overlap with the ENVISAT observation period (2002–2010).

Moreover, we performed an analysis on precipitation by averaging data acquired by all the five available stations during the 2002–2010 period (Figure 8C). The five stations show similar climate conditions from 2000 to 2010 with a mean annual rainfall ranging from 283 mm of Karapinar to 330 mm of Cumra. The analysis of 3-months cumulative precipitation data pointed out a typical seasonal trend, with less than 25 mm in July–September and up to 120 mm in November–January, while the 1-year cumulative precipitation data analysis revealed a minimum rainfall period from 2004 to 2008, followed by a wet period in 2009–2010.
5. Discussion

In this Section, we analyze the SBAS-DInSAR results in relation to the information derived from independent (remote sensed and ground-based) data, in order to understand the role of different factors (i.e., type and thickness of lithology, groundwater depletion induced by natural causes and human activities) in predisposing and/or triggering land subsidence (Figure 9).
Figure 9. (A) Subsidence area revealed by the SBAS-DInSAR analysis; (B) Thickness of compressible layers (i.e., clay and silt) from the available cores. Location of the a-b geological profile reported in Figure 10 is also shown; (C) Groundwater level changes measured in the available wells.

5.1. Correlation between Land Subsidence and Lithology

As already mentioned in Section 2, alluvial sediments, mainly made up of clay and silt interbedded by sandstone and gravel, are widely present in the Konya plain. Such clay and silt units are
characterized by high compressibility, playing as predisposing factor in the processes of subsidence induced by groundwater extraction.

Due to the number and distribution of stratigraphic cores within the plain, it was not possible to quantitatively investigate the correlation between subsidence and lithology. Only six of the available cores are characterized by a significant thickness of compressible layers (which reaches up to 130 m) and, therefore, the regression analysis carried out to estimate the relationship between subsidence rate and thickness of such cores is not meaningful. However, the joint analysis of Figure 9A,B allows a qualitative estimate of the relationship between the spatial distribution of DInSAR derived surface deformations and the thickness of compressible layers.

Such relationship is better pointed out in Figure 10 showing a simplified geological section against the vertical deformation rate along a profile extending from Hotamis to Karapinar (for location see Figure 9). A spatial correlation between land subsidence and alluvial deposits can be observed. Moreover, fastest subsidence affects the areas with thickest sediments as shown in Table 10, which reports the vertical deformation rate in relation to the thickness of compressible units.

![Figure 10. (A) Simplified geological profile (see Figure 9 for location) vs. (B) profile of DInSAR vertical deformation rate. The groundwater level for 2002 and 2010 is estimated from the nearest piezometric well (ID 52258).](image)

<table>
<thead>
<tr>
<th>Compressible Layer Thickness (m)</th>
<th>Number of Wells</th>
<th>Vertical Deformation Rate (cm/Year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obruk plateau</td>
<td>3</td>
<td>Min 0.03 Max 0.25 Mean 0.15</td>
</tr>
<tr>
<td>&lt;10</td>
<td>2</td>
<td>Min 0.06 Max 0.47 Mean 0.26</td>
</tr>
<tr>
<td>50–100</td>
<td>2</td>
<td>Min 0.30 Max 0.56 Mean 0.41</td>
</tr>
<tr>
<td>&gt;100</td>
<td>2</td>
<td>Min 0.43 Max 0.49 Mean 0.46</td>
</tr>
</tbody>
</table>

5.2. Correlation between Land Subsidence and Groundwater Depletion

Groundwater level variations in the study area have been investigated by analyzing data acquired by a local network of piezometric wells. Figure 9C shows the cumulative water table depletion during the period 2003–2010, estimated from wells located in the subsidence area. Wells recorded a significant
decrease of the piezometric level up to about 14 m, except for well 47,761 located at south of Karapinar, which measured a lower decrease (about 2 m).

Analysis of time-series of piezometric data reveals a seasonal variation reaching a maximum piezometric level approximately at the end of winter season (end of March) and a minimum piezometric level in October (Figure 11A–D). Such a variation, which can be related to climatic conditions and irrigation cycles, is observed in all the wells located along the plain expect for, as aforementioned, the well 47761 which shows not significant groundwater changes, probably due to the presence of a local shallow aquifer (Figure 11E).

Figure 11. (A–E) Groundwater level measured in four wells; (F) Water storage anomaly changes derived from GRACE data.
As regards the general trend, it shows a maximum decrease of piezometric level from 2004 to 2008, which corresponds to a minimum rainfall period (Figure 8), as pointed out from the meteorological data analysis (see Section 4.3). Subsequently, despite the 2009–2010 time interval was affected by heavy precipitation events (Figure 8), the groundwater table level remains stationary or shows a little increase, due to a condition of aquifer overexploitation.

Groundwater level changes occurred at regional scale have been analyzed by using GRACE satellite data. The water storage anomaly estimated by GRACE (Figure 11F) shows a trend in general agreement with wells data. In particular, a seasonal trend (about 150 mm of storage water anomaly oscillation), probably also related to surface water (reservoir), is present. This is superimposed to a general trend showing a decrease of storage level until 2008 (about −100 mm of water storage) followed by a weak increase in the 2009–2010 period. It is important to remark that GRACE observations are characterized by a footprint larger than our study area and may be influenced by various processes beyond the groundwater level variations (see Section 3.5), thus cannot be used for a direct comparison with the piezometric measurements available for our study area. However, they can be profitably exploited for a qualitative assessment of the water resources depletion occurring at regional scale.

As regards the piezometric measurements, we noticed that a groundwater level decrease is measured in the wells located along the Konya plain and on the Obruk plateau as well. However, such a groundwater drawdown had a different impact on the surface in terms of deformation.

To investigate the susceptibility of these two geological settings (i.e., plain and plateau) to ground instability, we performed a joint analysis of DInSAR deformation time-series and piezometric data time-series for four wells, two located in the alluvial Konya plain (wells 52268 and 47761) and two on the limestone Obruk plateau (wells 42278 and 52258; Figures 9 and 12). We remark that, despite the uncertainties related to the few available data which hampered us to carry out a quantitative correlation analysis between land subsidence and groundwater level changes, our analysis allowed pointing out the concomitant effect of lithological characteristics and groundwater depletion on surface displacements.

The well located on the limestone plateau (Figure 12A) and the well located on the boundary between plateau and plain (Figure 12B) recorded a significant decrease of the piezometric level reaching about 2 m and 12 m, respectively. However, the deformation times-series associated to SAR points located near these two wells show a quasi-stable trend with a subsidence rate lower than 0.2 cm/year. Therefore, such result shows no correlation between groundwater drawdown and subsidence process in the area where limestone crops out.

On the other hand, as expected, in the plain area where soft soils are present, a correlation between piezometric level changes and land subsidence is observed. This is particularly evident by comparing time-series of DInSAR deformations and piezometric data acquired by the 52268 well (Figure 12C). For this well, the groundwater level decreases of about 12 m during the 2003–2010 period and the times-series of near SAR points reveal that, in this time interval, a cumulated displacement of about 5 cm occurs in the compressible layer of alluvial sediments. The high deformability of such materials is confirmed by the DInSAR time-series of SAR points located in the proximity of well 47761 (Figure 12D). Indeed, such well, probably due to local differences of the aquifer characteristics, recorded a low piezometric level decrease (<2 m) but it is located in an area affected by moderate subsidence (0.4 cm/year).
Figure 12. Average of DInSAR vertical deformation time-series (red squares) associated to SAR measure points located in similar geological and morphological units, within a distance variable from 1 to 2 km from the relevant well, vs. groundwater level measurements (blue squares). (A,B) refer to wells located on the limestone plateau and on the boundary between plateau and plain area, respectively. (C,D) refer to wells located within the plain area.

5.3. Karst Processes and Land Subsidence

Although the groundwater drawdown has not caused detectable progressive subsidence where limestone crops out, subsidence is also present in the latter area in the form of karst collapses. As mentioned above, these processes represent a special hazard in the study area and they are also linked with the decline of the groundwater levels [9]. Unfortunately, we have not detected displacements associated to dissolution-induced subsidence, even though the time span covered by SAR measurements (2002–2010) has coincided with a period of high collapse occurrence (13 collapses between 2006 and 2009). Castañeda et al. [61] were able to detect karst subsidence in the Ebro Valley (Spain) with a similar resolution (90 × 90 m²) and using the same DInSAR processing technique. However, the karst processes described in their study area are widely different from those observed in the Konya basin. The Ebro Valley shows an evaporite karst where large depressions (>50,000 m²) are developed because of a progressive subsidence associated to the dissolution of gypsum, anhydrite, and most specially, glauberite and halite. These highly soluble minerals are responsible of a rapid dissolution phenomenon that accelerates the subsidence and influences its displacement regime (continuous, episodic or mixed). This is not the case of the Konya basin where a limestone karst
has been developed and the observed landforms indicate that the processes that control the karst subsidence seem to have an episodic regime (i.e., collapse). The large sinkhole collapses (the “old” obruks) and their surroundings appeared to be inactive as our DInSAR results indicate. On the other hand, the spatial resolution of our displacement maps is a priori inadequate for detecting movements associated to the “young” obruks. Baer et al. [62] based on ERS data of similar resolution failed to detect sinkhole induced deformation along the Dead Sea shores, an area widely affected by karst subsidence processes. Galve et al. [26] show that the minimum detectable sinkhole area in their test site, using a technique that provides 369.8 points/km², was approximately 2500 m². Our analysis provides 17–21 measure points/km² and the extent of the largest “young” obruk (<500 m²) is also below the mentioned minimum detectable dimensions. Nof et al. [63] demonstrate that small sinkhole induced deformation or collapse precursory displacements can be recognized with DInSAR techniques using high resolution data such as the SAR images provided by the COSMO-SkyMed satellite (pixel size of 3 m). Moreover, a higher spatial resolution is needed for recognizing sinkhole collapse precursory displacements. Nevertheless, the displacement maps produced in this research could be used for assessing the sinkhole hazard in the study area. As the “young” obruks are related to the lowering of the water table, maps of groundwater drawdown could be derived from our displacement maps and the latter could be used for assessing the areas with the highest sinkhole hazard.

6. Conclusions

Combining remote-sensed and in situ data and information, we explored the relationship between land subsidence and its main controlling factors in one of the most vulnerable regions of Turkey, the Konya plain.

Through space-borne SBAS analyses, we detected a wide subsidence phenomenon extending for about 1200 km² in the plain, with a deformation rate up to 1.5 cm/year in the 2002–2010 period. The SBAS technique allowed achieving a significant density of measure points, demonstrating the capability of this DInSAR approach to effectively investigate displacements also in semi-rural settings. Moreover, the availability of ascending and descending SAR datasets allowed us to investigate the direction of the ground movement and retrieve vertical displacements reaching up to 10 cm in the observed time interval.

The specific role of the geological and hydrogeological conditions has been explored, revealing that the consolidation of compressible alluvial sediments associated to the groundwater level decrease drives the subsidence and controls its temporal and spatial evolution over the plain. Groundwater lowering is also one of the factors that trigger sinkholes in the Obruk plateau causing damage to infrastructures and villages.

Groundwater depletion is mainly caused by overexploitation due to uncontrolled irrigation and unsustainable agricultural practices, as pointed out by our preliminary land cover change analysis showing an increase of irrigation fields of about 120 km² during the 2000–2012 period. The analysis of data acquired by piezometric wells revealed a groundwater level decrease up to about 14 m in correspondence with the minimum precipitation period (from 2004 to 2008) detected by analyzing meteorological data. Moreover, the analysis of climatic data shows an increase of temperature of 0.75 °C in the last sixty-five years, with an higher rate starting from early 1990s; such an increase probably further contributed to the rise of water needs for irrigation.

The use of GRACE data allowed us to point out that the groundwater depletion is not limited to our study area but affects a wider region in the Anatolian Plateau, demonstrating that such satellite data can be profitably exploited for assessing water resources at regional scale.

Our study has relevant local and global implications. Locally, it provides a holistic view over the anthropogenic and environmental dynamics in the Konya plain, encouraging the development of an integrated monitoring system which may play a relevant role for the protection of the area. At a more global scale, our approach may be extended to investigate land subsidence induced by groundwater overexploitation in different climatic and geological settings worldwide.
Acknowledgments: The work was financed by the project Space Advanced Project Excellence in Research and Enterprise (SAPERE) within the Italian Program “Sviluppo e Potenziamento dei Cluster tecnologici Nazionali”. The activities were carried out through the Infrastructure of High Technology for Environmental and Climate Monitoring-PONa3_00363 project of Structural Improvement financed under the National Operational Programme for “Research and Competitiveness 2007–2013,” supported with the European Regional Development Fund and National Resources. ENVISAT ASAR data were provided by European Space Agency (ESA) in the framework of the CAT-1 ESA No. 10050 project.

Author Contributions: F.C. performed SAR data processing and analyses, and wrote the paper; D.N. analyzed the data and contributed to write the paper; J.P.G., S.A., T.G., A.P. and F.B.S. contributed to analyze the data and write the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Şen, O.L.A. Holistic View of Climate Change and Its Impacts in Turkey; IPC (Istanbul Policy Center), Sabanci University: Istanbul, Turkey, 2013.
5. Gokmen, M.; Vekerdy, Z.; Lubczynski, M.W.; Timmermans, J.; Batelaan, O.; Verhoef, W. Assessing groundwater storage changes using remote sensing-based evapotranspiration and precipitation at a large semiarid basin scale. J. Hydrometeorol. 2013, 14, 1733–1753. [CrossRef]
10. Ozdemir, A. Investigation of sinkholes spatial distribution using the weights of evidence method and GIS in the vicinity of Karapinar (Konya, Turkey). Geomorphology 2015, 245, 40–50. [CrossRef]


42. Kocyigit, A. The Denizli graben-horst system and the eastern limit of Western Anatolia continental extension: Basin fill, structure, deformational mode, throw amount and episodic evolutionary history SW Turkey. Geol. Acta 2015, 18, 167–208. [CrossRef]


50. KNMI Climate Explorer. Available online: https://climexp.knmi.nl/getprcpall.cgi?id=someone@somewhere&WMO=17244 (accessed on 18 May 2016).


56. Longuevergne, L.; Wilson, C.; Scanlon, B.R.; Crétaux, J.F. GRACE water storage estimates for the Middle East and other regions with significant reservoir and lake storage. Hydrol. Earth Syst. Sci. 2013, 17, 4817–4830. [CrossRef]


Coastal Subsidence Monitoring Associated with Land Reclamation Using the Point Target Based SBAS-InSAR Method: A Case Study of Shenzhen, China

Bing Xu 1, Guangcai Feng 1,*, Zhiwei Li 1, Qijie Wang 1, Changcheng Wang 1 and Rongan Xie 2

1 School of Geosciences and Info-Physics, Central South University, Changsha 410083, China; xubing@csu.edu.cn (B.X.); zwli@csu.edu.cn (Z.L.); qjwang@csu.edu.cn (Q.W.); wangchangcheng@csu.edu.cn (C.W.)
2 Guangdong Institute of Surveying and Mapping of Geology, Guangzhou 510800, China; 13902396961@139.com

* Correspondence: fredgps@gmail.com; Tel.: +86-731-888-30573

Academic Editors: Zhenhong Li, Roberto Tomas, Zhong Lu and Prasad S. Thenkabail
Received: 30 May 2016; Accepted: 5 August 2016; Published: 13 August 2016

Abstract: Shenzhen, the first special economic zone of China, has witnessed earth-shaking changes since the late 1980s. In the past 35 years, about 80 km² of land has been reclaimed from the sea in Shenzhen. In order to investigate coastal vertical land motions associated with land reclamation, we proposed an elaborated Point Target (PT) based Small Baseline Subset InSAR (SBAS-InSAR) strategy to process an ENVISAT ASAR ascending and descending orbits dataset both acquired from 2007 to 2010. This new strategy can not only select high density PTs but also generate a reliable InSAR measurement with full spatial resolution. The inter-comparison between InSAR-derived deformation velocities from different orbits shows a good self-consistency of the results extracted by the elaborated PT-based SBAS-InSAR strategy. The InSAR measurements show that the reclaimed land is undergoing remarkable coastal subsidence (up to 25 mm/year), especially at the Shenzhen Airport, Bao’an Center, Qianhai Bay and Shenzhen Bay. By analyzing the results together with land reclamation evolution, we conclude that the ground deformation is expected to continue in the near future, which will amplify the regional sea level rise.

Keywords: coastal subsidence; land reclamation; PT-based SBAS-InSAR; sea level rise; Shenzhen

1. Introduction

The Coastal Zone [1] plays a very important role in transportation, the circulation of resources and funds, etc. To promote urbanization and economic development, land reclamation [2] from the sea is a common practice for many countries in the world, such as USA [3], Japan [4], Italy [5], the Netherlands [6,7], and China [8–10].

Over the past decades, Synthetic Aperture Radar Interferometry (InSAR) [11,12] has been proven to be a powerful geodetic technique, and been extensively exploited to investigate ground deformation induced, for example, by earthquake [13–15] or volcano activities [16,17]. Furthermore, recently remarkable improvements in Multi-Temporal InSAR (MT-InSAR), including the Small Baseline Subset InSAR (SBAS-InSAR) [18,19] and the Persistent Scatterer InSAR (PS-InSAR) [20–23], enable us to monitor long period surface deformation caused by permafrost [24], ground water extraction [25–28], mining [29], and land reclamation [30]. The MT-InSAR exploits a set of high phase-quality pixels [20], other than the original full two-dimensional image grid, to map the deformation history and the corresponding average deformation velocity. It can overcome most intrinsic temporal and geometric
decorrelations [25] of conventional InSAR, and simultaneously estimate the atmospheric effect [26–29] to generate high accuracy deformation results. In recent years, researches that investigated the ground deformation related to land reclamation by MT-InSAR have been carried out in reclaimed lands around the world, such as the Chek Lap Kok Airport in Hong Kong [30], Sibari in Southern Italy [31], and the Venice coastland [32]. Besides the MT-InSAR techniques, geodetic techniques [33], such as GPS and leveling, have been widely used in ground deformation monitoring. However, the MT-InSAR could provide a better spatial resolution to capture the deformation details compared with GPS or leveling.

Shenzhen is a city built on the weak alluvial clay [34] of sludge and reclaimed land. Due to high compressibility, low permeability, and weak strength of the alluvial clay [35], the soil becomes compacted under the pressure from constructions and clay self-consolidation. The compaction process causes ground deformation [36], threatening the stability of ground constructions and underground man-made facilities [37]. Therefore, ground deformation monitoring is very important for the reclaimed land utilization in Shenzhen. Dense Point Targets (PTs) for deformation monitoring can be selected from built-up area [20] that maintains long-time high coherence, while for the non-built-up area, e.g., reclaimed land, for the area surrounding buildings built on the reclaimed land, few PTs can be selected [38]. Although previous studies [36,39] have tried to capture the deformation in urban areas with MT-InSAR, the following two issues remain unsolved: (1) When traditional MT-InSAR methods are used, the ground deformation is not well determined due to the low PTs density in the reclaimed land; (2) The deformation associated with land reclamation has not yet been analyzed. So how to select the appropriate number of usable PTs for MT-InSAR deformation monitoring and the analysis of the deformation results needs to be further studied for the reclaimed land of Shenzhen.

In this study, an elaborated PT-based SBAS-InSAR strategy is proposed to select usable PTs and generate an InSAR measurement with full spatial resolution. Based on the selection of PTs, two independent ENVISAT ASAR datasets along the ascending and descending orbits were used to map the coastal land deformation in Shenzhen between 2007 and 2010. Inter-comparison of the InSAR measurements from different orbits was carried out to check the self-consistency of the measurements. Finally, we discussed the implications and potential subsidence hazards caused by land reclamation in this area.

2. Background of Study Area

2.1. Geological Environment

Shenzhen is located at the southeast of the Pearl River Delta (PRD) region in China. The city has complicated topography and is surrounded by the sea in the south and the west. Its elevation ranges between 1 and 10 m in the southwest of Shenzhen, reaches 30 m in the heart of the city, and goes up to 80 m at the northeast (see Figure 1). The marine sediments of the Quaternary period widely spread across the low-lying coastal land areas. This sediment layer is of high water content and high compressibility, big void ratio, and weak shearing strength [35]. The thickness of this layer is generally 3–10 m, but is up to 20 m in some particular areas [40].
Figure 1. Study area and SAR data coverage. The study area is outlined by the black rectangle. Blue rectangles represent the coverage of the SAR dataset used. Red lines denote the fault lines surrounding [41,42]. Inset, the approximate location of Shenzhen in China. The right sub-figure is the distribution of the sediment layer in Shenzhen. The colors indicate the thickness.

2.2. Historical Land Reclamation

Due to its geographical location, Shenzhen is the direct link between the China mainland and Hong Kong. With its economic development and boom [43], the demand for land use [44] is increasing. The comparison of Landsat satellite images (see Figure A1) shows the dramatic landscape changes in Shenzhen from 1979 to 2015. Small hills were leveled, and thousands of high buildings have replaced previous agricultural and vegetated areas. Meanwhile, 80 km² of land has been reclaimed from the sea along the coastal zone in the past 35 years [45]. The temporal and spatial evolution of land reclamation [43] is shown in Figure 2. Land reclamation has brought considerable economic benefits [46] to the government and led to further economic development in Shenzhen. However, long-time compaction of the reclaimed land causes ground deformation [30], which affects the stability of the surface and underground constructions, thereby threatens the safety of people’s lives and assets.

According to an investigation conducted by the Geological Bureau of Shenzhen, Shenzhen is in a stable tectonic environment with low tectonic hazard risk [41]. However, local ground deformation [1,42] related to anthropogenic activity has become the major geohazard in Shenzhen, especially in the coastal reclaimed areas.
Figure 2. Land reclamation and its evolution in Shenzhen. The reclaimed land for every period is spatially overlapped on the Landsat-8 background. Inset shows the temporal evolution. The coastlines are extracted from a set of co-registered Landsat images shown in Figure A1. Firstly, coastlines are automatically extracted by image segmentation with an appropriate threshold. Then the initial coastlines are overlaid to the Landsat image in ArcGIS software—we edited the inaccurate coastline manually.

3. Dataset and Methodology

3.1. Dataset

Two sets of ENVISAT ASAR data acquired between 2007 and 2010, along the ascending (31 scenes, see Table A1 for detail) and descending orbits (28 scenes, see Table A2 for detail), were employed to analyze the coastal deformation. Two dense networks of SBAS interferometric pairs (see Figure 3) with both short perpendicular and temporal baselines were used in the data processing. The available multi-orbit datasets allow us to validate the self-consistency of the monitoring results by inter-comparison.

Figure 3. SBAS networks of the datasets used. (a) Track 025; (b) track 175. Each black dot denotes one image. The line between each two dots represents one interferometric pair, whose color indicates the average coherence of the corresponding interferogram.
3.2. PT-Based SBAS-InSAR Strategy

In our study area, due to the vegetation covering and humid subtropical climate, the application of SAR interferometry is mainly limited by temporal decorrelation [25,47] and inhomogeneity in the tropospheric path delay [27,28]. Besides, the land reclamation related deformation has limited spatial range, thus we do not multi-look the interferograms to retain deformation details. We processed the datasets with an elaborated PT-based SBAS-InSAR, and below (Figure 4) is a flow chart showing the strategy, which consists of four major processing steps.

3.2.1. Pre-Processing

Data processing starts with a stack of Single Look Complex (SLC) images [48]. To minimize the effect of mis-coregistration caused by geometric, temporal, and Doppler decorrelations [25], the images acquired on 28 January 2009 and 17 August 2008 for track 025 and track 175, respectively, were chosen as geometry references for SLC coregistration using the Formula (4.1) of reference [49]. Interferometric pairs with both short perpendicular and temporal baselines were selected to construct networks of the SBAS interferometric pairs. Also, two dense networks of SBAS interferometric pairs (for details see Table 1 and Figure 3) were established. The topographic phase in each interferogram was removed with the 3-arc SRTM DEM [50]. To preserve the detail deformation pattern, the multi-looking process was not applied to any interferogram pairs. A reference point, black star in Figures 5 and 6, was chosen from an inland low-rise building in a stable area.

<table>
<thead>
<tr>
<th>Orbit Type</th>
<th>Track</th>
<th>No. of SLC Images</th>
<th>No. of Pairs</th>
<th>Temporal Span</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ascending</td>
<td>025</td>
<td>31</td>
<td>69</td>
<td>28 February 2007–15 September 2010</td>
</tr>
<tr>
<td>Descending</td>
<td>175</td>
<td>28</td>
<td>65</td>
<td>24 June 2007–22 August 2010</td>
</tr>
</tbody>
</table>

Figure 4. Flow chart of the elaborated PT-based SBAS-InSAR.

Table 1. Summary information of the networks of SBAS interferometric pairs.
Figure 5. Deformation rates generated with track 025 dataset for the whole study area. The black star locates the reference point, and the inset shows a more detail location of it. Black dashed lines represent the Shenzhen Metro Lines. The yellow line denotes the coastline in 1979. The background is a Landsat-8 image acquired in 2015. Regions A, B, C, and D marked with black rectangles will be further analyzed together with land reclamation in the discussion section. Points marked with white cross (×) are selected to show time series deformation.

Figure 6. Deformation rates generated with track 175 dataset for the whole study area. Other remakes are the same as those in Figure 5.
3.2.2. Improved Point Target Selection

Recently, many researches demonstrated that exploiting point-like scatterers [20,21] instead of a two-dimensional grid for deformation monitoring can overcome the effect of interferometric decorrelation [25,51]. The selection of more usable PTs in a low-coherence area [21] for deformation retrieval is a key issue in MT-InSAR data processing. In our study, the coastal reclaimed land may undergo significant deformation and the backscattering coefficients may vary significantly in the time domain. As a consequence, these two factors limit the application of the dispersion of amplitude (DA) method [20] or the distributed scatterers (DS) method [38] to select sufficient usable PTs. Besides, as a phase quality indicator, the coherence maps associated with the interferometric phase could be exploited to identify usable targets. Point target with coherence greater than a given threshold could be selected as a usable PT. However, the coherence is underestimated, as the phase values are non-stationary. Thus non-stationary signals, including DEM, atmospheric effect [26] and possibly the target motion should be estimated and thus be removed. The space averaging of the data inside a certain window (e.g., 50 m × 50 m) is used as the estimation of these non-stationary signals [21] for the center pixel of the window. However, this is not a common case because the average value may not be the genuine estimate [51].

To select adequate usable PTs, an elaborated strategy based on Spectral Diversity of SLC images [48] and the SBAS network [18] is exploited. The concept of spectral diversity originates from Interferometric Point Target Analysis (IPTA) [52], whose principle is that the energy of usable PTs remains almost the same when processing different looks with fractional azimuth and different range bandwidths. We calculated the spectral diversity map $\gamma$ for each co-registered SLC image, and selected a point target as a usable PT, which has the weighted average of the time series spectral diversity values greater than a given threshold (e.g., $\gamma > 0.38$). Given $M$ SLC images, and $N$ interferometric pairs in the SBAS network, then $\gamma$ can be expressed as

$$\bar{\gamma}(r,a) = \frac{1}{2N} \sum_{m=1}^{M} n_m \cdot \gamma_m(r,a)$$

where $\gamma_m(r,a)$ is the spectral diversity value at pixel location $(r,a)$ (range, azimuth) for the $m$-th SLC image, $n_m$ denotes the frequency of the $m$-th SLC image used in the SBAS network, and

$$2N = \sum_{m=1}^{M} n_m$$

Our PT selection strategy is consistent with the one based on the mean interferometric coherence of interferogram stack for SBAS-InSAR. In the coherence-based method, the more frequently a SLC image is used to combine an interferometric pair, the larger the weight factor would be assigned to this image in calculating the mean interferometric coherence for the SBAS network interferogram stack. Similarly, using the frequency of a SLC image as the weight factor in the calculation of mean spectral diversity map $\bar{\gamma}$, would also adjust the influence of that image’s spectral diversity map accordingly. The spectral diversity map could be treated as the Signal-to-Noise Ratio (SNR) indicator of the SLC pixel’s phase. Two pixels with high spectral diversity values would generate high quality interferometric phases, and vice versa. As a consequence, the weighted spectral diversity map $\bar{\gamma}$ is the equivalence of the SNR of the interferometric phases for the given SBAS network, which allows a more accurate PT selection in SBAS-InSAR.

After the PT selection, the data information, including the interferometric phases, spatial coherence etc., is extracted from the two-dimensional grid and stored in vector format for subsequent data processing.

3.2.3. Phase Filtering, Orbital Phase Error Removal and Phase Unwrapping

In order to reduce the phase noise, the differential interferometric phase for PTs, was spatially filtered by an improved SUSAN filter [53,54] with a search window size of 11 × 11 (range × azimuth)
pixels. This method can avoid noise contamination from unusable PTs. The spatial coherence for each PT was estimated with the same window. Subsequently, the smoothed phase of each interferometric pair from the SBAS network was unwrapped by the method based on network programming [55] with the spatial coherence as the weight factor.

In InSAR processing, the flat-earth phase component is removed with precise orbits. For ENVISAT, it has an accuracy of 5–7 cm radially and 10–15 cm cross-track [56], and its residual orbital phase error [57] with long wavelength character may remain in the interferogram. Such error is easy to discriminate from the local deformation in our study area. Hence, we estimated the orbital error using quadratic polynomial models [57] and removed it the error from the unwrapped interferogram. It needs to be noted that the quadratic models would also absorb long-wavelength components of atmospheric effect [27,28].

3.2.4. Modeling of Deformation and Unwrapping Error Checking

The Singular Value Decomposition (SVD) method was employed to solve the rank-deficient problem [18] in SBAS-InSAR deformation modeling. The raw time series deformations that are contaminated by atmospheric delay and phase noise and DEM error correction can be obtained based on the unwrapped phase obtained in Section 3.2.2. The reliability of phase unwrapping was evaluated via the phase difference between the unwrapped phase and the one simulated with raw time series deformation, especially for Qianhai Bay and Shenzhen Bay where sparsely distributed PTs are identified. If the interferometric phase was unwrapped with error, then the phase difference would be significantly large. We then computed the histograms of phase difference for statistics (see Tables A3 and A4). We can see that over 82% of PTs were unwrapped with an error less than 0.5 rad for Qianhai Bay, and 83.3% and 86.4% of the PTs were unwrapped for track 025 and track 175. For these two areas, over 95% of the PTs were unwrapped with error less than 1.0 rad. Additionally, the result showed that over 95% of PTs were unwrapped with error less than 1.0 rad for the whole study area. As a result, we believe that the interferometric phase was unwrapped without significant error.

In order to remove the error of atmospheric delay and phase noise from the raw time series deformations, the linear regression with the raw time series deformations was firstly carried out to inverse the mean deformation velocity map [18]. Then the inverted linear deformation was subsequently subtracted from the raw time series deformations calculated previously. The remaining components are the non-linear ground deformation, atmospheric delay, and random phase noise. Non-linear ground deformations might have some degree of correlation (low-pass) in the temporal domain, while atmospheric delay and phase noise are both high-pass, thus non-linear deformation can be separated by the temporal triangle weighted low-pass filter [20] with a certain length of time window (e.g., 180 days). Finally, we obtained the total deformation by adding up the linear deformation and non-linear deformation.

4. Results

Due to the side-looking geometry, the InSAR measurement is the sum of the projections of ground three-dimensional (3-D) ground deformations in the (Line-of-Sight) LOS direction. In Shenzhen coastal zone, the subsidence of the reclaimed land and sediment layer dominates the ground deformation. Hence, the deformation in our study area was assumed mainly in the vertical direction and the LOS displacement was directly back-projected into the vertical direction using the local incidence angle [36].

4.1. Deformation Rate Maps

The deformation rate maps generated from the ascending and the descending data are between −25 mm/year and 15 mm/year, as shown in Figures 5 and 6, respectively. The deformation patterns are very similar, suggesting that vertical subsidence dominates the coastal deformation. The deformation rate maps demonstrate that most of the inland area within the coastline in 1979 is stable, which means the selected reference point did not undergo obvious deformation. However, the reclaimed land area
along the coastline in 1979 (the yellow line in Figures 5 and 6) experienced significant subsidence, especially at Shenzhen Airport, Bao’an Center, Qianhai Bay, and Shenzhen Bay, marked by black rectangles in Figures 5 and 6. Obvious deformation was not observed in the Chiwan and Shekou harbors, because the foundation of the wharf apron space was enhanced by using special rammer of small base [35]. Besides, we also observed non-uniform deformation in the mean deformation rate maps. The phenomenon is particularly evident in the developed urban region B, Bao’an Center, where many roads and high rise buildings cover the reclaimed land. In addition to the ground subsidence, there are several patches with remarkable uplift located around the right bottom corner in region A. The positive deformation signal may be related to the rise of groundwater level [46]. This area is an old industrial district, where a great amount of groundwater was extracted for production. Now the factories are moving out, decreasing the groundwater extraction [46]. As a result, the groundwater level is recovering.

To validate the non-uniform subsidence, field work was carried out at Bao’an Center (around point TP4) in 2011. Figure 7a,b show remarkable effects of the differential subsidence rates observed at a road and a building. This is because they were built on different foundations. The black arrows in Figure 7 point to the side with solid foundations, which may have slight subsidence. While, red arrows point to the side with unreinforced foundations, where remarkable subsidence occurred.

Figure 7. Field work pictures showing non-uniform subsidence. (a) Site of a road; (b) site of a building.

4.2. Time Series of Selected Points

Several points were selected to show the time series deformation (see Figure 8). The image dataset start time is usually taken as the reference time of the time series in the SBAS-InSAR technique. As seen in Table 1, the two datasets have different start times. For comparison, we set one reference time for these two datasets as 28 February 2007. As a consequence, the time series deformations mapped in the descending orbit was shifted with a constant, which was calculated according to the deformation rate and time difference (116 days). Clearly, in the reclaimed land, TP1, TP2, TP4, PT5, and TP6, are subsiding linearly (see Figure 8). TP3, located inland according to the coastline in 1979, has a different temporal pattern, showing that the TP3 was in rapid deformation (−31.7 mm/year) before January 2009, then its deformation was slowed down (−18.6 mm/year) between January 2009 and January 2010, and finally the area became stable. This deformation is very consistent with the pattern from predictions of the consolidation theory of unsaturated soil [58]. Although the subsidence rate at each selected point is slightly different, the time series obtained from the ascending and descending orbits matched well.
Figure 8. Time series of the deformation at selected points. (a) TP1; (b) TP2; (c) TP3; (d) TP4; (e) TP5; (f) TP6. Purple dots and blue triangles represent the ascending (025) and descending (175) tracks. Purple lines and blue lines represent the linear deformations. The corresponding slopes are given in the lower left corner.

5. Discussion

5.1. Self-Consistency Checking by Inter-Comparison between the Ascending and Descending InSAR Measurements

In order to quantitatively check the self-consistency of InSAR measurements, inter-comparison of the deformation rates between ascending and descending results was carried out. Due to the uncertainty in SAR geo-location [59], the PTs from different orbits were unlikely to be at the exact same locations. The resolution of the results is down-sampled to reduce the effect of the geo-location uncertainty [60]. Firstly, a data grid with a resolution of 50 m × 50 m was created with the tradeoff between the deformation detail and resolutions. Then InSAR measurements from the ascending and descending orbits were resampled respectively to the defined data grid by Pixel Aggregate [61]. Finally, the InSAR measurements from different orbits were compared. We assumed that the InSAR measurements on the defined grid had a common deformation signal. The inter-comparison results between the ascending and descending InSAR-derived deformation rates showed that correlation between the two InSAR measurements is 0.92, revealing that the InSAR measurements mapped by
our PT-based SBAS-InSAR strategy are in good agreement with each other. The Root-Mean-Square Error (RMSE) of the difference between the ascending and descending orbits is 1.8 mm/year, implying that the InSAR measurements are highly self-consistent. However, there are points lying away from the fitted model (the purple line in Figure 9), which may be caused by the errors in InSAR measurements [57], errors in SAR geo-location [59] (mismatch of the measurements from different viewing geometries), horizontal displacement and possibly the non-linear deformation process [20].

Figure 9. Inter-comparison of the deformation rates between track 025 and track 175. Blue dots denote the measured points on track 025 and track 175. The purple line is the fitted linear model, i.e., $f(x) = 0.9245x - 0.07292$.

5.2. Deformation Associated with Land Reclamation

As the first special economic zone of China, Shenzhen has witnessed great changes since the 1980s, accompanied by extensive land reclamation over the past 30 years. The InSAR measurements of regions A, B, C, and D (Figures 5 and 6) are zoomed in Figures 10–13. According to the InSAR measurements, we can observe remarkable deformation in most of the reclaimed land, and the maximum rate is up to 25 mm/year. Generally, ground deformation is induced by primary compaction and secondary compression of the alluvial clay beneath the reclamation [30]. Previous investigations [30,58] showed that 90% primary compaction of alluvial clay in Hong Kong Chek Lap Kok Airport, a site that has very similar land reclamation activity to our study area, took about 10 to 20 years, and longer for a thickness sediment layer. As shown in Figure 2, about 34.3 km$^2$ of land, accounting for 42.6% of the total, was reclaimed from the sea during the past 15 years (period of 2000–2015), on which large infrastructure has been built. Thus, ground settlement is expected in most of the reclaimed area.

Figure 10 shows different subsidence patterns for the new airport built on reclaimed land and the old airport built inland. The new airport was built between years 2000 and 2010 (See Figure 11), thus few PTs were selected from that region. Numerous PTs were identified along the dam to the west part of the new airport. Although the new airport was reclaimed gradually (see the coastline change in Figure 10, the yellow line, yellow dash line, magenta dash line, and the white dash line), remarkable deformation was observed during different phases. For region B (Figure 11), Bao’an center, where many high-rise buildings stand, remarkable subsidence signals in the reclaimed land and along the Metro Lines were detected in both ascending and descending datasets (Figure 11). The Bao’an Center was mainly reclaimed in the period 1979–1990 (the area along the coastlines between 1979 and 1990, see Figure 11). However, it still undergoes remarkable ground deformation, which may be induced
by engineering construction. Slight subsidence was observed on the newly reclaimed land (the area outside the coastline of 1990), as little urban construction is in progress. If a new construction were to be built, the temporary equilibrium could be broken, which may cause additional ground deformation. The Qianhai Bay, an unexploited reclaimed land area, has a few sites with high deformation rates (see Figure 12) caused by compaction of the reclaimed foundation and the alluvial soil beneath. The detail subsidence rate maps of Shenzhen Bay are shown in Figure 13. The results revealed that the body of Shenzhen Bay Check Point is in a weak deformation state \([-6, -3]\) mm/year, while evident settlements were observed in some sites of the reclaimed land, especially the north part of the Check Point. High deformation rates were also detected along the Shenzhen Metro Line (black dash line in Figure 13), which was centered on the reclaimed land. Based on these, we can conclude that remarkable subsidence has occurred on the lands reclaimed 15 years before, such as the Bao’an Center.

**Figure 10.** Subsidence rate of Region A of Figures 5 and 6, Shenzhen Airport. (a) Track 025; (b) track 175. Black dash line with black triangles represents the Shenzhen Metro Line, and the black triangle denotes the Metro station. The yellow line denotes the coastline in 1979. Background is a Landsat-8 image acquired in 2015.

**Figure 11.** Subsidence rate of Region B of Figures 5 and 6, Bao’an center. (a) Track 025; (b) track 175. Other remakes are the same as in Figure 10.

In addition, the engineering constructions, e.g., the construction of Shenzhen Airport and urbanization of Bao’an, would also contribute to subsidence. Considering the deformation maps and
time series at selected points, we believe that the four major land reclamation regions are undergoing remarkable subsidence, which can be expected to continue in the near future. Work to further monitor the ground subsidence shall focus on the coastal land reclamation region, especially Shenzhen Airport, Bao’ an Center, Qianhai Bay, and Shenzhen Bay.

Figure 12. Subsidence rate of Region C of Figures 5 and 6, Qianhai Bay. (a) Track 025; (b) track 175. Other remakes are the same as in Figure 10.

Figure 13. Subsidence rate of Region D of Figures 5 and 6, Shenzhen Bay. (a) Track 025; (b) track 175. Other remakes are the same as in Figure 10.

5.3. Effects of Coastal Subsidence Coupled with Sea Level Changes

The subsidence of the reclaimed land threatens infrastructure, and contributes to the relative sea level changes [62], which may cause saltwater intrusion [63,64] and destroy coastal biodiversity [37]. Assuming a mean sea level rise of 2.5 mm/year [33,65], the selected regions A, B, C, and D (shown
in Figures 10–13) would fall to sea level in 114, 121, 225, and 313 years (see Table 2). Compared with previous studies [38], our deformation rate is slightly larger due to the fact that more PTs have been selected from significant deformation areas in our method. As a consequence, the large deformation rate areas have no measurement of their results whereas appropriate PTs have been identified by our method for deformation mapping. So the difference between their results and ours is expected. The risk assessment based only on sea level rise had been underestimated in the Shenzhen coastal zone because of the amplification effect from coastal subsidence. To decelerate the relative sea level rise, effective measures, such as reinforcing the ground foundation, need to be taken to prevent subsidence in the coastal zone [66].

Table 2. Estimation of the elapsed time of land surface falling sea level regarding coastal subsidence in selected regions.

<table>
<thead>
<tr>
<th>Region</th>
<th>Elevation 1 H (m)</th>
<th>Sea Level Rise ( v_1 ) (mm/year)</th>
<th>Subsidence Rate(^2) ( v_2 ) (mm/year)</th>
<th>Elapsed Time(^3) Estimated ( T ) (year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shenzhen Airport</td>
<td>1.5</td>
<td>2.5</td>
<td>-10.7</td>
<td>114</td>
</tr>
<tr>
<td>Bao’an Center</td>
<td>1.8</td>
<td>2.5</td>
<td>-12.4</td>
<td>121</td>
</tr>
<tr>
<td>Qianhai Bay</td>
<td>1.8</td>
<td>2.5</td>
<td>-5.5</td>
<td>225</td>
</tr>
<tr>
<td>Shenzhen Bay</td>
<td>2.0</td>
<td>2.5</td>
<td>-3.9</td>
<td>313</td>
</tr>
</tbody>
</table>

1 average elevation of region; 2 average subsidence rate of region; 3 \( T = H \times 1000/(v_1 - v_2) \).

5.4. Potential and Limits of InSAR in Mapping Coastal Subsidence

The InSAR measurements from both ascending (track 025) and descending (track 175) orbits showed a very similar pattern in the deformation rate maps (Figures 5, 6 and 10–13) derived by PT-based SBAS-InSAR, and the displacement time-series of the selected points (Figure 8). In order to properly map the coastal deformation associated with land reclamation, the effect of low-coherence will be considered in the implementation of data processing. To illustrate this problem, the dataset from track 175 was also processed with the STAMPS PS method. The comparison (see Appendix B for detail) between STAMPS result and ours showed that few PTs (see Figures B1–B3) have been selected by STAMPS due to long time decorrelation. Although the two results (see Figures B2 and B3) have a similar deformation trend in slight deformation areas, our method has much denser measurements in both significant deformation and stable urban areas, because the PT-based SBAS-InSAR method utilizes multi-reference to overcome the decorrelation effect of single reference interferometric phases, whereas the STAMPS could only select coherent PTs over a long time. As a consequence, such difference is expected. The results show that our method has great potential in mapping coastal deformation. However, our PT-based SBAS-InSAR technique also has its limitations. First, the threshold for PT selection needs to be carefully determined. Second, for a reliable InSAR measurement, a network of dense SBAS interferometric pairs needs to be constructed. Besides, the phase unwrapping error needs to be checked carefully, because such an error would become significant on the data margin for phase unwrapping on the network of irregular PTs.

It is worth remarking that not all coastal zones present subsidence signals in the ENVISAT InSAR measurements. It can be explained by temporal decorrelation [25,47], causing few usable PTs to be available. For example, there are several parts in our study area which tend to decorrelate due to a construction process. Compared with the long revisit cycle (35 days) of ENVISAT, the new generation of SAR platforms have shorter revisits, such as the ALOS-2 (14 days) and Sentinel-1A (12 days). SAR images from these new platforms generally maintain much higher coherence. In addition, only vertical displacement is retrieved, based on the assumption of non-horizontal displacement in this study. However the horizontal displacement cannot be neglected, especially for areas with a large subsidence rate. So combined multiple-pass InSAR datasets with different viewing angles can be considered to reveal ground 3-D deformations further.
6. Conclusions

To meet the demand of urbanization and economic development, about 80 km² of land has been reclaimed from the sea in Shenzhen over the past three decades. The reclaimed land undergoes local heterogeneous deformation that affects the stability of both ground constructions and underground man-made facilities. In order to monitor the reclaimed land deformation, we proposed a PTs-based SBAS-InSAR method, which has the ability to detect large amplitude deformation and choose high PTs density in a low coherence area. The spatial and temporal variability of the deformation was discussed over the lands reclaimed over the past decades in Shenzhen. Inter-comparison of the InSAR measurements from different orbits showed that the correlation between these measurements is 0.92, and the RMSE of the difference between these measurements is 1.8 mm/year, implying a good self-consistency of the InSAR measurements. The InSAR measurements showed that the reclaimed land underwent remarkable deformation (up to 25 mm/year in the LOS direction), especially at the Shenzhen Airport, Bao’an Center, Qianhai Bay, and Shenzhen Bay, and the deformation is expected to continue in the near future. Therefore, continuous deformation monitoring is necessary in these reclaimed area.

Acknowledgments: This work is supported by Shenghua Yuying fund of Central South University and the National Natural Science Foundation of China (41474007 and 41404013) and the Fundamental Research Funds of Central South University (2014zzts049). The Landsat images are provided by USGS. General Mapping Tools (v5.1.2) and Gnuplot (v4.4) are exploited to produce several Figures. The authors would like to thank the four anonymous reviewers for their constructive remarks and suggestions.

Author Contributions: Bing Xu, Guangcai Feng and Zhiwei Li conceive the research work, and Bing Xu wrote the first draft of the paper. Guangcai Feng, Zhiwei Li, Qijie Wang, Changcheng Wang, Rongan Xie contributed to experiment implementation and result interpretation. All authors contributed to paper writing and revision.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASAR</td>
<td>Advanced Synthetic Aperture Radar</td>
</tr>
<tr>
<td>DA</td>
<td>Dispersion of Amplitude</td>
</tr>
<tr>
<td>DEM</td>
<td>Digital Elevation Model</td>
</tr>
<tr>
<td>DS</td>
<td>Distributed Scatterers</td>
</tr>
<tr>
<td>ENVISAT</td>
<td>Environmental Satellite</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>InSAR</td>
<td>Synthetic Aperture Radar Interferometry</td>
</tr>
<tr>
<td>LOS</td>
<td>Line-Of-Sight</td>
</tr>
<tr>
<td>Land</td>
<td>The process of creating new land from ocean,</td>
</tr>
<tr>
<td>Reclamation</td>
<td></td>
</tr>
<tr>
<td>MT-InSAR</td>
<td>Multi-Temporal InSAR</td>
</tr>
<tr>
<td>No.</td>
<td>Number</td>
</tr>
<tr>
<td>PS-InSAR</td>
<td>Persistent Scatterer InSAR</td>
</tr>
<tr>
<td>PT</td>
<td>Point Target</td>
</tr>
<tr>
<td>SBAS</td>
<td>Small Baseline Subset</td>
</tr>
<tr>
<td>SBAS-InSAR</td>
<td>Small Baseline Subset InSAR</td>
</tr>
<tr>
<td>SLC</td>
<td>Single Look Complex</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>SRTM</td>
<td>Shuttle Radar Topography Mission</td>
</tr>
<tr>
<td>TS</td>
<td>Time Series</td>
</tr>
<tr>
<td>3-D</td>
<td>Three-Dimensional</td>
</tr>
</tbody>
</table>
Appendix A

Figure A1. Land reclamation in Shenzhen recorded by the Landsat images. The gray color denotes the urban areas, green represents natural vegetation, and blue represents water. The yellow line is the coastline in 1979. The red line denotes the coastline of the corresponding years.

Table A1. Image datasets of track 025 (ascending).

<table>
<thead>
<tr>
<th>No.</th>
<th>Acquisition Date</th>
<th>Orbit Number</th>
<th>Perpendicular Baseline (m)</th>
<th>Temporal Baseline (day)</th>
<th>Doppler Baseline (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>28 February 2007</td>
<td>26131</td>
<td>1.200</td>
<td>−700</td>
<td>25.565</td>
</tr>
<tr>
<td>2</td>
<td>4 April 2007</td>
<td>26632</td>
<td>−44.517</td>
<td>−665</td>
<td>−95.094</td>
</tr>
<tr>
<td>3</td>
<td>9 May 2007</td>
<td>27133</td>
<td>53.848</td>
<td>−630</td>
<td>24.165</td>
</tr>
<tr>
<td>4</td>
<td>13 June 2007</td>
<td>27634</td>
<td>83.181</td>
<td>−595</td>
<td>−35.858</td>
</tr>
<tr>
<td>5</td>
<td>18 July 2007</td>
<td>28135</td>
<td>249.955</td>
<td>−560</td>
<td>−13.692</td>
</tr>
<tr>
<td>6</td>
<td>22 August 2007</td>
<td>28636</td>
<td>330.813</td>
<td>−525</td>
<td>−36.740</td>
</tr>
<tr>
<td>7</td>
<td>31 October 2007</td>
<td>29638</td>
<td>168.819</td>
<td>−455</td>
<td>10.479</td>
</tr>
<tr>
<td>8</td>
<td>5 December 2007</td>
<td>30139</td>
<td>−198.362</td>
<td>−420</td>
<td>10.734</td>
</tr>
<tr>
<td>9</td>
<td>9 January 2008</td>
<td>30640</td>
<td>248.996</td>
<td>−385</td>
<td>−19.445</td>
</tr>
<tr>
<td>10</td>
<td>19 March 2008</td>
<td>31642</td>
<td>255.402</td>
<td>−315</td>
<td>22.999</td>
</tr>
<tr>
<td>11</td>
<td>23 April 2008</td>
<td>32143</td>
<td>141.737</td>
<td>−280</td>
<td>8.279</td>
</tr>
<tr>
<td>12</td>
<td>2 July 2008</td>
<td>33145</td>
<td>−93.116</td>
<td>−210</td>
<td>−109.998</td>
</tr>
<tr>
<td>13</td>
<td>6 August 2008</td>
<td>33646</td>
<td>216.613</td>
<td>−175</td>
<td>23.050</td>
</tr>
<tr>
<td>14</td>
<td>15 October 2008</td>
<td>34645</td>
<td>45.448</td>
<td>−105</td>
<td>20.116</td>
</tr>
<tr>
<td>15</td>
<td>19 November 2008</td>
<td>35149</td>
<td>−180.867</td>
<td>−70</td>
<td>−93.809</td>
</tr>
<tr>
<td>16</td>
<td>28 January 2009</td>
<td>36151</td>
<td>0.000</td>
<td>0</td>
<td>0.000</td>
</tr>
<tr>
<td>17</td>
<td>4 March 2009</td>
<td>36652</td>
<td>67.309</td>
<td>35</td>
<td>27.292</td>
</tr>
<tr>
<td>18</td>
<td>8 April 2009</td>
<td>37153</td>
<td>−218.997</td>
<td>70</td>
<td>−72.524</td>
</tr>
<tr>
<td>19</td>
<td>13 May 2009</td>
<td>37654</td>
<td>−7.443</td>
<td>105</td>
<td>7.185</td>
</tr>
<tr>
<td>20</td>
<td>17 June 2009</td>
<td>38155</td>
<td>−199.092</td>
<td>140</td>
<td>5.291</td>
</tr>
</tbody>
</table>
Table A1. Cont.

<table>
<thead>
<tr>
<th>No.</th>
<th>Acquisition Date</th>
<th>Orbit Number</th>
<th>Perpendicular Baseline (m)</th>
<th>Temporal Baseline (day)</th>
<th>Doppler Baseline (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>22 July 2009</td>
<td>38656</td>
<td>121.769</td>
<td>175</td>
<td>67.987</td>
</tr>
<tr>
<td>22</td>
<td>26 August 2009</td>
<td>39157</td>
<td>274.131</td>
<td>210</td>
<td>−1.431</td>
</tr>
<tr>
<td>23</td>
<td>30 September 2009</td>
<td>39658</td>
<td>−279.599</td>
<td>245</td>
<td>−92.997</td>
</tr>
<tr>
<td>24</td>
<td>13 January 2010</td>
<td>41161</td>
<td>42.797</td>
<td>350</td>
<td>3.044</td>
</tr>
<tr>
<td>25</td>
<td>17 February 2010</td>
<td>41662</td>
<td>−116.154</td>
<td>385</td>
<td>−77.225</td>
</tr>
<tr>
<td>26</td>
<td>24 March 2010</td>
<td>42163</td>
<td>167.457</td>
<td>420</td>
<td>7.420</td>
</tr>
<tr>
<td>27</td>
<td>28 April 2010</td>
<td>42664</td>
<td>−17.800</td>
<td>455</td>
<td>−89.608</td>
</tr>
<tr>
<td>28</td>
<td>2 June 2010</td>
<td>43165</td>
<td>201.274</td>
<td>490</td>
<td>20.132</td>
</tr>
<tr>
<td>29</td>
<td>7 July 2010</td>
<td>43666</td>
<td>−183.277</td>
<td>525</td>
<td>−36.164</td>
</tr>
<tr>
<td>30</td>
<td>11 August 2010</td>
<td>44167</td>
<td>153.339</td>
<td>560</td>
<td>8.991</td>
</tr>
<tr>
<td>31</td>
<td>15 September 2010</td>
<td>44668</td>
<td>311.510</td>
<td>595</td>
<td>−0.209</td>
</tr>
</tbody>
</table>

Table A2. Image datasets of track 175 (descending).

<table>
<thead>
<tr>
<th>No.</th>
<th>Acquisition Date</th>
<th>Orbit Number</th>
<th>Perpendicular Baseline (m)</th>
<th>Temporal Baseline (day)</th>
<th>Doppler Baseline (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24 June 2007</td>
<td>27784</td>
<td>−54.492</td>
<td>−420</td>
<td>−9.713</td>
</tr>
<tr>
<td>2</td>
<td>29 July 2007</td>
<td>28285</td>
<td>−120.473</td>
<td>−385</td>
<td>−6.183</td>
</tr>
<tr>
<td>3</td>
<td>2 September 2007</td>
<td>28786</td>
<td>64.293</td>
<td>−350</td>
<td>−10.988</td>
</tr>
<tr>
<td>4</td>
<td>7 October 2007</td>
<td>29287</td>
<td>−177.061</td>
<td>−315</td>
<td>−3.644</td>
</tr>
<tr>
<td>5</td>
<td>11 November 2007</td>
<td>29788</td>
<td>87.049</td>
<td>−280</td>
<td>−11.855</td>
</tr>
<tr>
<td>6</td>
<td>16 December 2007</td>
<td>30289</td>
<td>−204.988</td>
<td>−245</td>
<td>−9.418</td>
</tr>
<tr>
<td>7</td>
<td>20 January 2008</td>
<td>30790</td>
<td>−212.565</td>
<td>−210</td>
<td>−12.067</td>
</tr>
<tr>
<td>8</td>
<td>24 February 2008</td>
<td>31291</td>
<td>−282.373</td>
<td>−175</td>
<td>−6.489</td>
</tr>
<tr>
<td>9</td>
<td>30 March 2008</td>
<td>31792</td>
<td>137.053</td>
<td>−140</td>
<td>−7.713</td>
</tr>
<tr>
<td>10</td>
<td>4 May 2008</td>
<td>32293</td>
<td>−85.548</td>
<td>−105</td>
<td>−14.254</td>
</tr>
<tr>
<td>11</td>
<td>8 June 2008</td>
<td>32794</td>
<td>66.104</td>
<td>−70</td>
<td>−8.627</td>
</tr>
<tr>
<td>12</td>
<td>13 July 2008</td>
<td>33295</td>
<td>174.508</td>
<td>−35</td>
<td>−3.449</td>
</tr>
<tr>
<td>13</td>
<td>17 August 2008</td>
<td>33796</td>
<td>0.000</td>
<td>0</td>
<td>0.000</td>
</tr>
<tr>
<td>14</td>
<td>21 September 2008</td>
<td>34297</td>
<td>−184.108</td>
<td>35</td>
<td>2.401</td>
</tr>
<tr>
<td>15</td>
<td>4 January 2009</td>
<td>35800</td>
<td>36.880</td>
<td>140</td>
<td>−7.600</td>
</tr>
<tr>
<td>16</td>
<td>8 February 2009</td>
<td>36301</td>
<td>−227.607</td>
<td>175</td>
<td>−7.695</td>
</tr>
<tr>
<td>17</td>
<td>15 March 2009</td>
<td>36802</td>
<td>395.745</td>
<td>210</td>
<td>−2.732</td>
</tr>
<tr>
<td>18</td>
<td>19 April 2009</td>
<td>37303</td>
<td>−159.084</td>
<td>245</td>
<td>−16.068</td>
</tr>
<tr>
<td>19</td>
<td>24 May 2009</td>
<td>37804</td>
<td>21.721</td>
<td>280</td>
<td>−17.713</td>
</tr>
<tr>
<td>20</td>
<td>28 June 2009</td>
<td>38305</td>
<td>156.203</td>
<td>315</td>
<td>−9.648</td>
</tr>
<tr>
<td>21</td>
<td>2 August 2009</td>
<td>38806</td>
<td>−116.555</td>
<td>350</td>
<td>−9.147</td>
</tr>
<tr>
<td>22</td>
<td>6 September 2009</td>
<td>39307</td>
<td>133.858</td>
<td>385</td>
<td>−8.236</td>
</tr>
<tr>
<td>23</td>
<td>20 December 2009</td>
<td>40810</td>
<td>−153.050</td>
<td>490</td>
<td>−11.066</td>
</tr>
<tr>
<td>24</td>
<td>28 February 2010</td>
<td>41812</td>
<td>−269.032</td>
<td>560</td>
<td>−1.572</td>
</tr>
<tr>
<td>25</td>
<td>4 April 2010</td>
<td>42313</td>
<td>232.788</td>
<td>595</td>
<td>−21.591</td>
</tr>
<tr>
<td>26</td>
<td>9 May 2010</td>
<td>42814</td>
<td>171.201</td>
<td>630</td>
<td>−35.714</td>
</tr>
<tr>
<td>27</td>
<td>13 June 2010</td>
<td>43315</td>
<td>143.627</td>
<td>665</td>
<td>−15.045</td>
</tr>
<tr>
<td>28</td>
<td>22 August 2010</td>
<td>44317</td>
<td>−191.774</td>
<td>735</td>
<td>−7.937</td>
</tr>
</tbody>
</table>

Table A3. Statistics of the phase unwrapping error in track 025.

<table>
<thead>
<tr>
<th>Phase Unwrapping Error (rad)</th>
<th>Qianhai Bay Percentage (%)</th>
<th>Shenzhen Bay Percentage (%)</th>
<th>Whole Area Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[−0.5, 0.5]</td>
<td>82.70</td>
<td>83.33</td>
<td>85.12</td>
</tr>
<tr>
<td>[−1.0, 1.0]</td>
<td>95.35</td>
<td>96.06</td>
<td>96.53</td>
</tr>
<tr>
<td>[−1.5, 1.5]</td>
<td>98.98</td>
<td>99.14</td>
<td>99.26</td>
</tr>
</tbody>
</table>
Table A4. Statistics of the phase unwrapping error in track 175.

<table>
<thead>
<tr>
<th>Phase Unwrapping Error (rad)</th>
<th>Qianhai Bay</th>
<th>Shenzhen Bay</th>
<th>Whole Area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percentage (%)</td>
<td>Percentage (%)</td>
<td>Percentage (%)</td>
</tr>
<tr>
<td>[−0.5, 0.5]</td>
<td>82.81</td>
<td>86.43</td>
<td>87.15</td>
</tr>
<tr>
<td>[−1.0, 1.0]</td>
<td>96.25</td>
<td>97.52</td>
<td>97.65</td>
</tr>
<tr>
<td>[−1.5, 1.5]</td>
<td>99.46</td>
<td>99.58</td>
<td>99.63</td>
</tr>
</tbody>
</table>

Appendix B

To evaluate the effectiveness of the proposed strategy, the dataset from track 175 is processed with the STAMPS PSI method [21,49] (see Figure B1). Numerous PTs in built-up areas have been detected by the STAMPS method and the proposed method. However, in low-coherence areas, the STAMPS method can select few PTs due to the decorrelation effect, while our strategy can select an appropriate number of PTs. The density of PTs selected by STAMPS is 301 points/km², however, it is 779 points/km² by our method. Such difference is particularly evident in reclaimed land areas, such as Bao’an Center, the north to the Shenzhen Bay Checkpoint, and the dam of Shenzhen Airport.

For more detailed comparison, the regions A and B marked with a black rectangle (see Figure B1) were further analyzed with zoom-in view shown in Figures B2 and B3. The deformation rate maps obtained by STAMPS and our method have a very similar deformation trend (see Figures B2 and B3). However, our deformation rate can detect much larger deformation than that of the STAMPS results. As shown in Figures B2 and B3, few PTs have been selected from areas with large deformations by the STAMPS method. As a consequence, the area that undergoes rapid deformation has few measurements on STAMPS results, while appropriate PTs have been identified by our method for deformation mapping.

Figure B1. Deformation rate maps extracted by (a) PT-based SBAS-InSAR and (b) STAMPS. We can see that more PTs have been selected by our method (PT-based SBAS-InSAR). Regions A and B are further analyzed.
Figure B2. Deformation rate maps for region A, Bao’an Center extracted by (a) PT-based SBAS-InSAR and (b) STAMPS. Numerous PTs have been selected in the urban area, upper-right part of the Figure. However, the density of PTs selected by our method is larger than that of STAMPS, especially in areas with large deformation. However, in (b) the significant deformation could not be well mapped, due to the sparse PTs density of the STAMPS method.

Figure B3. Deformation rate maps for region B, Shenzhen Bay extracted by (a) PT-based SBAS-InSAR and (b) STAMPS. Numerous PTs have been selected in the urban area. However, our method can select more PTs than the STAMPS method in areas with a large deformation rate, especially the northern part of the Shenzhen Bay Check Point, which enables us to map the significant deformation.

References


45. Li, X.; Damen, M.C. Coastline change detection with satellite remote sensing for environmental management of the Pearl River Estuary, China. *J. Mar. Syst.* 2010, 82, S54–S61. [CrossRef]


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
Investigation on Mining Subsidence Based on Multi-Temporal InSAR and Time-Series Analysis of the Small Baseline Subset—Case Study of Working Faces 22201-1/2 in Bu’ertai Mine, Shendong Coalfield, China

Chao Ma 1,2, Xiaoqian Cheng 1,2,*, Yali Yang 1, Xiaoke Zhang 3, Zengzhang Guo 2 and Youfeng Zou 2

1 Department of Remote Sensing Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China; mac@hpu.edu.cn (C.M.); d_insar@126.com (Y.Y.)
2 Key Laboratory of Mine Spatial Information Technologies of SBSM, Henan Polytechnic University, Jiaozuo 454000, China; gzc@hpu.edu.cn (Z.G.); zouyf@hpu.edu.cn (Y.Z.)
3 School of Public Administration, Hohai University, Nanjing 210098, China; zhang_xiao_ke@126.com
* Correspondence: cxq@hpu.edu.cn; Tel.: +86-391-398-7661

Academic Editors: Zhenhong Li, Roberto Tomas, Richard Gloaguen and Prasad S. Thenkabail
Received: 9 June 2016; Accepted: 8 November 2016; Published: 16 November 2016

Abstract: High-intensity coal mining (large mining height, shallow mining depth, and rapid advancing) frequently causes large-scale ground damage within a short period of time. Understanding mining subsidence under high-intensity mining can provide a basis for mining-induced damage assessment, land remediation in a subsidence area, and ecological reconstruction in vulnerable ecological regions in Western China. In this study, the mining subsidence status of Shendong Coalfield was investigated and analyzed using two-pass differential interferometric synthetic aperture radar (DInSAR) technology based on high-resolution synthetic aperture radar data (RADARSAT-2 precise orbit, multilook fine, 5 m) collected from 20 January 2012 to June 2013. Surface damages in Shendong Coalfield over a period of 504 days under open-pit mining and underground mining were observed. Ground deformation of the high-intensity mining working faces 22201-1/2 in Bu’ertai Mine, Shendong Coalfield was monitored using small baseline subset (SBAS) InSAR technology. (1) DInSAR detected and located 85 ground deformation areas (including ground deformations associated with past-mining activity). The extent of subsidence in Shendong Coalfield presented a progressive increase at an average monthly rate of 13.09 km² from the initial 54.98 km² to 225.20 km², approximately, which accounted for 7% of the total area of Shendong Coalfield; (2) SBAS-InSAR reported that the maximum cumulative subsidence area reached 5.58 km² above the working faces 22201-1/2. The advance speed of ground destruction (7.9 m/day) was nearly equal to that of underground mining (8.1 m/day).

Keywords: multi-temporal InSAR; small baseline subset (SBAS); deformation monitoring; damage assessment; mining subsidence; Shendong Coalfield; Bu’ertai Mine

1. Introduction

Shendong Coalfield is one of the largest coal production bases in the world. The high-density fields and high-intensity mining in Shendong Coalfield will profoundly influence the ecological environment of northwestern China. Therefore, assessing mining-induced damage and understanding mining subsidence in the high-intensity mining environment of Shendong Coalfield are important for the land remediation and ecological reconstruction of mining subsidence areas. Traditional mining subsidence
monitoring approaches mainly include geodesy, global navigation satellite system measurements, and electronic distance measurements. These approaches exhibit the following shortcomings. (1) They are labor-intensive, time-consuming, costly, and difficult to keep the monitoring flags in good condition for a long time; (2) The survey crew has to enter the region being monitored, which increases the difficulty and risks of the task; (3) These approaches are inapplicable to fast and accurate large-scale real-time monitoring of mines and cannot monitor subsidence in unknown regions because of the limited spatial extent of monitoring, the low spatial resolution, and the long work cycle; (4) Theoretical analysis is restricted within-data observation of discrete points; hence, identifying the ground deformation characteristics while satisfying the actual requirements of mining subsidence prediction and disaster prevention is difficult [1,2].

Interferometric synthetic aperture radar (InSAR) offers a novel Earth observation approach and can provide all-weather, all-day, and cloud influence free monitoring. Differential InSAR (DInSAR), as an extension of InSAR in terms of monitoring ground deformation, is mainly used to capture centimeter-level or smaller ground deformations along the line of sight (LOS) of a radar satellite. Considerable developments related to DInSAR monitoring of earthquake deformation [3,4], volcanic activities [5,6], glacial shift [7,8], urban water-loss settlement [9,10], mining subsidence [11,12], and landslides [13,14] have been achieved. However, DInSAR technology can easily cause interference decorrelation for mining subsidence with immense deformation and short deformation period. Moreover, for the conventional DInSAR, it is a challenging task to distinguish, and thus, effectively eliminate orbit residue error, residual terrain, atmospheric delay, and other phase errors. Nevertheless, several successful applications of the conventional DInSAR have been reported. In the past 10 years, many scholars have monitored mining subsidence in selected mines using DInSAR and have conducted corresponding experimental research. Perski et al. successfully measured ground settlement caused by underground mining using DInSAR [15–20]. Other scholars have successfully applied DInSAR to mining subsidence monitoring experiments, thereby confirming the feasibility of using DInSAR in large-scale mining subsidence monitoring and geological disaster assessment [21–28]. Therefore, the conventional DInSAR coupled with an appropriate interferometric strategy can recognize ground deformation associated with past-mining activities and atmospheric effects (e.g., water vapor in the troposphere and fluctuations in the ionosphere). It can also reduce data costs and increase observation efficiency [29].

To overcome the shortcomings of DInSAR and to acquire accurate subsidence data, the deformation velocity of highly coherent target points was computed via least squares (LS) estimation [30]. In 2002, Berardino et al. proposed the small baseline subset (SBAS) algorithm based on the LS model [31–33]. SBAS was developed from the traditional DInSAR. It combined multiple small baseline subsets via the singular value decomposition (SVD) method based on InSAR data pairs with small spatial and temporal baselines. Therefore, SBAS does not only inherit the advantages of the traditional DInSAR, but can also effectively solve the time discontinuous problem caused by extremely long spatial baselines among SAR data sets, thereby obtaining subsidence values with high coherence targets during different periods. SBAS becomes increasingly optimized and its monitoring accuracy increases accordingly with the development of three-dimensional (3D) phase unwrapping algorithms, extraction algorithms of coherent target points, and error elimination algorithms, as well as the improvement of the time-series deformation model [34–40]. Nevertheless, it still has low computational efficiency. Moreover, precise observation is limited within a small scale, and thus, a substantial part of mining subsidence (i.e., maximum subsidence) cannot be covered.

Therefore, given the limited data set, macroscopic and microscopic studies on global and local ground displacement features caused by mining subsidence were performed by combining conventional and advanced interferometric techniques (DInSAR and SBAS-InSAR). On the one hand, the large-scale dynamic characteristics of mining subsidence were identified through a comparative analysis of multi-temporal data by combining two DInSAR interferences. On the other hand, a time-series analysis of a typical working face was conducted based on SBAS-InSAR and several
ground displacement parameters were determined. These processes not only provide complete control of the maturity and stability of DInSAR to perform a fast assessment of mining-induced damage in the entire coalfield, but also fully utilize the technical advantages of SBAS-InSAR in increasing the time sampling rate as well as in inhibiting terrain and atmospheric delay influences on the algorithm. Consequently, an accurate analysis of the subsidence feature of a typical working face is realized.

2. Site Selection and Methods

2.1. Site Selection

2.1.1. Study Area 1: Shendong Coalfield

Shendong Coalfield, an immense coal production base in China, is located in the northern region of Shenmu County, the western region of Fugu County (Yunlin City, Shaanxi Province) and Ejin Horo County, and the southern region of Dongsheng District and Junggar County (Ordos City, Inner Mongolia). Its geographical coordinates are 38°52′–39°41′N and 109°51′–110°46′E. Its north–south length is approximately 38–90 km and its east–west width is approximately 35–55 km, thereby covering an area of approximately 3481 km². The proven reserves are 727 Gt. Shendong Coalfield covers a cluster of the mines of 10-million-tons annual production that comprise 14 underground coal mines, including Bu’ertai (20.0 Mt/year), Daliuta (25.0 Mt/year), Bulianta (21.0 Mt/year), Halagou (12.0 Mt/year), Shangwan (10.0 Mt/year), and Shigetai (10.0 Mt/year) mine, as well as 5 open-pit mines, including Ha’erwusu (30.0 Mt/year) and Heidaigou (31.0 Mt/year) mine. Shendong Coalfield is one of the largest combined open-pit–underground mining enterprises with the highest degree of modernization in the world (Figure 1). Shendong coalfield is suffering from a large-scale, centralized distribution of mining subsidence and ground damage because it is being large-scale exploitation, the main mineable coal beds is particularly thick, and the ratio of mining depth/mining thickness is smaller. (Figure 2).

Figure 1. Geographical location of Shendong Coalfield. Background data: Satellite imagery acquired by Landsat 7 on 10 April 2013.
2.1.2. Study Area 2: 22201-1/2 Working Face

The 22201-1/2 long wall belongs to Bu’ertai Mine and is composed of two columns of working faces. 22201-1 was exploited in 2012, whereas 22201-2 was exploited in 2013 (Figure 3). The mining width of the working face that corresponded to the subsidence area was about 300 m. The mining length from 20 January 2012 to June 2013 was about 3950 m. The 22# coal bed was exploited. The mining seam thickness was 2.2–2.9 m (2.5 m on average), and the average mining depth was about 260 m [41].

![Figure 3. Mining layout plan of 22201-1/2 long wall.](image)

In Figure 2, the mining boundaries (the purple vertical line) of the working face, while the RADARSAT-2 were imaging, were calculated according to the monthly stopping lines (the green vertical line marked by the arrow symbol) of the coal excavation plans.

2.1.3. Data used

RADARSAT-2 data from 20 January 2012 to June 2013 were used, and the data revisit period was set to 24 day. The spatial resolution was 5 m. Moreover, this study adopted the C wave band as the working wave band, HH polarization mode, multilook fine beam, and 50 × 50 km² coverage area. The digital elevation model (DEM) data required for processing were 90 m elevation data of the Shuttle Radar Topography Mission (SRTM DEM). Images that covered the entire Bu’ertai Mine (within 39°24′–39°26′N and 109°58′–110°1′E) were collected by subset 18 SAR scenes via region
of interests (ROIs) Data from September 2012, 20 January13, and May 2013 were missing, but the remaining revisit period data were all collected (Table 1).

Table 1. Imaging parameters and geometry of RADARSAT-2 MF6 data used in this work.

<table>
<thead>
<tr>
<th>No.</th>
<th>Orbit</th>
<th>Frame</th>
<th>Acquisition Date (yyyymmdd)</th>
<th>Cumulative Pattern</th>
<th>Consecutive Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Time Baseline (Days)</td>
<td>Perp. 1 Baseline (m)</td>
</tr>
<tr>
<td>1</td>
<td>53987</td>
<td>513409</td>
<td>20120120</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>53987</td>
<td>513410</td>
<td>20120213</td>
<td>24</td>
<td>406.3809</td>
</tr>
<tr>
<td>3</td>
<td>53987</td>
<td>513411</td>
<td>20120308</td>
<td>48</td>
<td>483.7567</td>
</tr>
<tr>
<td>4</td>
<td>53987</td>
<td>513412</td>
<td>20120401</td>
<td>72</td>
<td>566.6843</td>
</tr>
<tr>
<td>5</td>
<td>53987</td>
<td>513413</td>
<td>20120425</td>
<td>96</td>
<td>431.9236</td>
</tr>
<tr>
<td>6</td>
<td>53987</td>
<td>513414</td>
<td>20120519</td>
<td>120</td>
<td>169.2050</td>
</tr>
<tr>
<td>7</td>
<td>53987</td>
<td>513415</td>
<td>20120612</td>
<td>144</td>
<td>399.0324</td>
</tr>
<tr>
<td>8</td>
<td>53987</td>
<td>513416</td>
<td>20120706</td>
<td>168</td>
<td>518.7924</td>
</tr>
<tr>
<td>9</td>
<td>53987</td>
<td>513417</td>
<td>20120730</td>
<td>192</td>
<td>480.6470</td>
</tr>
<tr>
<td>10</td>
<td>53987</td>
<td>513418</td>
<td>20120823</td>
<td>216</td>
<td>216.5596</td>
</tr>
<tr>
<td>11</td>
<td>53987</td>
<td>513419</td>
<td>20121010</td>
<td>264</td>
<td>242.0272</td>
</tr>
<tr>
<td>12</td>
<td>53987</td>
<td>513420</td>
<td>20121127</td>
<td>312</td>
<td>461.3537</td>
</tr>
<tr>
<td>13</td>
<td>53987</td>
<td>513421</td>
<td>20121221</td>
<td>336</td>
<td>470.9812</td>
</tr>
<tr>
<td>14</td>
<td>45789</td>
<td>439762</td>
<td>20130207</td>
<td>384</td>
<td>94.9292</td>
</tr>
<tr>
<td>15</td>
<td>45789</td>
<td>439765</td>
<td>20130303</td>
<td>408</td>
<td>226.1092</td>
</tr>
<tr>
<td>16</td>
<td>57396</td>
<td>539004</td>
<td>20130327</td>
<td>432</td>
<td>463.3215</td>
</tr>
<tr>
<td>17</td>
<td>57396</td>
<td>539005</td>
<td>20130420</td>
<td>456</td>
<td>489.2963</td>
</tr>
<tr>
<td>18</td>
<td>57396</td>
<td>539006</td>
<td>20130607</td>
<td>504</td>
<td>593.3395</td>
</tr>
</tbody>
</table>

1 An abbreviation of the perpendicular component of the spatial baseline.

2.2. Methods

In view of the lack of field data, we will adopt a joint calculation and mutual authentication method, such as consecutive DInSAR, cumulative DInSAR and SBAS-InSAR, in order to achieve a high reliability by the way.

2.2.1. Multi-Temporal InSAR Technology

The interferometric principle of mining subsidence based on DInSAR has been discussed in many studies [11,12,15–29], and will not be duplicated in this work. Instead, multi-temporal InSAR technology was applied and two interferometric processing techniques were used for the entire coalfield: the consecutive (i.e., adjacent acquisitions) DInSAR interferometry and the cumulative DInSAR interferometry. These techniques require various physical values and mathematical methods. They use different temporal and spatial baseline combinations, and thus, the corresponding results provide varying geological significance. The consecutive DInSAR interferometry, which is defined that the two adjacent scenes of time series of SAR data (their temporal baseline is the shortest) carry out an interferometry one another and eventually achieve a complete time series interferometric results (e.g., $\phi_{1,2}$, $\phi_{2,3}$, ..., $\phi_{n-1,n}$), is convenient for the quantitative analysis of interferometric phase changes and can obtain a subsidence extent within a single period. The cumulative DInSAR interferometry, which is defined that the master image is fixed the first scene of time series of SAR data, meanwhile the next scene is selected (their temporal baseline is gradually lengthened) to implement interferometry with it and eventually achieve a complete time series interferograms (e.g., $\phi_{1,2} + \phi_{2,3} + ... + \phi_{n-1,n}$), is convenient for identifying the phase change features in a mining area and can obtain the maximum subsidence magnitude. This technique is conducive to evaluating the variation characteristics for mining subsidence on a deformation field from different spatio-temporal perspectives.
2.2.2. Brief Description of SBAS-InSAR

To improve the monitoring accuracy of the InSAR technique, one of the advanced InSAR techniques termed short baseline subsets (SBAS) InSAR was developed and has been frequently applied since 2002 [31–33,42,43]. This algorithm uses interferograms with small baselines that overlap in time in order to reduce spatial decorrelation, and to mitigate atmospheric artifacts and topographic errors in time-sequential interferometric pairs. The differential phase for a generic coherent pixel of the range and azimuth coordinates \((x, r)\) in interferogram \(j\) that is generated by combining SAR acquisitions at times \(t_B\) and \(t_A\) is

\[
\delta \phi_j(x, r) = \phi(t_B, x, r) - \phi(t_A, x, r) \approx \frac{4\pi}{\lambda} [d(t_B, x, r) - d(t_A, x, r)] + \frac{4\pi}{\lambda} \frac{B_{ij} \Delta z}{R \sin \theta} + [\phi_{atm}(t_B, x, r) - \phi_{atm}(t_A, x, r)] + \Delta n_j, \forall j = 1, \ldots, M
\]

(1)

where \(\lambda\) is the transmitted radar wavelength; \(\phi(t_B, x, r)\) and \(\phi(t_A, x, r)\) are the phases that correspond to times \(t_B\) and \(t_A\); and \(d(t_B, x, r)\) and \(d(t_A, x, r)\) are the radar LOS projection of the cumulative deformation referenced to the first scene, which implies \(\phi(t_0, x, r) = 0\). We also include a phase term related to possible errors \(\Delta z\) in the applied DEM used to generate differential interferograms. This phase component is proportional to the perpendicular baseline for each interferogram \(B_{ij}\), range distance \(R\), and beam incident angle \(\theta\). A possible atmospheric signal is included in the terms \(\phi_{atm}(t_B, x, r)\) and \(\phi_{atm}(t_A, x, r)\). Decorrelation effects and other noise sources are included in the last term \(\Delta n_j\).

Therefore, Equation (1) can be expressed as follows:

\[
\delta \phi_j(x, r) = \delta \phi_{disp}^j(x, r) + \delta \phi_{topo}^j(x, r) + \delta \phi_{atm}^j(x, r, \forall j = 1, \ldots, M
\]

(2)

where \(\delta \phi_{disp}^j\) is the slant deformation phase between \(t_B\) and \(t_A\); \(\delta \phi_{topo}^j\) represents the topographic phase error of the external DEM used for differential interferogram generation; \(\delta \phi_{atm}^j\) accounts for temporal atmospheric variation at different SAR acquisition \(t_B\) and \(t_A\); and \(\delta \phi_{noise}^j\) denotes temporal decorrelation, orbital errors, thermal noise effects, etc.

The cumulative deformation between \(t_B\) and \(t_A\) can also be expressed as follows:

\[
\delta \phi_{disp}^j(x, r) = \frac{4\pi}{\lambda} \sum_{k=A}^{B-1} v_{k-1, k+1} (t_{k+1} - t_k)
\]

(3)

where \(k\) indicates the sequence index of SAR acquisition time between \(t_B\) and \(t_A\), and \(v\) represents the mean phase velocity in the period from \(k\) to \(k+1\).

Given \(M\) unwrapped interferograms, the cumulative deformation at different SAR acquisition times can be achieved using the LS or SVD methods.

3. Results

3.1. DInSAR Data Processing

3.1.1. Spatio-Temporal Baselines of Two Interferometric Strategies

The temporal baseline of the consecutive DInSAR interferometry can be controlled efficiently within 24–48 days (Figure 4a), which will produce an ideal interferometric effect. However, the temporal baseline of the cumulative DInSAR interferometry increased continuously from the initial 24 days to the final 504 days (Figure 4b), which indicated that the influence of temporal decorrelation increasingly intensified.
3.1.2. Results of Consecutive and Cumulative DInSAR Interferometries

(1) For the consecutive DInSAR interferometry, no image is fixed to a master image. The advanced 17 images were in turn used as the master image, whereas the subsequent image was used as the slave image. The temporal baseline was the same with the revisit period. The interferogram was characterized by independent differential phase changes in various periods and a forward movement of the footprint of the deformation phase of mining subsidence (Figure 5).

Figure 4. Distribution of the spatiotemporal baseline. (a) Spatiotemporal baseline distribution of consecutive DInSAR; (b) Spatiotemporal baseline distribution of cumulative DInSAR.

Figure 5. Cont.
Figure 5. The results of the consecutive DInSAR pattern (January 2012–June 2013).

An ionospheric disturbance occurred during the interference process. Large reverse phase regions were observed from two interferometric pairs, namely, 20120425–20120519 and 20120519–20120612, which indicated that an atmospheric effect was present during the imaging of 20120519. The large reverse phase of 20120612–20120706 and 20120706–20120730 implied the presence of an atmospheric effect during the imaging of 20120706.
(2) For the cumulative DInSAR interferometry, the first image was set as the master image (Master, 20 January 2012) and the subsequent 17 images were alternately used as slave images. The interferogram indicated that the differential phase changes covered all previous phase changes, the deformation phase center of mining subsidence moved forward, and the deformation region expanded monthly (Figure 6).

Figure 6. Cont.
The cumulative DInSAR interferometry was influenced simultaneously by temporal and spatial decorrelations. The interferogram had many unstable patches and several phase residues. It failed in retaining coherence and phase unwrapping during the last three periods.

3.2. SBAS-InSAR Data Processing

To eliminate spatiotemporal decorrelation and atmospheric effects, an experimental plot was selected to test SBAS-InSAR, which was expected to reduce or eliminate the influences of decorrelation factors during DInSAR data processing by selecting appropriate spatiotemporal baseline thresholds.

3.2.1. Spatiotemporal Baseline Optimization of SBAS

All interferometric pairs within the thresholds of the spatial and temporal baselines were screened based on a preset spatial baseline threshold, a temporal baseline threshold, and relevant input parameters, thereby generating the SAR data pair connection diagram (Figure 7). In this experiment, the maximum critical baseline percentage was set at 5%. The maximum perpendicular component of the spatial baseline was 422 m. The maximum temporal baseline was set 200 days. The system automatically screened the super master image (20120823). Interferometric pairs with low coherence and poor unwrapping were eliminated, and finally, 47 interferometric pairs were selected.

Figure 6. The results of the cumulative DInSAR pattern (January 2012–June 2013).
Figure 7. Spatiotemporal baseline distribution of the connection diagram and unwrapping results: (a) Spatiotemporal baseline distribution of the connection diagram; (b) 3D unwrapping sub-pairs.

3.2.2. Interferometry Results of SBAS

Considering significant computational efforts and necessary disk space, SBAS was applied finely in a single working face of the 22201-1/2 long wall. Interferometry was performed on each image pair according to their connection relationship based on the aforementioned optimized baselines of SBAS, and 47 interferograms were obtained. A total of 17 time-series cumulative phase deformation diagrams were collected through orbit refining, relatting, phase unwrapping, and geocoding. Subsequently, a −10 mm subsidence contour map was acquired based on the further processing of these diagrams (Figure 8a–q).

Figure 8. Cont.
Figure 8. Cont.
4. Discussions

4.1. The Reliability Validation of the Interferometry

Because of the past mining subsidence and lack of field data, the reliability of interferometric measurements needs to be verified and validated. We randomly selected three subsidence basins (basin A, B, C) and some understanding are gotten by comparing the results of the two interferometry measurements (i.e., SABS-InSAR and the consecutive DInSAR).

See Figures 9–11, using the same profile line, we extract the subsidence values in the two subsidence basins by SABS-InSAR and the consecutive DInSAR, and then perform the overlay analysis. The unwrapping of the central area of the subsidence basin is erroneous due to the influence of decoherence. So we only compare the information at the edge of the subsidence basins. We found that the results of two interferometric measurements coincide each other fairly well. For example, the Root
Mean Square Error (RMSE) of the subsidence values on both sides of basin A are ±0.6 mm (left) and ±0.3 mm (right) respectively (see Figure 9), the RMSE of the subsidence values on both sides of basin B are ±0.6 mm (left) and ±0.2 mm (right) respectively (see Figure 10), and the RMSE of the subsidence values on both sides of basin C are ±0.4 mm (left) and ±0.1 mm (right) respectively (see Figure 11). Therefore, through the above cross-validation, we believe that our data processing scheme is highly reliable, and it can be used for large-scale investigation of past mining subsidence.

Figure 9. Example I: verification and validation of Interferometric Measurement.
Figure 10. Example II: verification and validation of Interferometric Measurement.

Figure 11. Cont.
4.2. Mining Subsidence Status in Shendong Coalfield

4.2.1. Interpretation based on the consecutive DInSAR interferometry

The results of the 17-temporal the consecutive DInSAR interferometry were interpreted using the subsidence value of $-10$ mm recommended by a literature as the outer boundary of mining subsidence [44]. Polygons with low speckle in the interpreted subsidence were further screened. Some polygons, which did not change during three revisit periods (According [44], the zone, where mining subsidence is less than $-50$ mm during 3 months, is considered to be the residual deformation zone.), will be removed, thereby obtaining the time-series subsidence map of the consecutive DInSAR interferometry (Figure 12).

The interpretation results indicated that Huhehewusu Mine, Erlintu Mine, Zhugaita Mine, and Dahaize Mine in Shendong Coalfield had been hardly any exploited, whereas the remaining mines produced an average of $45 \pm 3$ dynamic subsidence regions during each observation period from 20 January 2012 to 7 June 2013. The average mining extent during each observation period reached as high as $57.00 \pm 0.24$ km$^2$, which indicated the active production activities in the coalfield. Approximately 45 mining areas were exploited simultaneously, which caused serious mining subsidence. Interpretation was a difficult task accomplished through team cooperation. The concerned subsidence quantity and area would differ to a certain extent given that the team members had varying work experiences.
4.2.2. Interpretation Based on the cumulative DInSAR Interferometry

Similarly, the results of the 14-temporal the cumulative DInSAR interferometry were interpreted using the subsidence value of $-10$ mm recommended by [44] as the outer boundary of mining subsidence. The polygons with low speckle in the interpreted subsidence were further screened, and those that did not change successively will be removed. Finally, the time-series subsidence map was obtained (Figure 13).

![Figure 12. Interpretation results of the consecutive interferometry (January 2012–June 2013).](image1)

![Figure 13. Interpretation results of the cumulative interferometry (January 2012–June 2013).](image2)

Interpretation became more difficult given the influence of temporal decorrelation, and the interpretable degree was lower than that of the consecutive DInSAR interferometry. Evidently,
the characteristic of the time series of the cumulative DInSAR interferometry was that polygons with low speckle in the subsidence range expanded to a fixed direction. An average of 46 ± 5 subsidence areas expanded stably during each study period from 20 January, 2012 to 21 December 2012. The subsidence area increased from an initial value of 54.98 km²/24 day to 225.20 km²/504 day throughout the observation period (20 January 2012 to June 2013), which indicated an average rate of 13.09 km².

In general, it is still difficult to obtain a continuous series of subsidence areas. On the one hand, mine production regarded mine as a unit, which was not uniform. Several subsidence areas were caused by ground deformation associated with post-mining activities, others resulted from the preparation of alternate long walls, whereas some areas were produced by trial mining. On the other hand, mining developments in different mines varied because of the different geological conditions of the mines. Moreover, surface construction might also cause misjudgment of interpretations.

4.3. Mining Subsidence Characteristics in Shendong Coalfield

To identify the mining subsidence characteristics in high-intensity mines, 19 continuous deformation areas with mining subsidence in the west bank of Wulanmulun River were investigated. These study areas were divided into A (A1, A2, A3, A4, and A5), B (B1, B2, and B3), C (C1), D (D1, D2, D3, and D4), E (E1, E2, and E3), F (F1 and F2), and G (G1) (Figure 14 A–G).

Figure 14. (A–H) Subsidence areas obtained from the InSAR time-series analysis (Western region to the Wulanmulun River).
In order to ensure that the section lines of each period would pass through the maximum subsidence point, the coherence images of the consecutive DInSAR interferometry were used, in which mining subsidence areas were determined as low-coherence areas. The centroids (i.e., geometric centers) of low-coherence areas were connected, thereby forming the principal sectional line of the maximum subsidence. Subsequently, this sectional line was used to extract the subsidence values under the consecutive DInSAR interferometry and the cumulative DInSAR interferometry. Then, the subsidence characteristic curve of 19 continuous deformation areas was drawn (Figure 15).

Only several representative working faces were illustrated and analyzed in this paper given the limit in article length.

In Figure 14, the subfigures ('A'–'G') show 7 sub-regions of the focused investigation in the west bank of Wulanmulun River, and they contain a total of 19 continuous deformation areas. The Subfigure 'H' is a regional aerial view that represents the relative spatial relationships of the 7 sub-regions.

In Figure 15, in a sectional graph way, interferometric results from 6 working faces (A3, B2, C1, D2, E1, and G1) are shown concentrically, which comprise that of the consecutive DInSAR interferometry located in left and that of the cumulative DInSAR interferometry located in right. The 6 working faces are located in Cuncaota No. 2 Mine, Cuncaota No.1 Mine, Bu’ertai Mine, Buliantai Mine, Shangwan Mine, and Huojitu Mine respectively (see also Figures 1 and 14).

**Figure 15. Cont.**
Along the centroids of subsidence basin of the working faces (see also Figure 14), 6 observation lines (length difference from 2200 m (B2) to 4840 m (G1)) were set, and re-sampled at an interval of 20 m along the line. So a series of subsidence curves were drawn along the time axis (interval difference from 336 day (A3, B2, and E1) to 504 day (C1 and G1)).

In Figure 15, a common attribute is in that the results of the consecutive DInSAR interferometry reappear some phenomena of a forward movement of the footprint of the deformation area with a series of perfect subsidence basins of subcritical mining such as the Figure 15a,c,e,g. Only in Figure 15i,k it represents a bit of fluctuate due the influence of near mining. Mean over, another common attribute is in that the results of the cumulative DInSAR interferometry represent a pattern of the subsidence basin area continuously expanded and the center of mining subsidence continuously moved forward (such as the Figure 15b,d,f,h,j,l).

Nevertheless, the Figure 15 also represents the characteristic of high-intensity mining subsidence. Firstly, it is in that maximum subsidence cannot be obtained from the results of the consecutive DInSAR interferometry due to strong phase discontinuities caused by the large deformation gradient. In other words, the subsidence curve is incorrect in the bottom of the bowl. Secondly, the results of the consecutive DInSAR interferometry express three different patterns of subsidence that the subcritical mining (Figure 15d,j), the critical mining (Figure 15b,l), and the supercritical mining respectively (Figure 15f,h). Finally, the results of the average subsidence rate during the observation period were range from 240 mm/year to 720 mm/year, which is a far less than that of the Poland [43] and of the Australia (a private communications, Ge, L.L., 2016.7 [45]).
4.4. Mining Subsidence Characteristics in Working Face 22201-1/2 in Bu’erTai Mine

Time-series analysis was conducted to discriminate ground deformation associated with past-mining activity and atmospheric effects, and consequently, acquire the spatial distribution characteristics, deformation pattern, development direction, evolution property, and temporal correlation of the multi-temporal mining subsidence phase-changing areas. For the determined boundaries of mining-induced damages, the classification of ground displacement and the calculated ground displacement parameters have practical significance. The experimental research only focused on working face 22201-1/2 in Bu’erTai Mine considering the heavy computing and analysis workloads. Multi-temporal DInSAR interferometry and SBAS-InSAR interferometry were adopted to calculate the available parameters. The subsidence patterns obtained using these two interferometric strategies were compared. The results are shown in Figure 16.

![Figure 16. Comparative analysis between multi-temporal DInSAR and SBAS-InSAR. (a) Time-series subsidence areas of the consecutive DInSAR; (b) Time-series subsidence curves of DInSAR; (c) Time-series subsidence areas of SBAS-InSAR; (d) Time-series subsidence curves of SBAS-InSAR.](image)

The subsidence curves of multi-temporal DInSAR interferometry with SBAS-InSAR interferometry were compared, and the following conclusions were drawn.
(1) In Figure 16a,b, the 17-temporal subsidence curves were obtained from the superposition of multi-temporal DInSAR interferometry, which could reflect general mining subsidence characteristics. That is, underground mining caused subsidence. With the expansion of the mining area, the subsidence area gradually expanded toward the advancing direction and subsidence continuously increased. The maximum subsidence became stable when it reached a certain extent, thereby forming a critical mining-induced subsidence bowl. However, rapid considerable subsidence caused by high-intensity mining (large mining height, thin bedrock, and fast advancing) brought serious decorrelation, and the principal value of phase change could not be effectively extracted. Accordingly, only the first temporal subsidence curve was connected smoothly because subsidence was within the unwrapping capacity of DInSAR ($\lambda/4 = 14$ mm) [46].

By contrast, in the subsequent observation the large deformation gradient resulted unwrapping failure and had a noise phase that was beyond the calculation capacity of the DInSAR. Such subsidence curves had no practical significance and could not obtained maximum subsidence. Furthermore, the effect of temporal decorrelation increased with temporal baseline, and the subsidence curve became insensitive. This condition was accompanied by a certain delay in the middle-temporal and late-temporal subsidence curves. The accumulation of residual phase changes also caused unexpected subsidence in non-mining areas.

(2) As shown in Figure 16c,d, SBAS-InSAR interferometry optimized the spatio-temporal baselines and improved the spatio-temporal sampling rates. The obtained results were more reliable than those of DInSAR. The 17-temporal subsidence curves of SBAS-InSAR did not only reflect the mining subsidence characteristics more accurately, but was also closer to the subsidence curve of probability integral estimation and maintained good continuity. However, it still failed to extract the principal value of phase changes because of the large deformation caused by mining subsidence. The acquired maximum subsidence was smaller than 200 mm, which was significantly different from that in practical situations.

(3) The mining area (1.23 km²) and subsidence area (5.58 km²) of the 22201-1/2 working face were calculated by further interpreting the results of SBAS-InSAR. The ratio of the mining area to the subsidence area was 1:4.5. The daily average advance rate of the working face was 8.1 m/day, and the daily average advance rate of subsidence area was 7.9 m/day. The average maximum subsidence angle was 67.3°. The advance angles of influence of DInSAR and SABS-InSAR were 53.5° and 50.5°, respectively.

5. Conclusions

In this study, InSAR and SBAS-InSAR techniques allow us understand the annual growing characteristics of subsidence disasters under high-intensity mining in Shendong Coalfield, and recognize the ground deformation features above a single working face.

(1) Conventional DInSAR is an effective approach for investigating ground damages caused by mining subsidence in coal mining areas. It demonstrates our tremendous destructive capability of the high-intensity coal exploitation. Over 85 phase change regions are identified within 504 days (20 January 2012 to June 2013) in Shendong Coalfield, including at that time and in the past, and including open-pit mining and underground mining areas. Annual subsidence area is about 150 km², approximately, which accounted for 4% of the total area of Shendong Coalfield (3481 km²).

(2) The time-series analysis of SBAS-InSAR achieves more reliable results and more accurately reflects the mining subsidence characteristics of a single working face. It makes some important parameters (such as the advance distance of influence and the advance angle of influence, which parameters can be used to predict the extent of mining subsidence so that action can be taken to prevent buildings from being damaged or other measures taken to control potential geohazards, etc.) is possible to be obtained.

(3) We find that the ground damage rate (7.9 m/day) caused by the high-intensity mining, is almost equal to mining rate of the working face (8.1 m/day), which should be a unique phenomenon occurred only at high-intensity mining conditions. In addition, we find that the advance distance of influence
is much greater than the distance of the exploitation, but the advance angle of influence (50.5°–53.5°) is much smaller than the angle of conventional mining (67.0°–70.0°), this may be due to the rate of water-loss settlement is much than that of mining subsidence.

(4) We find that some uplift are occurred at the front of working face in the direction of advancing, which may be caused systematically by orbit error, or by an inappropriate selection of the reference points. It is more likely a unique phenomenon associated with high-intensity mining. Although we have not yet to find a well-received explanation, it should be the focus of our future work, and we will try to understand and explain this phenomenon by means of the in-situ field measurements.

Acknowledgments: This research is jointly supported by the project of the National Natural Science Foundation of China and the Shenhua Coal Industry Group Co., Ltd. (Grant No. U1261106: Scattering Variation of Mining Subsidence Area Based on High-resolution SAR and Grant No. U1261206: The Mechanism of Surface Ecological Environment Evolution and Its Control due to High-Intensity Coal Mining) and the project of the Henan Province University Technological Innovation Team Support Program in 2013 (Grant No. 13IRTSTHN029: Monitoring and protection of mining-induced deformation in coal mining area). The authors would like to thank Ge, L.L., (UNSW) for his English editing in enhancing the quality of the paper.

Author Contributions: All the authors participated in the writing of the manuscript and have read and approved the final manuscript. Ma, C. and Zou, Y. conceived and designed the experiments. Cheng, X. and Yang, Y. processed the SAR data and performed the experiments. Zhang, X. implemented the coordinate transformation; Ma, C. and Guo, Z. analyzed the data.

Conflicts of Interest: The authors declare no conflict of interest.

References


43. Przyłucka, M.; Herrera, G.; Graniczny, M.; Colombo, D.; Bejar-Pizarro, M. Combination of conventional and advanced DInSAR to monitor very fast mining subsidence with TerraSAR-X data: Bytom City (Poland). *Remote Sens.* 2015, 7, 5300–5328. [CrossRef]


Investigating the Ground Deformation and Source Model of the Yangbajing Geothermal Field in Tibet, China with the WLS InSAR Technique

Jun Hu 1, Qijie Wang 1,*, Zhiwei Li 1, Rong Zhao 1 and Qian Sun 2

1 School of Geosciences and Info-Physics, Central South University, Changsha 410083, China; csuhujun@csu.edu.cn (J.H.); zwli@csu.edu.cn (Z.L.); zhaorong1018@126.com (R.Z.)
2 College of Resources and Environmental Science, Hunan Normal University, Changsha 410081, China; sunqian200241@aliyun.com

* Correspondence: qjwang@csu.edu.cn; Tel.: +86-731-8883-0573

Academic Editors: Zhenhong Li, Roberto Tomas, Richard Gloaguen and Prasad S. Thenkabail
Received: 11 December 2015; Accepted: 16 February 2016; Published: 26 February 2016

Abstract: Ground deformation contains important information that can be exploited to look into the dynamics of a geothermal system. In recent years, InSAR has manifested its strong power in the monitoring of ground deformation. In this paper, a multi-temporal InSAR algorithm, WLS InSAR, is employed to monitor and characterize the Yangbajing geothermal field in Tibet, China, using 51 ENVISAT/ASAR images acquired from two overlapping descending tracks. The results reveal that the WLS InSAR algorithm can suppress the adverse effects of seasonal oscillations, associated with the freezing-thawing cycle of the permafrost in the Qinghai-Tibet Plateau. Deformations of up to 2 cm/yr resulting from the exploitation of the geothermal resource have been detected in the southern part of the Yangbajing field between 2006 and 2010. A source model inversion of the subsurface geothermal fluids was carried out based on the elastic half-space theory using the accumulated deformations. It was found that most geothermal fluid loss has occurred in the southern part of the shallow reservoir as the pore space beneath the northern part of field was recharged by the ascending flow from the deep layers of the reservoir through well-developed faults in the region.

Keywords: InSAR; Yangbajing Geothermal Field; deformation monitoring; source model inversion; Qinghai-Tibet Plateau

1. Introduction

The Yangbajing geothermal field, located 94 km northwest from Lhasa City, the capital of the Tibet Autonomous Region, is characterized by the highest reservoir temperature in China. Since its initial exploration in 1976, the Yangbajing geothermal field has played an important role for the electricity power supply in Lhasa [1]. It is widely acknowledged that the production from the geothermal reservoir has caused considerable surface deformation in the Yangbajing basin, which could cause environmental and structural hazards, such as damage to the power plants [2]. However, the surface deformation provides useful information about the volume, pressure and temperature variations within the geothermal system, which makes it possible to study its dynamics [3]. The relationship between ground deformation and geothermal dynamics is straightforward, i.e., stress change caused by injection or production of fluid leads to surface deformation. Therefore, it is essential to employ geodetic techniques for monitoring the surface deformation in order to minimize the associated damage and infer the dynamics of the geothermal field and the injection/extraction processes.
Recent decades have witnessed a significant development of geodetic tools for ground deformation measurements. In addition to ground-based measurements (e.g., leveling and GPS), active geothermal fields can now be monitored using InSAR technology and satellite data. The advantages of InSAR, e.g., large-scale, high spatial resolution, low cost, all-day and all-weather imaging capability, have been widely proven in monitoring ground deformation caused by natural disasters [4–7] and anthropological activities [8–10]. In addition, many advanced InSAR algorithms have been developed since the beginning of this century [11–15]. By analyzing a time series of SAR images acquired along the same track, we find that these advanced InSAR algorithms enhance the accuracy and reliability of the derived average deformation and enable us to measure the evolution of surface deformation over a long period of time. Many geothermal fields have also been successfully observed by InSAR. For instance, the deformation and seismicity in the Coso geothermal field in eastern California have been observed by processing 10 ERS-1/2 images acquired between 1993 and 1998 [16]. The ground subsidence at the Tauhara and Ohaaki geothermal fields in New Zealand have been analyzed with 12 ALOS PALSAR images between 2007 and 2009 [17]. The deformation and its temporal variation at the Geysers geothermal field in California has been resolved by using two distinct sets of InSAR data (i.e., ERS-1/2 and TerraSAR-X), which agreed well with the coupled numerical modeling results [18].

In this paper, we present a study of the Yangbajing geothermal field during the time period 2006–2010 by analyzing a set of ESA ENVISAT C-band SAR data collected from two overlapping descending tracks. The Yangbajing geothermal field has already been monitored by the SBAS algorithm with C-band SAR data and the PS algorithm with X-band SAR data, respectively [19,20]. However, the seasonal oscillations of the permafrost, which have been detected in the Qinghai-Tibet Plateau due to the freezing-thawing cycle of the permafrost [21], were not previously considered in the retrieval of the displacements caused by geothermal exploitation. In this study, a WLS InSAR algorithm [22], a modified version of the SBAS algorithm, is used to distinguish the geothermal exploitation-related displacements and seasonal oscillations. Furthermore, the resolved deformation is used to characterize the source model, which can describe the behavior of the Yangbajing geothermal system. The source model parameters are searched by using a nonlinear optimization method (i.e., genetic algorithm), followed by a linear least square adjustment.

2. Study Area

The Yangbajing basin (Figure 1a), bordered by Pangduo Mountain to the southeast and Nynchen Tongliha Mountain to the northwest, is one of the most developed and largest fault basins in the Qinghai-Tibet Plateau, with a length of ~90 km and an area of ~450 km² [23]. The terrain of the basin is characterized by steep slopes in the southern area of the Zangbo River and gentle relief in the northwestern area, where several creeks have developed [2]. The China-Nepal Highway and Zangbo River pass through the basin. There has been intensive tectonic activity in the basin. For instance, 30 earthquakes with a magnitude ≥4.75 have occurred from 1921–1991 in the Yangbajing geothermal field [24]. As no evidence shows any occurrence of Quaternary volcanic eruptions in the basin, the Yangbajing geothermal field belongs to the non-volcanic high-temperature fields in the Qinghai-Tibet Plateau [24]. The basin is covered by different types of frozen soil, including seasonally frozen ground and permafrost [23].

The Yangbajing geothermal system includes two reservoirs at distinct depths [1,24,25]. The shallow reservoir, divided into northern and southern parts by the China-Nepal Highway, covers an area of approximately 14.6 km² at a depth of 180 ~ 280 m (Figure 1b). The temperature in the shallow reservoir ranges from 150 °C to 160 °C, and the northern part is generally a little hotter than the southern part. As the primary electricity generation source, the shallow reservoir yields thermal water at a rate of 1200 tons/day to the Yangbajing geothermal power plant [24]. The deep reservoir is a typical bedrock-fractured geothermal reservoir and covers an area of approximately 3.8 km² in the northern part of the field. The deep geothermal fluid is stored in fracture zones and tectonic fissures in the rocks, and migrates along the faults in the region. Downhole temperature measurements show that
the deep reservoir is characterized by higher temperatures and has at least two layers [2]. The upper layer, below a depth of at least 950 m, has a temperature of 250–275 °C and a thickness of more than 350 m. The lower layer is found below a depth of 1850 m and has a temperature higher than 300 °C. Although the shallow and deep reservoirs are distinguished depthwise by different temperatures, they belong to the same geothermal system. It has been confirmed that the occurrence of the Yangbajing geothermal field is due to a magmatic heat source [26]. The recharge of the Yangbajing geothermal fluids occurs because of the snowmelt water from the Nyenchen Tonglha Mountains at altitudes of 4400–5800 m [1]. The primary discharge paths were the hot springs before the power plant was built. Since the Yangbajing geothermal field started to produce electricity in 1977, intense exploitation dominates the discharge of the geothermal fluids [25].

![Figure 1](image_url)

**Figure 1.** (a) Optical image of the Yangbajing Basin. Dashed boxes outline the footprints of the ENVISAT ASAR descending images. The black rectangle indicates the study area. The inset provides the location of the study area in the Qinghai-Tibet Plateau; (b) Optical image of the study area, i.e., the Yangbajing geothermal field.

### 3. Datasets and Processing

A total of 51 SAR images acquired from two ENVISAT descending tracks are analyzed in the investigation of the Yangbajing geothermal field. The datasets span a time interval of four and a half
years from April 2006–September 2010. The footprints of the two sets of SAR images can be seen in Figure 1a. Clearly, more than half of the frames overlap, providing an opportunity to conduct a quantitative comparison. The datasets include 29 scenes from Track 176 and 22 scenes from Track 405, allowing us to generate 121 and 68 interferograms, respectively, after applying thresholds for the spatial-temporal baselines (Figure 2). The selected thresholds are a temporal baseline smaller than 700 days (approximately 2 years) and a perpendicular baseline shorter than 150 m. Precise orbital data provided by ESA are employed to refine the orbit vectors of the ENVISAT SAR images.

The standard two-pass differential InSAR approach is used to process each small baseline interferogram. The 90-m SRTM DEM data were used to remove the topographic contribution in the interferograms. A multi-look processing (i.e., 2 looks in range and 10 looks in azimuth directions) is adopted to suppress the decorrelation noise. In addition, the decorrelation noise in the generated differential interferograms is further minimized by using an improved Goldstein filter [27]. After a reference pixel is selected, the differential interferograms are then unwrapped by a minimum-cost flow algorithm in order to retrieve the real phase at the coherent pixels.

3.1. WLS InSAR Estimation

The phase at a pixel in a differential interferogram, including the contributions from the surface deformation, topographic residuals, orbital errors, atmospheric artifacts, and remaining phase noises, can be written as:

\[ \phi_{\text{int}} = \frac{4\pi}{\lambda} \Delta d_{\text{los}} + \frac{4\pi B_{\perp}}{\lambda R \sin \theta} \Delta \varepsilon + \phi_{\text{orbit}} + \phi_{\text{atmo}} + \phi_{\text{noise}}, \]  

(1)

where \( \phi_{\text{int}} \) is the phase in the differential interferogram, \( \Delta d_{\text{los}} \) denotes the phase component due to the LOS surface deformation between the master and slave acquisitions, \( \Delta \varepsilon \) is phase component due to the topographic residuals, \( \phi_{\text{orbit}} \) and \( \phi_{\text{atmo}} \) are the phase components due to the inaccurate orbit and atmospheric delays, respectively (which usually behave as phase ramp throughout the whole image and can be removed by the polynomial fitting method [28,29]), \( \phi_{\text{noise}} \) is the phase component due to the remained noises, \( \lambda \) is the wavelength, \( B_{\perp} \) is the perpendicular baseline, \( R \) is the slant range between the sensor and the target, and \( \theta \) is the incidence angle.

According to Equation (1), it is obviously difficult to resolve the deformation of interest from other contributions (e.g., the topographic residuals), if only a single differential interferogram were used. However, when many differential interferograms along the same track are available, a time series analysis can be conducted to estimate the LP temporal component of the surface deformation. This is how most current advanced InSAR algorithms work. Although the production-induced deformation can generally be characterized by gradual accumulation, the seasonal variations require consideration in this study due to the freezing-thawing cycle of the permafrost in the Qinghai-Tibet
Plateau. Therefore, the LP deformation component can be modeled by the following superimposed function [22]:

\[
\Delta d_{los} = \alpha_1 \Delta t + \alpha_2 \sin \left( \frac{2\pi}{T} \Delta t \right) + \alpha_3 \cos \left( \frac{2\pi}{T} \Delta t \right)
\]  

(2)

The first term on the right side of this equation is the linear deformation component, the second and third terms reflect the seasonal deformation component, \(\Delta t\) is the time interval between the master and slave acquisitions, \(T\) is the period of the seasonal deformation (one year in this study), and \(\alpha_1, \alpha_2\) and \(\alpha_3\) are the amplitudes of the base functions requiring estimation. Assuming there are \(m\) differential interferograms available along a track, a linear system with four unknowns and \(m\) observations can be constructed as:

\[
\Phi = B \cdot X
\]  

(3)

where \(\Phi = [\phi_{int,1}, \phi_{int,2}, \cdots, \phi_{int,m}]^T\) is the observation matrix, \(X = [\alpha_1, \alpha_2, \alpha_3, \Delta t]^T\) is the unknown matrix, and \(B\) is the design matrix

\[
B = \begin{bmatrix}
\frac{4\pi}{\lambda} \Delta t_1 & \frac{4\pi}{\lambda} \sin \left( \frac{2\pi}{T} \Delta t_1 \right) & \frac{4\pi}{\lambda} \cos \left( \frac{2\pi}{T} \Delta t_1 \right) & \frac{4\pi}{\lambda R \sin \theta} B_{1,1} \\
\frac{4\pi}{\lambda} \Delta t_2 & \frac{4\pi}{\lambda} \sin \left( \frac{2\pi}{T} \Delta t_2 \right) & \frac{4\pi}{\lambda} \cos \left( \frac{2\pi}{T} \Delta t_2 \right) & \frac{4\pi}{\lambda R \sin \theta} B_{1,2} \\
\vdots & \vdots & \vdots & \vdots \\
\frac{4\pi}{\lambda} \Delta t_m & \frac{4\pi}{\lambda} \sin \left( \frac{2\pi}{T} \Delta t_m \right) & \frac{4\pi}{\lambda} \cos \left( \frac{2\pi}{T} \Delta t_m \right) & \frac{4\pi}{\lambda R \sin \theta} B_{1,m}
\end{bmatrix}
\]  

(4)

This linear system can be resolved using the WLS adjustment when \(m \geq 4\):

\[
X = (B^T \Sigma^{-1}_\Phi B)^{-1} B^T \Sigma^{-1}_\Phi \Phi,
\]  

(5)

where \(\Sigma_\Phi\) reflect the stochastic properties of the observation matrix, composed by the variances of the differential interferograms. For each pixel, the variance estimate is based on the probability density function of the differential phase, which depends on the interferometric coherence and multi-look number [22].

The phase residuals, contributed by the HP temporal phase components and the unmodeled atmospheric phase components (i.e., turbulent atmosphere), can be separated from the phase measurements after removing the estimated LP deformation components and topographic residuals. A least-squares inversion is then applied to retrieve the time series of the phase residuals, corresponding to the SAR images, assuming its mean is zero. Subsequently, spatial median and temporal triangle filtering are carried out to separate the turbulent atmospheric phase components from the time series of the phase residuals. Thus the estimated turbulent atmospheric phase components and topographic residuals are subtracted from the differential interferograms. Finally, the refined differential interferograms are processed using the WLS adjustment again, in order to estimate the time series of surface displacements.

3.2. Source Model Inversion

Once the surface deformation due to the dynamics of the geothermal system is accurately measured, it should be inverted to retrieve the geothermal parameters at depth. In this study, we assume a volumetric source component for the Yangbajing geothermal system based on the elastic half-space theory [3]. This means that the surface deformation at the geothermal field is caused by the movement of fluids and gases generated by the geothermal activity beneath the field. It should be noted that although not all tectonic stresses can be described, the volumetric source component is an acceptable assumption for a geothermal system [30]. Therefore, when the Earth deforms elastically,
a linear relationship between the surface displacement and the subsurface volume change can be constructed as [31]:

$$d_i(x) = \int_V G_i(x, y) \Delta v(y) dy,$$

(6)

where $d_i(x)$ is the $i$th component of surface displacement at pixel $x$ ($i = 1, 2, 3$ correspond to the up, east, and north components, respectively), $\Delta v(y)$ is the fractional volume change of block $y$ with a source volume $V$, and $G_i(x, y)$ is the Green’s function of the observation pixel $x$ and source block $y$ using a homogeneous elastic half-space assumption [32], the Green’s function can be expressed as follows:

$$G_i(x, y) = \frac{(v + 1) (x_i - y_i)}{3\pi S^3},$$

(7)

where $S = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$ is the distance between the observation pixel $x$ and the source block $y$, and $v$ is Poisson’s ratio (0.25 in this study).

For more details about the presented forward modeling of the volumetric source component, see [33]. However, it should be noted that other sources, such as finite rectangular source [34], nucleus-of-strain formulation [35] and its developers in geophysics [36] and reservoir engineering [37], can also be used to construct the relationship between the surface displacement and the subsurface dynamics.

Since the InSAR-derived LOS measurement is related to the up, east, and north deformation components according to the radar imaging geometry [38,39], Equation (6) can be rewritten as:

$$d_{\text{los}}(x) = \int_V (u_1 G_1(x, y) + u_2 G_2(x, y) + u_3 G_3(x, y)) \Delta v(y) dy,$$

(8)

where $u_1$, $u_2$ and $u_3$ are the components of unit vector, projecting from the up, east, and north directions onto the LOS direction, respectively:

$$\begin{align*}
    u_1 &= \cos \theta \\
    u_2 &= -\sin \theta \cos (\alpha - 3\pi/2) \\
    u_3 &= -\sin \theta \sin (\alpha - 3\pi/2)
\end{align*},$$

(9)

where $\theta$ is the radar incidence angle and $\alpha$ is the azimuth angle (clockwise from north).

Since a number of displacement measurements are provided by InSAR, the set of Equation (6) can be written into a vector-matrix form:

$$D = GV,$$

(10)

where $D$ is the observation vector containing the InSAR LOS displacement measurements at the surface points, and $V$ is the unknown vector containing the fractional volume changes within a grid of subsurface blocks. Apparently, this system can be solved by the linearization approach (i.e., least-squares adjustment) if the design matrix $G$ is accurately determined and the number of the observations is larger than that of the surface blocks. The latter condition can be easily satisfied, since we can divide the subsurface volume into a coarse grid (several times the InSAR pixel size) and a few layers (fewer than three layers). However, the former condition depends on the a priori geometrical information about the source volume, i.e., the depth and thickness of the geothermal system. We can employ a nonlinear optimization method (e.g., genetic algorithm) to estimate the best-fitted geometrical parameters in case they are not accurately provided. In order to balance the estimation accuracy and the computation burden of the nonlinear inversion we assume that the fractional volume changes are uniform at all blocks, so that only three unknowns (depth, thickness, and uniform volume change) require estimation. In addition, a quadtree down-sampling algorithm is applied to reduce the large number of InSAR point observations, while retaining sufficient information for inversion.
As the geometric parameters of the geothermal system are either provided in advance or estimated by the genetic algorithm, the fractional volume change at each block can be estimated by applying the least-squares adjustment to Equation (10). Note that the original sampling InSAR point observations are used in the linear inversion in order to make use of as many observations as possible. However, although the linear system is over-determined, the solution of the fractional volume changes could be unstable, due to the observation errors and the errors of design matrix. A roughness penalty should be included in the system to stabilize the least-squares inversion [3,40]:

\[
\begin{bmatrix}
D \\
0
\end{bmatrix} = \begin{bmatrix}
G & \Pi \\
W_r & 0
\end{bmatrix} \begin{bmatrix}
V
\end{bmatrix},
\]

(11)

where \( \Pi \) is the penalty matrix that estimates the spatial derivative of the fractional volume change, and \( W_r \) is a weighting factor controlling the relative importance of the roughness penalty term with respect to the data misfit term, which can be solved by the method of trial and error in a number of inversions [3]. Therefore, continuity can be preserved for the estimated fractional volume changes, which ensures that the geothermal system represents a connected network rather than a collection of isolated blocks. The roughness penalty is however not necessary if the a priori knowledge about the reservoir is provided [41,42].

Besides the presented two-step scheme, the inversion can also be conducted by some integrated schemes (e.g., the Ensemble-based approaches [42,43]), which can avoid the possible inappropriate results yielded by the nonlinear inversion. The two-step scheme is preferred in this study since its applicability has been widely demonstrated in many geophysical inversion problems related to such seismic events. In addition, it is quite simple and efficient to conduct the two-step scheme.

4. Results

The linear LOS surface displacement velocity maps are first estimated by the WLS InSAR algorithm for Tracks 176 and 405, respectively, as shown in Figure 3a,b. Note that the exposed rock at the foot of the Nyenchen Tonglha Mountain is selected as the reference area, where the deformation can be neglected (see black square in Figure 4a). In addition, the areas with low coherence (<0.35) have been masked out to avoid unwrapping errors induced by large decorrelation noise. Figure 3 shows that most areas in the Yangbajing basin were rather stable, with LOS displacement rate below 0.5 cm/yr. However, the linear displacement velocity estimated by the data from Track 176 (Figure 3a) shows that the Yangbajing geothermal field (black box in Figure 1a), has experienced obvious local surface deformation. This ground subsidence is also detected by the data from Track 405, although only the right half of the geothermal field is covered (Figure 3b). As mentioned above, a linear ground deformation is expected to be caused by the withdrawal of subsurface fluid from the Yangbajing geothermal system. Figure 4a shows the superimposed displacement velocity fields estimated by both sets of SAR data, where the overlapping regions are difficult to distinguish due to their good resemblance. An enlarged view of the Yangbajing geothermal field is shown in Figure 5a. Most of the ground subsidence occurs in the southern part of the Yangbajing geothermal field, with a maximum LOS velocity of ~2 cm/yr. The northern part experiences gentle, smaller-scale ground subsidence, with LOS velocities generally smaller than 1 cm/yr. It is also observed that the ground subsidence increases from the margins to the center of the geothermal field. However, the displacements near the hydrothermal explode hollow and the Zangbo River are unavailable due to decorrelation of the InSAR observations. This is somewhat expected since the moist and changeable features of the ground surface at the discharge area of the geothermal field is quite unfavorable for InSAR observations. It is possible that even larger ground subsidence have occurred in that area.
From the LOS surface displacement evolution estimated by the WLS InSAR algorithm for Tracks 176 and 405, we found that the whole Yangbajing basin experiences obvious seasonal oscillations, i.e., ground uplift in the cold season and subsidence in the warm season. This is to be expected, since this area is covered by the permafrost, deforming with the freezing-thawing cycle of the active layer [23]. However, the surface displacements induced by geothermal production are somewhat obscured due to the seasonal oscillations. In order to present the details of the local deformation evolution, the deformations at two selected points marked A and B in Figure 5a, are shown in Figure 6a,b, respectively. It is observed that the displacement evolutions estimated from Track 176 and 405 agree quite well with each other. Point A shows seasonal variations in the margins of the geothermal field, where the cumulative displacement velocities are smaller than 2 mm/yr. Point B at the center of the geothermal field is characterized by both seasonal variation and cumulative displacement. The cumulative displacement velocity at point B, about −1.5 cm/yr, reveals that the Yangbajing geothermal field experiences obvious ground subsidence. Profile CC’ across the field, whose location is marked with a dashed line in Figure 5a, is also shown in Figure 6c, exhibiting the progress of the deformation along the profile. In order to better represent the cumulative deformation associated with the production of the Yangbajing geothermal system, four deformation evolutions acquired on nearly the same dates (but different years), i.e., 15 April 2007, 4 May 2008, 19 April 2009, and 9 May 2010, are selected to minimize the effects of the seasonal oscillations. The reference time of the four deformation evolutions is 30 April 2006. As displayed in Figure 6c a subsidence funnel was clearly formed in the Yangbajing field that kept growing with time.
Figure 5. Maps of measured and simulated displacement velocities at the Yangbajing geothermal field, as superimposed on the optical map. (a) Linear displacement velocity map from InSAR; (b) Down-sampled linear displacement velocity; (c) Simulated linear displacement velocity; (d) Differences between the InSAR-estimated and the simulated linear displacement velocity. Triangles mark selected points A and B (see text). The dashed line shows the location of profile CC’. Black and white solid lines mark the China-Nepal Highway and Zangbo River, respectively.

Figure 6. Point time series of deformation and deformation profiles. (a,b) Evolutions of LOS surface deformation at points A and B marked in Figure 5a. Circles and triangles represent the results of Track 176 and 405, respectively; (c) LOS surface deformation (with respect to 30 April 2006) along profile CC’. Deformations acquired on similar dates in four consecutive years are presented, in order to minimize the effects of seasonal variations.
The estimated LOS linear surface displacements at the Yangbajing geothermal field are further down-sampled by the quadtree algorithm to provide less-dense observations for the nonlinear inversion of the geometrical information about the source volume. In the quadtree down-sampling, the LOS displacement field is first divided by the blocks with a dimension size of $64 \times 64$. These initial blocks are then divided into four equal-sized square blocks if the difference between the maximum and the minimum displacements of the block elements is less than 2 mm, until the dimension size of block reaches $4 \times 4$. The result is shown in Figure 5b, achieving 844 data points, 2% of the original 43,684 displacement observations. Using the down-sampled InSAR displacement observations, the best fitted geometrical parameters are inverted by using a genetic algorithm with 51 iterations. These estimates are 189.1 and 81.0 m for the reservoir depth and thickness, respectively. Subsequently, the shallow reservoir is divided into a rectangular grid of size $100 \times 128$, i.e., two times coarser than that of the InSAR pixel grid. This means that the grid spacing of the shallow reservoir is ~90 m. The fractional volume change at each block is then determined by using a least-squares adjustment with a roughness penalty, as shown in Figure 7. The total value of the fractional volume change is estimated at $-21.5$, which is comparable to that of the uniform volume change ($i.e., -24.3$). Since no a priori information about the reservoir can be used to constrain the distribution of volume change, the roughness penalty term $W_r$ of 10 is used in the inversion, on the basis of an examination of the data misfit. In other words, we pick $W_r$ such that the difference between the InSAR-measured and the simulated displacement velocity fields is insignificant. In addition, we adopt an iterative strategy to ensure the robustness of the inversion. In the initial processing, the variety range of uniform volume change is relatively large to cover most situations. In the iterative processing, the variety range of uniform volume change is set to smaller by using the average value of the fractional volume change estimated from the last iteration as the reference. The iterative processing will be terminated when the standard deviation of the difference between the two consecutively estimated fractional volume changes is smaller than a pre-determined threshold (0.002 in this study).

Figure 7. Fractional volume changes of the Yangbajing geothermal system inverted from the InSAR displacement observations. Black and white solid lines represent the China-Nepal Highway and Zangbo River, respectively. Red dashed lines and black dots mark fault locations and the locations of production wells [2].
The linear displacement velocity field simulated by the inverted source model is shown in Figure 5c and then compared with the InSAR-estimated one. The two displacement velocity fields are rather similar. Figure 5d shows the differences between the InSAR-estimated and the simulated displacement velocities. Most of the differences range from −1 to 1 mm, demonstrating that the inverted source model for the shallow reservoir can well explain the linear surface displacements measured in the Yangbajing geothermal field. Despite the overall agreement, some relatively large residuals (up to 4 mm) can be found around the decorrelated areas. They are characterized by a grid-shape, mostly due to the InSAR observation noises and errors associated with the sparse grid used for the geothermal model inversion.

As expected, obvious volume changes are found under the ground where subsidence occurs (Figure 7). The total volume change of the Yangbajing geothermal system was estimated at $-1.41 \times 10^7$ m$^3$/yr, indicating that a significant volume of geothermal fluids had been extracted from the shallow reservoir during the investigated period. Most of the source volume reduction occurred in the southern part of the field, between the Nepal Highway and the Zangbo River.

5. Discussion

Field surveys have been carried out in the Yangbajing field since 1983 to monitor the ground subsidence associated with the geothermal exploitation. The observed mean annual velocity reached ~3.5 cm/yr at the southern part of the field [2]. Field work was suspended in 1994 due to shortage of financial support, so there are no in situ data to assess the accuracy of the InSAR-measured surface deformation. However, the decrease of the mean annual velocity during the investigated period is reasonable, since the main geothermal exploitation gradually transferred from the southern to the northern field [44]. The observed deformation velocities are comparable to other measurements with the InSAR technique [19,20]. Furthermore, we performed quantitative comparisons in the overlapping area between the displacement fields observed using SAR data from Tracks 176 and 405. As shown in Figure 4b, the differences basically range from −2 to 2 mm/yr, with a standard deviation (STD) of 0.1 mm/yr and a mean of 0.4 mm/yr. This good agreement indicates that the WLS InSAR algorithm is suitable to estimate the linear surface displacements in the geothermal field. The residuals can be explained by the differences between the incidence angles of the two sets of SAR images. Compared to conventional survey methods, InSAR requires much less human and financial support, and therefore can be used as a routine tool in the investigation of the geothermal field.

Also, a 4-km portion of the China-Nepal Highway passes through the subsidence region. Although the highway and railway constructed in the permafrost area have been designed to prevent hazards related to the freezing-thawing cycle of the active layer, the cumulative inhomogeneous deformation induced by the geothermal production represents a more serious threat to road infrastructure. Therefore, the InSAR-derived displacements should be taken into account in any further exploitation plan for the Yangbajing geothermal field.

Although it has been reported that the Yangbajing geothermal system consists of shallow and deep reservoirs [25], in this study we only consider the shallow layer for the source model inversion. This simplification is realistic for three reasons. First, most thermal water used for electricity generation was extracted from the shallow reservoir during the study period [24]. Second, the deep geothermal fluid is stored in the fracture zones and tectonic fissures at least 950 beneath the surface m, where the space and pressure are fairly steady. Therefore, it has quite limited effect on the ground surface deformation, which depends on the combination of pressure reduction and the effective elastic modulus of the reservoir. Thirdly, the deep reservoir is expected to be located in the northern part of the Yangbajing field and to generate the shallow reservoir in the Quaternary pores through a side flow recharge of hot fluids [2]. Therefore, monitoring the dynamics of the shallow reservoir reflects in a certain degree the dynamics of the deep reservoir.

The production wells are distributed uniformly in the Yangbajing field, 24 in the northern part and 31 in the southern part (see dots in Figure 7). The localization of the volume reduction in the
southern part indicates that the deep reservoir is located in the northern field and recharges the shallow reservoir there. This result supports the conceptual reservoir models for the Yangbajing geothermal system proposed in [1,2]. Red dashed lines in Figure 7 denote the observed and inferred faults in the Yangbajing field. We found that rhombic block structures are crossly formed along the northeast and the northwest striking faults. The longest northeast striking fault is located near the Nepal Highway, serving as a boundary between the northern and southern geothermal systems. The mature faults in the northern part provide ideal conditions for the ascending flows from the deep reservoir. However, only one of the northwest striking faults extends to the southern field. This could explain why the fluid loss in the southern part of the shallow reservoir cannot be immediately replenished by the flow from the deep reservoir.

6. Conclusions

Monitoring the surface deformation induced by the geothermal production is essential to prevent possible geo-hazards, as well as to infer the dynamics of the subsurface flow fluids. In this study we employ an advanced InSAR algorithm, termed WLS InSAR, to investigate the Yangbajing geothermal field in the Qinghai-Tibet Plateau. We used a set of ENVISAT ASAR data acquired from two adjacent descending tracks from 2006 to 2010. By introducing a seasonal deformation model, the accumulated displacements caused by the geothermal production are separated from the seasonal oscillations associated with the freezing-thawing cycle of the permafrost. The elastic half-space theory is then applied to characterize the source model of the Yangbajing geothermal system from the WLS InSAR displacement measurements. Two main conclusions can be summarized from this study:

1. The WLS InSAR algorithm can be applied to detect and monitor ground deformation occurring in geothermal fields. Although in situ observations are unavailable during the study period, the pattern and amplitude of the InSAR-measured displacements are reasonable according to the field survey between 1983 and 1994. A cross-validation has been conducted between the two adjacent tracks’ InSAR measurements, indicating an accuracy of millimeter or even sub-millimeter level. It is observed that the southern part of the Yangbajing field experiences considerable linear deformation along the LOS direction, of up to ~2 cm/yr, which behaves as ground subsidence and radial contraction. In addition, almost 8 cm peak-to-peak seasonal oscillations have been inferred from the deformation evolution results.

2. InSAR-measured displacements are very useful for the inversion of the source model of the subsurface geothermal fluids. The geometrical parameters and the fractional volume changes of the Yanbajing geothermal system have been determined by a genetic algorithm and least squares adjustment, respectively. The results reveal that geothermal fluids of $1.41 \times 10^7$ m$^3$ had been extracted from the shallow reservoir every year in the period 2006–2010. The fluid loss in the southern part of the shallow reservoir dominates the source volume reduction. In the northern part, the discharged fluid has been refilled by ascending flow from the deep reservoir. The reason for these assumptions is that the deep reservoir is located in the northern field and has several developed faults serving as natural pipelines for the thermal fluids.

Since field surveying is always a time- and labor-consuming work, InSAR should be viewed as a routine tool to monitor and characterize the Yangbajing geothermal field in the future. In particular, the newer TerraSAR-X, COSMO-SkyMed and Sentinel-1 SAR data have finer resolution and shorter revisit period than the ENVISAT ASAR data, and the use of them will significantly enhance the spatial and temporal resolutions of the measured surface displacements. Although the volume change in the shallow reservoir is proven sufficient to fit the InSAR-measured displacements, the interactive behavior of the fluid flows between the shallow and deep reservoirs should be considered in order to investigate the interior dynamics of a geothermal system, especially when more surface displacement observations and subsurface geometrical information are provided.
Acknowledgments: The ENVISAT ASAR data are provided by the European Space Agency through ESA EO Projects No. 14662 and 29255. This work was supported by the National Basic Research Program of China (Nos. 2013CB733303, 2012CB719903), the Key Laboratory for National Geography State Monitoring (National Administration of Surveying, Mapping and Geoinformation) (No. 2013NGCM07), the Key Laboratory of Videometric and Vision Navigation of Hunan Province, China (No. TXCL-KF2013-002), and Nature Science Foundation of China (Nos. 41404011, 41404013, 41222027 and 41474007). The comments from three anonymous reviewers and editor have greatly improved the quality of the manuscript.

Author Contributions: Jun Hu, Qijie Wang and Zhiwei Li conceived the study, performed the experiments and produced the results. Jun Hu drafted the manuscript. Qijie Wang and Zhiwei Li contributed to the discussion of results. Rong Zhao helped to analysis the permafrost seasonal oscillations. Qian Sun helped to conduct the source model inversion. All authors reviewed and approved the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

WLS Weighted Least Squares
InSAR Interferometric Synthetic Aperture Radar
ESA European Space Agency
PS Persistent Scatterers
SBAS Small Baseline Subsets
LOS line-of-sight
LP low-pass
HP high-pass

References

4. Li, Z.; Elliott, J.R.; Feng, W.; Jackson, J.A.; Parsons, B.E.; Walters, R.J. The 2010 Mw 6.8 Yushu (Qinghai, China) earthquake: Constraints provided by InSAR and body wave seismology. J. Geophy. Res. 2011, 116. [CrossRef]


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
InSAR Observation and Numerical Modeling of the Earth-Dam Displacement of Shuibuya Dam (China)

Wei Zhou 1, Shaolin Li 1,2, Zhiwei Zhou 3,* and Xiaolin Chang 1

1 State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China; zw_Mxx@163.com (W.Z.); shaolin@whu.edu.cn (S.L.); changxl@whu.edu.cn (X.C.)
2 Changjiang Institute of Survey, Planning, Design and Research, Wuhan 430010, China
3 Global Navigation Satellite System Research Centre, Wuhan University, Wuhan 430079, China
* Correspondence: zhiwei8848@gmail.com; Tel.: +86-27-6877-8240; Fax: +86-27-6877-8971

Abstract: How to accurately determine the mechanical parameters of rockfill is one of the key issues of concrete-face rockfill dams. Parameter back-analysis using internal or external monitoring data has been proven to be an efficient way to solve this problem. However, traditional internal or external monitoring methods have limitations in efficiency and long-term monitoring. In this paper, the displacement of the Shuibuya concrete-face rockfill dam is monitored by the space-borne Interferometric Synthetic Aperture Radar (InSAR) time series method. Using the InSAR results and the finite element method, the back-analysis of the mechanical parameters of the rockfill dam is investigated, and the back-analysis results of InSAR and levelling are compared. A high correlation of 0.99 for the displacement results generated from InSAR and the levelling offers good agreement between the two methods. The agreement provides confidence that the external InSAR monitoring measurement allows producing a reliable back-analysis and captures the displacement properties of the dam. Based on the identified parameters from the InSAR results, the dam displacement is predicted. The prediction of the maximum settlement of the dam is 2.332 m by the end of 2020, according to the dam displacement characteristics, which agrees well with the results derived from the recorded internal monitoring data. Therefore, the external monitoring results from the InSAR observation can be used as a supplement for traditional monitoring methods to analyse the parameters of the dam.

Keywords: Interferometric Synthetic Aperture Radar (InSAR) time series; Concrete-Face Rockfill Dam (CFRD); numerical modelling; finite element method; back-analysis

1. Introduction

In the stability and security analysis of the concrete-face rockfill dams (CFRDs), the physical parameters of the rockfill are pre-requisite for settlement prediction. Because of the complexity of the dam characteristics, the material properties are influenced by the loading, load size, stress path and other factors in the engineering project. Currently, engineering analogy, expert experience and experimentation are used to determine the material parameters of the rockfill. However, engineering analogies and expert experience require users who have rich experience in engineering and detailed engineering data. Lots of subjective factors are contained in the results. Due to the scale effect, sample disturbance and the impact of random sampling, it is difficult to guarantee the representation of the experiment results [1,2]. Moreover, even though the accurate design parameters of the rockfill are obtained by experiment, the values are greatly different from the actual engineering parameters due to the influence of the construction process, construction technology, construction quality, the external
environment and other factors. Therefore, the stress–strain analysis, operational behaviour analysis and safety evaluation based on the experimental parameters cause great error in CFRDs.

Parameter back-analysis using the internal or external monitoring data has proven to be an efficient way to identify the physical parameters of the rockfill [3,4]. Based on the identified parameters, the stability and security analysis of the CFRDs can be investigated using the finite element method (FEM) [5]. The internal monitoring methods include tension wire alignment and hydraulic overflow settlement gauges [6–8]. However, these internal monitoring methods do not meet the safety monitoring requirements of large dams in terms of efficiency and long-term observations because of their low coverage and durability, and labour-intensive monitoring needs. The internal monitoring data, the optimization algorithm and the artificial neural network are used in traditional back-analysis to identify the parameters of the rockfill dam [3,9–11]. Generally, external monitoring methods include levelling or Global Position System (GPS) measurement, but they are rarely used in back-analysis. Traditional back-analysis requires many training samples when using the artificial neural network and is time-consuming and inefficient. To overcome these disadvantages, a response surface method (RSM) may be applied during the processing [4].

In the past two years, interferometric synthetic aperture radar (InSAR) has proven to be a powerful tool to measure the Earth’s surface movements. Few studies have focused on using InSAR to map the displacement of dam bodies and/or reservoir slopes [12–15], and back-analysis in high CFRDs based on InSAR observations has not been performed. In this study, we aim to use the space-borne InSAR time series results of Shuibuya (SBY), the RSM and the FEM to investigate the physical parameters of the rockfill of the SBY Dam and to predict its long-term displacement of the dam. The numerical results calculated using the back-analysis model parameter set based on the InSAR results are compared with the levelling monitoring data.

This paper is structured as follows: Section 2 presents the general description of the SBY Dam. Section 3 presents the SAR data, a short description of InSAR time series processing and the InSAR time series results. The constitutive model, back-analysis method, identified parameters set, validation, stress and displacement analysis are given in Section 4. Finally, the discussions and conclusions are presented in Sections 5 and 6.

2. Study Area

Located on the middle stream of the Qingjiang River in the Enshi Autonomous Prefecture, Hubei Province, China, the SBY CFRD (110.3377°N, 30.4374°E) is approximately 117 km upstream of the city of Enshi and 92 km from the downstream Geheyan Hydropower Plant (Figure 1). The height of the SBY CFRD is up to 233.5 m, and the crest length is 675 m. The normal water level of the reservoir is 400 m above sea level, and the maximum water storage of the reservoir is approximately 4.59 × 10⁹ m³. The SBY Dam has four generators in its underground power plant, with a total electric generating capacity of 1840 MW. It is a typical concrete-face rockfill dam, and its body is composed of six material zones: arbitrary fill material, bedding material, transition material, primary rockfill, secondary rockfill and downstream rockfill (Figure 2a). The upstream and downstream of the dam has the same slope of 1:1.4. The construction began in October 2002 and can be divided into three reference periods:

I  Dam construction period (before October 2006).
II  The first reservoir filling period (October 2006–September 2007, when the water level was between 205.06 m above sea level and 389.61 m).
III  Operation period (after September 2007).

Because of the significance of the project, an improved and detailed settlement monitoring system was implemented for the SBY CFRD, including eleven levelling monitoring stations distributed downstream of the dam. The eleven monitoring stations (WS1 to WS11) were established at elevations of 235 m, 265 m, 300 m, 335 m, and 370 m to monitor the settlement of the downstream dam surface (as shown in Figure 2b).
3. InSAR Data Analysis and Results

A detailed InSAR data processing steps and analysis about the SBY dam is performed in our previously paper [16]. Here, we present a short description of the InSAR processing and related results. Twenty-one ascending ALOS-1 PALSAR L-band (~23.6 cm wavelength) images acquired from the Japan Aerospace Exploration Agency were used to map the displacement of the SBY Dam. The ALOS-1 satellite was launched in January 2006 and terminated in April 2011. Only images from 28 February 2007 to 11 March 2011 were processed, representing the period after completion of the dam (Figure 3).
In the single-imaging mode, the resolution of the images is approximately 4.7 m in the slant range and 4.5 m in the azimuth direction. The incidence angle is approximately 38.7°. There are many archived ESA Envisat ASAR C-band (~5.6 cm) images, however the resolution is too low (~20 m). Considering the scale of the dam, ASAR data is not suitable for this study. Although the resolution of the TerraSAR-X and Cosmo-SkyMed X-band (~3 cm) images are much higher (up to 1 m), there is very little archived data in this area at present. That means these data are not suitable for the InSAR time series analysis.

![Figure 3](image1.png)

**Figure 3.** Interferogram distribution in terms of the spatial and temporal baseline, adapted from [16]. Each black circle represents one SAR image and the black solid line between two circles represents one interferogram.

The GAMMA processing package was used to generate single look complex (SLC) products from the set of raw SAR images [17]. All SLC images were co-registered to a chosen master image (2 June 2008) to match the dataset to its later interferogram generation. Next, an area of approximately 5 km × 5 km (1000 pixels × 1000 pixels) centred on the SBY Dam was cropped from the original 70 km × 70 km SLC images (Figure 4a). All possible interferograms were generated with a spatial baseline below 2000 m and temporal baselines below 30 months. Then, interferograms were investigated visually, and only interferograms with a valid phase and coverage greater than 70% were kept for later InSAR time series analysis. Finally, 26 interconnected interferograms were selected (Figure 3). Topographic fringes of each interferogram was subtracted by using the 1-arc second (~30 m) Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) [18], and the MCF algorithm was used to unwrap the phase [19]. Given the image resolution and the scale of the SBY Dam, multi-looking processing was not performed in this study.

![Figure 4](image2.png)

**Figure 4.** The cropped study area and the average coherence map in radar geometry (not geocoded): (a) The amplitude image of the study area; and (b) The average coherence map, where white indicates pixels with strong coherence and black indicates pixels with low or no coherence. This figure is quoted from [16].
Generally, the unwrapped phase consists of five terms: displacement signal, inaccurate orbital information, topographic error, atmospheric propagation delays (atmospheric phase screen, APS), and noise [20]. There are various approaches have been developed to reduce the latter four terms based on multi-interferogram analysis [21–25]. In this study, the InSAR Time Series with the Atmospheric Estimation Model (InSAR TS + AEM, [26]) is used to perform the time series analysis, which is based on the Small Baseline Subset algorithm (SBAS) (e.g., [22]). The temporal or intermittent coherent pixels strategy was applied to increase the density and distribution of stable pixel [27,28]. The Figure 4b shows the average coherence of the study area.

Because the study area is approximately 5 km × 5 km, a linear plane across the entire interferogram is sufficiently accurate to remove the orbital error [20]. For a given set of unwrapped interferograms, topographic error signals can be separated because they are proportional to the perpendicular baselines [20]. Therefore, the orbital and topographic error can be estimated and individually removed from the unwrapped phase. Taking into account the spatial structure of atmospheric effects (only correlated in space) (e.g., the power-law process [29,30]), the APS can be estimated using a temporary linear velocity (TLV) model and can be distinguished from non-linear surface motion, more details can be found from [26]. The algorithm was applied iteratively until convergence was achieved [26,31]. A small area (500 m × 500 m) in the village was assumed to be stable during the InSAR observations and was therefore selected as the reference site for the InSAR time series analysis (Figure 4a). For a given pixel, the LOS displacement contains three displacement components in east, north and up direction. It is a function of incidence angle and above three components:

\[ d_{\text{los}} = \cos \alpha \sin \vartheta d_e - \sin \alpha \cos \vartheta d_n - \cos \vartheta d_u \]  

where \(d_e, d_n, d_u\) are displacements in east, north and up directions respectively, \(\alpha\) is the azimuth of the satellite heading (positive clockwise from the north) and \(\vartheta\) is the radar incidence angle.

In this study, we assume that the horizontal components are negligible. Assuming this, the vertical component can then be converted from \(d_{\text{los}}\) using the local incidence angle: \(d_u = -d_{\text{los}}/\cos \vartheta\). The LOS mean velocity and LOS time series were both converted into vertical direction measurements (Figure 5a). Eleven points at the same locations as the eleven WS stations of the dam were used to investigate the accuracy of the InSAR measurements. These eleven locations were carefully selected by cross-reference of the optical image, SAR amplitude image, and the ground monitoring locations. Finally, a high correlation of 0.9 and an RMS of 1.5 cm/year provided confidence that InSAR can produce reliable monitoring results (Figure 5b).
4. Mechanical Parameter Back-Analysis Using the FEM and InSAR Results

4.1. Mechanical Parameters Back-Analysis Method

Rockfill model parameters back-analysis is used to find one group of parameters that make the calculated and observed displacement correspond. This process includes the following steps (shown in Figure 6):

1. Build the objective function using the InSAR results, as Equation (2). Equation (2) presents the InSAR results of \( i \)th monitoring point.
2. Perform the strain-stress analysis using the FEM (see Section 4.1.1) and the calculation of the RSM as defined in Section 4.1.2. Using the RSM, simulate the relationship between the parameter set and the displacement of each monitoring point.
3. Find the optimal parameter set of the objective function using the modified genetic algorithm (GA) introduced in Section 4.1.3. The optimal parameter set is a combination of the material parameters that minimize the objective function. During the process of searching, these RSM are used to replace the FEM to calculate the fitness of all the measurement points.

\[
f(\lambda_1, \lambda_2, \ldots, \lambda_n) = \left[ \frac{1}{n} \sum_{i=1}^{n} \left( \frac{u_i - u_i^*}{u_i^*} \right)^2 \right]^{0.5}
\]  

where \((\lambda_1, \lambda_2, \ldots, \lambda_n)\) is a group of constitutive parameters to be identified, \(u_i\) is the calculated displacement at the monitoring point \(i\), \(u_i^*\) is the corresponding measured displacement and \(n\) is the number of monitoring points used in the back-analysis.

4.1.1. The FEM Model and the Constitutive Model

To further extend the investigation of the settlement determined by the InSAR analysis, we used the finite element method to conduct mechanical parameter back-analysis and strain-stress analysis, which allowed us to account for all geological and geophysical information available for the considered area. As a numerical technique, the FEM is used to find approximate solutions to boundary value problems consisting of partial differential equations and boundary conditions. The FEM subdivides a large problem into smaller, simpler, parts, called finite elements. The simple equations that model these finite elements are then assembled into a larger system of equations that models the entire problem. Then, the FEM uses variational methods from the calculus of variations to approximate a solution by minimizing an associated error function. The principle of the variational methods has
been proven to be the equivalent integral of the differential equations and boundary conditions in mathematics. If the finite element conforms to the convergence criteria, the final approximate solution of the FEM converges to the exact solution of the original mathematical model [32]. The Duncan EB model is the most widely used nonlinear elastic constitutive model in the FEM analysis of soil structure [33,34]. In particular, we analysed the stress–strain characteristics of the dam in a 3-D FEM model to solve for the retrieved displacements (Figure 7). We defined the subdomain setting of the FEM model using the available geological and structural information. The process and the reservoir water filling were simulated in the analysis (Figure 8).

**Figure 7.** The 3D FEM mesh of the Shuibuya CFRD: (a) 3D FEM mesh, 1-1 is the biggest cross section of the SBY dam (0 + 212 m); and (b) 2D FEM mesh of the cross section of the 1-1 (0 + 212 m).

**Figure 8.** The process of construction and reservoir water filling. The numbers indicate the step of the dam construction and reservoir filling.

Due to its simplicity and the explicit physical meaning of its input parameters, the constitutive model proposed by Duncan has been widely used in civil engineering, especially for rockfill dams in China [33,34]. It can be described as follows:
where $E_t$ is the tangent modulus, $K$ is the modulus number, $p_a$ is the atmospheric pressure, and $\sigma_3$ is the minor principal stress. $n$ and $R_t$ are the exponent and failure ratio, respectively. $\varphi = \varphi_0 - \Delta \varphi \lg(\sigma_3/p_a)$ is the internal friction angle, and $c$ is the cohesive strength. $\sigma_1$ and $\sigma_3$ are the major and minor principal stresses, respectively. $B_i$ is the bulk modulus. The rockfill of the dam has obvious creep characteristics. In this paper, the Merchant five-parameter creep model is used to describe the creep behaviour of rockfill materials [35].

$$
\epsilon_t = \epsilon_t(1 - e^{-at}), \quad \epsilon_{df} = d \left( \frac{S_L}{1 - S_L} \right)
$$

where $\epsilon_t$ is the creep strain, $\epsilon_t$ is the limit of the creep strain, $a$ is a coefficient related to the creep strain rate, $\epsilon_{df}$ is the limit of shear creep strain, and $d, b, m_c$ and $\beta$ are the creep parameters.

### 4.1.2. The Response Surface Method

Because of the complexity of engineering problems, the implementation of back-analysis is often time-consuming for a large number of the finite element analyses. In this study, the RSM with strong nonlinear mapping ability, is used to simulate the relationship between the parameter sets and the displacement of the monitoring points [4].

Twenty-seven parameter sets are generated using the orthogonal algorithms. Then, the FEM program that takes the static and creep properties of material into account is adopted to calculate the displacement at different monitoring points. The process of construction and reservoir water filling are simulated in the FEM analysis as show in Figure 8. In Li’s [4] study, the RSM considers only the static displacement of the rockfill. In this paper, the creep displacement is also taken into consideration. The RSM is given as follows:

$$
S_L(\bar{x}) = a + \sum_{i=1}^{N} b_i \bar{x}_i + \sum_{i=1}^{N} c_i \bar{x}_i^2 + \sum_{j=1}^{M} d_j \bar{y}_j + \sum_{j=1}^{M} f_j e^{\bar{\varphi}_j}
$$

where $\bar{x}_i$ and $\bar{y}_j$ are the static and creep parameters to be identified, respectively. $N$ and $M$ are the numbers of the static and creep parameters, respectively $a, b_i, c_i, d_j$ and $f_j$ are the coefficients of the RSM to be identified. $\bar{x}_i$ and $\bar{y}_j$ can be calculated as follows:

$$
\bar{x} = \left\{ \frac{\varphi_0}{\varphi_0}, \Delta \varphi, K, \pi, R_t, K_b \right\}^T
$$

$$
\bar{\varphi}_0 = \frac{\varphi_{01}}{\varphi_0}; \quad \Delta \varphi = \frac{\Delta \varphi}{K_1}; \quad K = \frac{k_1}{K_1};
$$

$$
\pi = \frac{n_i}{m_i}, \quad R_t = \frac{R_t}{K_1}, \quad K_b = \frac{K_b}{K_1}
$$

$$
\bar{y}_j = \left\{ \bar{x}_i, \bar{F}, m_c, d_i, d_j \right\}^T
$$

$$
\bar{\pi} = \frac{a}{a_1}, \quad \bar{F} = \frac{b}{b_1}, \quad \bar{m_c} = \frac{m_c}{m_{c1}}, \quad \bar{d_i} = \frac{d_i}{d_{i1}}
$$

$$
\bar{b} = \frac{\beta}{\beta_1}, \quad \bar{d} = \frac{d}{d_1}
$$

where $\varphi_{01}, \Delta \varphi, k_1, n_1, R_t$ and $K_{b1}$ are the values of the Duncan EB model parameters obtained from the laboratory; $\varphi_0, \Delta \varphi, k, n, R_t$ and $K_b$ are the values after zooming in; $a_1, b_1, m_{c1}, \beta_1$ and $d_{11}$ are the values of the Merchant creep model parameters obtained from the laboratory; and $a, b, m_c, \beta$ and $d$ are the values after zooming in.
Then, the results of strain-stress analysis are used to calculate the coefficients of the RSM of every monitoring point using Equation (6). The RSM can simulate the relationship between the parameter set and the displacement of every monitoring point.

\[
\begin{align*}
S_1^1(x) &= S(x_1, x_2, \ldots, x_M) \\
S_2^1(x) &= S(x_1 + dx_1, x_2, \ldots, x_M) \\
S_3^1(x) &= S(x_1 - dx_1, x_2, \ldots, x_M) \\
& \vdots \\
S_{26}^1(x) &= S(x_1, x_2, \ldots, x_M + dx_M) \\
S_{27}^1(x) &= S(x_1, x_2, \ldots, x_M - dx_M)
\end{align*}
\]

(8)

4.1.3. The Modified Genetic Algorithm

The parameter back-analysis of rockfill dams can be abstracted as an optimization problem in mathematics. The GAs proposed by Holland in 1975 [36] have proven to be powerful for solving optimization problems. In this paper, a modified GA with global convergence is used to find the optimized parameter set of the objective function.

In GAs, the diversity of the population and the disproportionation between exploitation and exploration is conditioned by the crossover operator, which also influences the speed of convergence and determines the global convergence of the algorithm. The crossover operator is a basic operation with great importance. However, in the crossover operation of the traditional GA, the crossover points and the gene fragments are selected with randomness and blindness. It is difficult to generate new individuals when the selected gene fragments are highly similar to each other, that is, traditional crossover is not very effective in producing new generations that have great differences with their parents. This crossover operation is invalid and may decrease the diversity of the population.

Inspired by genetic engineering and the cloning of superior genes, a novel genetic crossover operator based on the sum of differences in gene fragments (SoDX) is proposed. The fragments’ crossing probabilities are first evaluated based on the differences. The gene fragments are then selected and exchanged according to the crossing probability. This process can reduce inbreeding and the possibility of invalid crossover operations. A new GA code named SoDX-NUM was implemented in MATLAB by combining the crossover operator SoDX and the non-uniform mutation (NUM).

4.2. The Back-Analysis Results and the Argumentation

In this paper, six monitoring points (WS06–WS11) from levelling and InSAR with high precision and large displacement are used to invert the parameters. Using the displacement back-analysis method proposed in this paper, the identified model parameters set of rockfill based on the levelling and InSAR is determined, as shown in Table 1. Identified 1# and Identified 2# present the back-analysis results based on the levelling and InSAR monitoring results, respectively. \(K_1, n_1, K_{b1}\) and \(m_1\) are the parameters of the Duncan EB model in the first material zone of the rockfill dam. \(K_2, n_2, K_{b2}\) and \(m_2\) are the parameters of the Duncan EB model in the second material zone of the rockfill dam. \(a, b, mc, \beta\) and \(d\) are the parameters of the creep constitutive model of the rockfill.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>(K_1)</th>
<th>(n_1)</th>
<th>(K_{b1})</th>
<th>(m_1)</th>
<th>(K_2)</th>
<th>(n_2)</th>
<th>(K_{b2})</th>
<th>(m_2)</th>
<th>(a)</th>
<th>(b) (%)</th>
<th>(mc)</th>
<th>(\beta) (%)</th>
<th>(d) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>1100</td>
<td>0.35</td>
<td>600</td>
<td>0.1</td>
<td>850</td>
<td>0.25</td>
<td>400</td>
<td>0.05</td>
<td>0.009</td>
<td>0.0098</td>
<td>1.1338</td>
<td>0.64</td>
<td>0.21</td>
</tr>
<tr>
<td>Identified 1#</td>
<td>876</td>
<td>0.264</td>
<td>512</td>
<td>0.115</td>
<td>967</td>
<td>0.203</td>
<td>320</td>
<td>0.0579</td>
<td>0.0069</td>
<td>0.0113</td>
<td>1.334</td>
<td>0.77</td>
<td>0.168</td>
</tr>
<tr>
<td>Identified 2#</td>
<td>958</td>
<td>0.31</td>
<td>550</td>
<td>0.1</td>
<td>816</td>
<td>0.225</td>
<td>375</td>
<td>0.05</td>
<td>0.0007</td>
<td>0.0088</td>
<td>0.907</td>
<td>0.512</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Identified 1# and Identified 2# are the back-analysis results of the levelling and InSAR, respectively.
Using the identified parameter sets of Identified 1# and Identified 2#, the FEM analysis is implemented. In the analysis, the process of dam construction and the water filling are simulated. The results are shown in the Figure 9, including the settlement process of the levelling, the identified parameter set based on the levelling, and the identified parameter set based on InSAR.

![Figure 9](image-url)

**Figure 9.** The displacement for WS05–WS11 based on the levelling, levelling-identified, InSAR and InSAR-identified parameters.

The traditional back-analysis methods use internal monitoring data. In this study, the external monitoring results are used as the objective function to perform the back-analysis. The correlation of the levelling monitoring data and the computed results based on Identified 1# of the six points are 0.97, 0.96, 0.95, 0.96, 0.93 and 0.96. This suggests that the displacement pattern of the calculated results using the identified parameters set of levelling agree well with the observed levelling values, which indicates that the identified parameters, in general, satisfactorily reflect the displacement properties of the dam. The external monitoring data could be used to map the displacement properties of the dam. The InSAR results are also used to perform the parameters back-analysis. Figure 9 shows the computed settlement based on Identified 2# and the InSAR monitoring results. They have similar magnitude and distribution.

The back-analysis results based on levelling data and InSAR monitoring data are also compared. The displacement from 28 February 2007 to 11 March 2011 is shown in Figure 10a,b and the higher elevation, the greater the displacement increment. In addition, the differences and the correlation are shown in Figure 10c,d. The computed displacement of the two identified parameter sets have little
difference with each other (Figure 10c). The correlation of the displacement results generated from InSAR and levelling is 0.99 (Figure 10d). The displacement in the middle of the dam is greater than at the two sides. The greatest total displacement increment of the FEM computed result is 0.467 m, which is similar to the InSAR results. It also shows that the computed results and the results of InSAR have similar displacement increment distributions, and the displacement computed using the parameters Identified 2# has similar magnitude and distribution as the InSAR monitoring results.

Combining the time series analysis and back-analysis, it can be concluded that InSAR technology can be used as a supplement for the traditional monitoring method to monitor high dams and can also be used as a method to invert the parameters of the dam.

Figure 10. Comparison of the displacement increment from 28 February 2007 to 11 March 2011: (a) Vertical displacement computed by the FEM model and the identified parameters based on levelling (1#); (b) Vertical displacement computed by FEM model and the identified parameters based on the InSAR results (2#); (c) The displacement difference between (a) and (b); (d) The Correlation between InSAR and levelling measurements, and black line in is a 1:1 line. XCoor means the direction of the river flow, YCoor means the direction from the left bank to the right bank. The units of X and Y-axes are in meters.
4.3. Settlement Prediction for the SBY Dam

The FEM model, the constitutive model and the back-analysis results can accurately represent the time-dependent displacement of the SBY dam. To study the post-construction settlement of the SBY dam, further FEM analysis is performed to predict the dam displacement using the parameters identified based on the InSAR results. The settlement contours of cross-section 0 + 212 m for 30 December 2020 are given in Figure 11. The maximum settlement is 2.332 m, approximately 1% of the dam height. The maximum settlement occurs in the centre of the dam body, whereas the settlement of upstream (left) and downstream (right) of the dam is smaller (<40 cm), and the settlement increases towards the centre. The biggest settlement located at the core of the dam mainly because of following two reasons: (1) Because of the gravity of rockfill in different layers and intervals between the different stages of concrete-face construction, most settlement of the lower layer filled rockfill is happened when the upper layer rockfill is completed on the dam. The SBY dam has different material zones, which are shown in Figure 2a). The rockfill is gradually filled according to layer during the construction period. For example, the first stage of concrete-face is constructed after the rockfill of the dam was decreased six or nine months, then the second stage concrete-face and the third concrete face. (2) Because of arching effects caused by different settlement rate of the dam body in the parallel and perpendicular direction of the river. In the direction of perpendicular to the river, due to the V-shaped valley, the arching effect is raised when the settlement rate of the central rockfill is faster than the two sides of the rockfill, which contributes an uplift force to the central rockfill. While in the direction of parallel to the river flow, the other arching effect is caused because of the settlement rate in the central is faster than the upstream and downstream of the dam, which also produces an uplift force to the central rockfill. Therefore, the maximum displacement occurs in the centre of the dam.

The settlement increments from 28 February 2007 to 11 March 2011 and to 30 December 2020 are given in Figure 12a,b, respectively, and their corresponding difference is shown in Figure 12c. Figure 12a,b shows that the displacement increment is approximately 60 cm at the dam crest, and the upstream is greater than the downstream. From the Figure 12c, it can be observed that the settlement at the top and upstream (left) is much greater than that downstream (right) from the dam. This difference is due to water filling; the load of the upstream (left) of the dam is greater than the downstream load (right). The settlement is smaller than 10 cm from 11 March 2011 to 30 December 2020 and is approximately 60 cm from 2007 to 2011. This suggests that the displacement increases rapidly in the early reservoir period, and the settlement trends toward stability in the later operation period. The displacement consists of instantaneous and creep displacement. Most of the later settlement is creep displacement, and the basic characteristic of creep displacement is that it decreases gradually with the time. These results agree with the displacement characteristics of the dam.

Figure 11. The settlement contours of cross-section 0 + 212 m for 30 December 2020. XCoor means the direction of the river flow. The units of X and Y-axes are in meters.
5. Discussions

In our previous study, the InSAR technique was used to investigate the displacement of the SBY Dam, and the reliability was compared with levelling measurements. The results provide confidence that the InSAR time series is a useful tool for CFRD displacement monitoring [16]. In this study, to further investigate the inside features of the SBY dam, the FEM, RSM and GA are used to identify the mechanical parameters of the rockfill, and then the associated displacement time series and settlement is simulated. The Duncan EB model is the most widely used nonlinear elastic constitutive model in the numerical analysis of soil structure. The triaxial shear tests show that the Duncan EB model is capable of simulating the stress–strain relationship of rockfill materials. The computed results of the previous rockfill dams and the in situ monitoring data have proven the precision of the Duncan EB model [4,9,33].

To examine the stability and security of the CFRD, the physical parameters of the rockfill are pre-requisite for settlement prediction. Due to the size effect, sample disturbance and the impact of random sampling, it is difficult to guarantee the represent activeness of the experiment results. Parameter back-analysis is an efficient way to identify parameters. In this paper, the levelling and InSAR measurement results are used to perform the parameter back-analysis. The results show that the trend, the magnitude and the distribution of the displacement of the SBY Dam downstream surface
agree well with each other (Figures 9 and 10). The high correlation of 0.995 and a low RMSE of 0.055 m between the results based on InSAR and the results based on the levelling measurements suggest that the back-analysis based on the InSAR technique is a reliable way to identify the parameters of a rockfill dam. The low RMSE indicates that the back-analysis method is accurate and useful for obtaining the mechanical parameters of the dam. The internal ground monitoring data continues to 16 March 2012. The maximum settlement from the ground monitoring data is 2.536 m, and it is 2.156 m based on the rockfill parameters from the InSAR results, which confirms that the combination of InSAR measurement and the FEM back-analysis is a reliable way to identify the parameters and to predict the displacement of rockfill dams. The results also suggest the stability and security of the SBY CFRD between 2007 and 2020.

The density and coverage of the InSAR measurement can be much higher than conventional survey methods, such as levelling and GPS. In addition to the displacement of the dam body, the stability of the surroundings of the dam can be determined by the InSAR method. This is important because the geologic stability of a dam’s surroundings has a great impact on its behaviour. Moreover, many high CFRDs will be built in the northwest of China, where earthquake frequently occur; thus, accurate monitoring is a guarantee of their security. Compared to conventional measurements, the InSAR is an efficient supplement technology to the traditional methods in the long term.

6. Conclusions

The physical parameters of the SBY dam are identified using the proposed back-analysis method by combining the FEM model and the InSAR time series results. The displacement results using the identified parameters agree well with the observed levelling values in both magnitude and distribution. The settlement prediction of the SBY Dam is performed using the identified parameters, and the maximum settlement is 2.332 m, approximately 1% of the dam height. The settlement of the SBY Dam becomes stable in the later operation period. The back-analysis method using the external monitoring data from InSAR can record the displacement properties of the dam. The InSAR technology can be used as supplement for the traditional monitoring methods to monitor high dams and can also be used to invert the parameters of the dam.

Acknowledgments: This research was supported by the National Natural Science Foundation of China (No. 1322905) and China Postdoctoral Science Funding (2015M570666). We also acknowledge Zhenhong Li for providing the InSAR + AEM InSAR time series package.

Author Contributions: Wei Zhou, Shaolin Li, Zhiwei Zhou, and Xiaolin Chang developed the main idea of this study, and Wei Zhou guided the project. Shaolin Li and Wei Zhou contributed to the ground monitoring data analysis, parameters back-analysis and FEM analysis. Zhiwei Zhou performed the ALOS PALSAR data interferometry processing and analysis. Zhiwei Zhou, Shaolin Li and Wei Zhou contributed to discussion of the final results and writing the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References


29. Li, Z.; Fielding, E.J.; Cross, P.; Muller, J.-P. Interferometric synthetic aperture radar atmospheric correction: GPS topography-dependent turbulence model. J. Geophys. Res. Solid Earth 2006, 111, B02404. [CrossRef]


31. Hammond, W.C.; Blewitt, G.; Li, Z.; Plag, H.P.; Kreemer, C. Contemporary uplift of the Sierra Nevada, western United States, from GPS and InSAR measurements. Geology 2012, 40, 667–670. [CrossRef]


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
Chapter 2:
New EO Techniques and Services
Article

Precise Positioning of BDS, BDS/GPS: Implications for Tsunami Early Warning in South China Sea

Kejie Chen *, Natalia Zamora, Andrey Y. Babeyko, Xingxing Li and Maorong Ge

GFZ German Research Centre for Geoscience, Telegrafenberg, 14473 Potsdam, Germany; zamora@gfz-potsdam.de (N.Z.); babeyko@gfz-potsdam.de (A.Y.B.); lixin@gfz-potsdam.de (X.L.); maor@gfz-potsdam.de (M.G.)

* Correspondence: kejie@gfz-potsdam.de; Tel.: +49-331-2882-8792

Academic Editors: Zhenhong Li, Roberto Tomas, Xiaofeng Li and Prasad S. Thenkabail

Received: 30 September 2015; Accepted: 23 November 2015; Published: 30 November 2015

Abstract: Global Positioning System (GPS) has been proved to be a powerful tool for measuring co-seismic ground displacements with an application to seismic source inversion. Whereas most of the tsunamis are triggered by large earthquakes, GPS can contribute to the tsunami early warning system (TEWS) by helping to obtain tsunami source parameters in near real-time. Toward the end of 2012, the second phase of the BeiDou Navigation Satellite System (BDS) constellation was accomplished, and BDS has been providing regional positioning service since then. Numerical results indicate that precision of BDS nowadays is equivalent to that of the GPS. Compared with a single Global Satellite Navigation System (GNSS), combined BDS/GPS real-time processing can improve accuracy and especially reliability of retrieved co-seismic displacements. In the present study, we investigate the potential of BDS to serve for the early warning system of tsunamis in the South China Sea region. To facilitate early warnings of tsunamis and forecasting capabilities in this region, we propose to distribute an array of BDS-stations along the Luzon Island (Philippines). By simulating an earthquake with Mw = 8 at the Manila trench as an example, we demonstrate that such an array will be able to detect earthquake parameters in real time with a high degree of accuracy and, hence, contribute to the fast and reliable tsunami early warning system in this region.

Keywords: BDS; BDS/GPS; South China Sea; tsunami early warning system

1. Introduction

As one of the most devastating natural coastal disasters, tsunamis are triggered mostly by shallow earthquakes in submarine subduction zones [1] producing the most damage in the near field but also propagating basin-wide. Consequently, traditional Tsunami Early Warning Systems (TEWS) mainly rely on seismic methods in order to evaluate source parameters (magnitude, epicenter) and forecast tsunamis as soon as possible. Corresponding seismic instrumentation mainly includes broadband seismometers as well as strong-motion sensors (accelerometers). Despite the fact that these traditional techniques have proved their reliability for most historical cases over the last decades, they have failed to provide correct rapid magnitude estimation for several important events like the 2004 Mw = 9.3 Great Sumatra earthquake, the 2010 Mw = 7.8 Mentawai tsunami earthquake and the 2011 Mw = 9.0 Great Tohoku event. For example, in the latter case, the true magnitude was significantly underestimated during the fifty minutes after the earthquake [2]. Magnitude underestimation, in turn, may result in the underestimation of the tsunami forecast in the near field [2].

The reasons for magnitude underestimation during the first several minutes after large earthquakes vary for different seismic instruments as well as earthquake evaluation procedures. Broadband seismometers saturate near the source in case of strong shaking so that the full rupture image can be analyzed only at larger epicentral distances, causing unavoidable delay due to the finite
speed of seismic wave propagation. Strong motion sensors, in turn, do not saturate by amplitude and can be employed close to the source; however, the usual band-pass filtering of waveforms aimed to avoid processing unambiguity caused by co-seismic tilting [3] effectively leads to magnitude saturation due to removal of long periods which are essential for huge earthquakes [4]. See, for example, [2,5] regarding the analysis of the strong-motion magnitude saturation during the Great Tohoku earthquake.

The 2004 Great Mw = 9.3 Sumatra earthquake and tsunami boosted the development of new seismic approaches that are aimed toward the faster and more reliable characterization of tsunamigenic earthquakes. During the GITEWS Project (German-Indonesian Tsunami Early Warning System), Bormann and Saul [6] proposed a fast, non-saturating cumulative magnitude estimator based on the P-wave train. The procedure was tested at epicentral distances starting from 5° and is currently operated by the INATEWS (Indonesian Tsunami Early Warning System) [7]. Additionally, Kanamori and Rivera [8] suggested a new broadband W-phase source inversion algorithm, which is currently running in operation at PTWC (Pacific Tsunami Warning Center). During the 2011 Great Mw = 9.0 Tohoku event, this algorithm automatically detected the focal mechanism and came with magnitude estimate of Mw = 8.8 22 min after the earthquake (the true magnitude Mw = 9.0 was estimated 40 min after the earthquake) [9]. Another approach was suggested by Lomax and Michelini [10], who introduced an original duration-amplitude procedure to estimate the tsunamigenic potential of earthquakes.

Significant efforts were also undertaken to improve the quality of real-time strong-motion data processing. New studies [4,11,12] proposed fast and effective procedures for automatic base-line correction. Furthermore, the novel technique of collocation of strong-motion and GPS sensors [13,14] should help to overcome the problem of magnitude saturation since the band-pass filtering is no longer required [15].

Rapid improvement of the real-time GPS processing precision put GPS-technology at the front end of current progress in the earthquake and tsunami early warning system. Compared with seismic approaches, continuous GNSS real-time data processing provides arbitrary ground displacements that can be directly inverted into source parameters in (near-) real time. Hence, it is considered to be a more trustworthy tool, especially for the near field tsunami early warning. Several previous studies have discussed or demonstrated the potential of using GPS data for tsunami early warnings. For example, Blewitt et al. [16] showed that the magnitude, mechanism, and spatial extent of rupturing of the 26 December 2004 Sumatra earthquake might have been accurately determined using only 15 min of GPS data following an earthquake initiation. Simultaneously, Sobolev et al. [17] pointed out that the reliable prediction of tsunami waves on the Indonesian coast can be issued within less than 5 min of an earthquake by incorporating special types of near-field GPS arrays (“GPS-Shield” concept for Indonesia). They also proposed deployment of such arrays for other tsunamigenic active regions. Following this idea, GITEWS (German-Indonesian Tsunami Early Warning System) became the first TEWS to implement real-time GPS for a tsunami early warning [18,19]. Recently, Hoechner et al. [20] replayed the Great Tohoku 2011 event in that they presented a complete processing chain starting from actual raw GPS data and fully simulated the situation as it would be in a warning center ending up with a very fast and qualified tsunami early warning. Now using real-time GPS for earthquake and tsunami early warnings is an active area of research (e.g., [18–26]).

To date, geo-hazard monitoring with GNSS is restricted to the oldest, largest and most-used GPS system. Around the end of 2012, China has completed the regional constellation of the BeiDou Satellite System (BDS), and the system has been running routinely since then. Even at its second-phase stage, numerous studies (e.g., [27–29]) have demonstrated that precise positioning performance of BDS is equal in match against that of GPS in the Asia-Pacific region. Compared with a single satellite navigation system, the fusion of GPS, GLONASS, GALILEO and BDS can significantly increase the number of observed satellites, reduce the position dilution of precision (PDOP) [29] and, thus, increase robustness and the reliability of observations. What is more, BDS, due to its similarity to GPS signal structure and frequencies, provides a much better chance for investigating the impact of multi-GNSS
capacity on real-time precise positioning in terms of accuracy and reliability. In the present study, we investigate the potential of the BDS system to contribute to the tsunami early warning system in the South China Sea region. In particular, we assess the accuracy of the real-time BDS processing and corresponding source inversion due to a hypothetical tsunamigenic earthquake at the Manila trench.

As shown in previous studies (e.g., [30–33]), the Manila subduction zone (Figure 1) has been identified as a zone of potential tsunami hazard in the SCS region. Despite the absence of clear historical evidence on large tsunamis generated at the Manila trench, tsunami deposits found at the Xischa Islands (~1024 AD) may be attributed to a large event in this source region [34]. According to the recently compiled historical tsunami database for the northeastern South China Sea [35], annual probability of a tsunami in this region from any source is very high (~33%). However, the likelihood of a damaging tsunami per year is much smaller (1%–2%). In the absence of large historical events from the Manila Trench, damaging tsunamis in the South China Sea have been previously studied by means of numerical modeling. Simulations [30,36] suggest that an earthquake with a magnitude of Mw = 8.0 or larger at the Manila trench may cause significant tsunami damage along the whole south-eastern coast of China, Southwestern Taiwan and West Philippines. Taking the potential large-scale disaster into account, the tsunami early warning system in the SCS region was suggested by [30,37,38] based on DART-type deep ocean buoys. In fact, China has installed two buoys in SCS in 2014. In the present paper, we propose to improve the regional TEWS by installing real-time continuous BDS-network at the Luzon Island, Philippines. Being integrated with traditional seismic networks, the proposed GNSS-array will contribute to fast and reliable source inversion, which is one of the most important tasks in the early warning and forecasting of tsunamis.

Figure 1. Seismotectonic map of the Southern China Sea. The red thick line represents the Manila trench of the seismogenic Manila subduction zone. The red points show the earthquakes that have occurred from 1976 to 2015 with a magnitude larger than Mw = 6.5 (GCMT catalog). The yellow line E2 indicates the rupture area used in this study. The two multi-GNSS stations used in Section 2 are also shown by green triangles.

In Section 2, we test the real-time precise positioning performance of BDS and compare it with GPS and joint BDS/GPS in the SCS region. In Section 3, we simulate a test scenario of an Mw = 8.0 earthquake along the Manila trench and its corresponding tsunami to demonstrate the feasibility of
incorporating BDS, BDS/GPS into the regional TEWS. In Section 4, we discuss additional issues related to the performance of BDS, BDS/GPS based TEWS. Finally, Section 5 summarizes results and presents an outlook.

2. Real-Time Kinematic Precise Positioning Performance of BDS in South-East Asia

Generally, there are two kinds of precise positioning algorithms, i.e., relative positioning and single-point positioning. With respect to relative positioning, most of the error budgets can be cancelled through double differencing between reference station and rover station; as a result, its mathematical model is relatively simple and easy to be conducted. Most importantly, double-differencing integer-cycle phase ambiguities can be resolved to their correct integer values, reliably ensuring its high precision. However, the reference station of relative positioning must be fixed, which requires that, during an earthquake, it should be located outside the zone of deformation. The latter cannot always be satisfied, especially during large tsunamiogenic earthquakes [21]. By contrast, the single-point positioning approach does not need a fixed reference station to form double-difference observations and, in theory, is more desirable for the retrieval of co-seismic displacements [25]. Nonetheless, it should be pointed out that for Precise Point Positioning (PPP), ambiguity resolution needs a convergence phase that may be as long as 20 min. Fortunately, the Variometric Approach for Displacements Analysis Stand-alone Engine (VADASE) [39] and the Temporal Point Positioning (TPP) [40] are able to overcome this disadvantage. The TPP has been employed in this study.

Up to now, no BDS or BDS/GPS data recorded during an earthquake is publicly available. Considering that performances of GNSS positioning are not different between earthquake-free and earthquake-breaking periods [13], to exploit the potential of BDS and BDS/GPS in retrieving co-seismic displacements, we selected two weeks of data (from 29 December 2013 to 11 January 2014) from IGS Multi-GNSS Experiment (MGEX) stations in the Asia-Pacific region. In this case, data from a 1-Hz sampling rate MGEX stations GMSD and SIN1 were used (see station distribution in Figure 1). Since no earthquake happened, we take zero displacement as ground truth. In this paper, real-time precise BDS/GPS orbits and clocks were produced according to the strategies suggested by Li et al. [29]. For every station, the length of each TPP session was set to 10 min, and, in each individual hour, we employed four sessions with evenly distributed start points; in total, we have 96 sessions per day for two weeks, which ensures a more trustworthy conclusion on the precise positioning performance. Figure 2 presents the TPP results in modes of BDS-only, GPS-only and BDS/GPS at station GMSD which can be taken as a typical example. As clearly shown, BDS-only and GPS-only show similar positioning performances while a BDS/GPS combined solution significantly improves the accuracy. Statistical Root Mean Square (RMS) is summarized in Table 1.

Table 1. Expected RMS of BDS, GPS, and joint BDS/GPS real-time precise positioning using TPP in South-East Asia.

<table>
<thead>
<tr>
<th></th>
<th>BDS</th>
<th>GPS</th>
<th>BDS/GPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>East</td>
<td>0.0095</td>
<td>0.0071</td>
<td>0.0049</td>
</tr>
<tr>
<td>North</td>
<td>0.0068</td>
<td>0.0101</td>
<td>0.0048</td>
</tr>
<tr>
<td>Up</td>
<td>0.0203</td>
<td>0.0207</td>
<td>0.0098</td>
</tr>
</tbody>
</table>
3. Testing the Feasibility of BDS Real-Time Network at the Luzon Island for the Tsunami Early Warning in the South China Sea

In this section, we demonstrate the feasibility of BDS for the tsunami early warning in the Southern China Sea. To do that, we simulate a hypothetical tsunamigenic earthquake along the Manila trench and try to invert resulting co-seismic displacements at a virtual BDS/GPS array distributed along the Luzon Island, northern Philippines. Previously, Kirby and Geist [41] and Liu et al. [30] assessed tsunamigenic potential of six rupture scenarios along the Manila trench. We follow their approach and place our fault model along the northern part of the Manila subduction zone (Figure 1). Fault parameters correspond to the rupture model E2 from Liu et al. (2006). This rupture scenario was selected due to its highest impact potential for the south Chinese coast. The accepted source model has a magnitude of $Mw = 8.0$ with the central point located at $119.8^\circ$E/$18.7^\circ$N. The length is 180 km, the width 35 km, and the strike and the dip angles of the rupture plane are fixed to 35 and 20 degrees. The only difference with the original E2 rupture model is in slip distribution: we use a bell-shaped distribution rather than a uniform slip. The rupture plane consists of 75 sub-faults with a symmetry maximum at the slip center, and with the rake angles varying between 80 and 100 degrees.

We use code QSSP developed by Wang [42] to calculate synthetic displacement waveforms at stations of our virtual GNSS-array. With respect to the Earth model, in this study, the average global 1-D reference earth model AK135Q [43] is adopted. Figure 3 demonstrates simulated displacement waveforms at the selected station B1 (see virtual array in Figure 4a). To make our test even more realistic, we contaminate synthetic co-seismic displacements with a typical real-time BDS and BDS/GPS processing noises derived from a random segment of the GMSD (or SIN1) residual displacement time series, and the corresponding “noisy” displacements are also shown at Figure 3. These “noisy” displacements will be used for source inversion.

Figure 4a presents an overview of our forward rupture model: It shows assumed slip distribution together with the corresponding co-seismic vertical deformation as well as horizontal static offsets at the virtual BDS/GPS network.
Figure 3. Simulated displacement waveforms in forward-model scenario (rupture E2, see text for description). Waveforms are shown at one selected virtual GNSS-station placed at the Luzon Island (contoured triangle on Figure 4a). Black lines represent original kinematic simulations; colored lines represent the simulations after an addition of typical TPP processing noise (Section 2). "Noisy" displacements will be used for source inversion.

Figure 4. Cont.
Figure 4. (a) Assumed slip distribution (colored dots) and correspondent co-seismic surface deformation for an event with a magnitude of Mw = 8.0 rupturing along the Northern Manila mega-thrust: our forward-model scenario. Red arrows show horizontal displacements computed at the virtual BDS/GPS network (black triangles). The yellow triangle (B1) marks the station from Figure 3. (b) Colored dots represent slip distribution as inverted from the simulated BDS-displacements (note, synthetic BDS-displacements include real-time processing noise). Additionally shown is the resulting vertical deformation; (c) Same as (b), but for the inversion of the joint BDS/GPS displacements.

Based on the suggested rupture, the corresponding tsunami scenario was simulated (Figure 5) using the easyWave wave propagation code [20]. easyWave follows the numerical algorithm by Goto et al. [44] for the simulation of linear long-wave propagation in spherical coordinates. Tsunami propagation was computed on the 1 arc minute GEBCO (General Bathymetric Chart of the Oceans) bathymetric grid [45], and the sea-level heights at offshore positions were projected to the coast using Green’s amplification law [46]. As expected, due to the predominantly SW-NE orientation...
of the rupture, main tsunami energy is radiated towards the southeastern coast of China, between the cities of Hong Kong and Shantou (Figure 5). Expected maximum tsunami wave heights may reach up to 10 m at the west coast of the Luzon Island. At the southwest coast of Taiwan and southeastern coast of China, tsunami wave heights may reach 1–2 m.

In order to issue a reliable tsunami early warning as soon as possible, seismic parameters (e.g., epicenter, magnitude, slip distribution) should be inverted from co-seismic signal immediately after the earthquake breaks. Since the Luzon Island is located close to the northern Manila trench, our scenario rupture causes significant horizontal co-seismic displacements throughout the island. As shown in Figure 4a, stations distributed along the northwest part of the Luzon Island will experience northwest-directed surface displacements in excess of 10 cm. Taking into account the horizontal accuracy of real-time BDS and GPS processing of about 1 cm (Table 1), we conclude that co-seismic displacements due to an earthquake with a size of Mw = 8.0 should be well-detected and accurately measured with the present-day BDS/GPS observation and processing precision, which, in turn, should enable reliable source inversion and tsunami forecasting.

![Simulated tsunami scenarios](image)

**Figure 5.** Simulated tsunami scenarios. The color map shows the maximum wave height in the forward model (Figure 4a) after 6 h of tsunami propagation. Vertical bars are maximum wave heights as recorded at selected coastal locations for the three propagation models: forward model (black); with source as inverted from BDS-displacements (dark grey in Figure 4b); with source as inverted from joint BDS/GPS-displacements (light grey, Figure 4c).

To illustrate this, we retrieved static offsets from our simulated kinematic BDS and BDS/GPS displacements (Figure 3) and invert them back into the source parameters (Figure 4b,c). Finally, we compute tsunami propagations from the “inverted sources” and compare them to the forward scenario (vertical bars on Figure 5). In order to get static offsets at virtual BDS/GPS stations from the contaminated displacement waveforms (Figure 3), we employ a method introduced by Ohta *et al.* [21]. Resulting static displacements were subsequently inverted into the slip distribution along the predefined geometry of the Manila trench using the SDM software [47]. Here, in order to make the inversion test closer to the real world scenario, we did not constrain the possible rupture geometry to the forward model area by length and width. Additionally, we employed alternative
fault discretization with $23 \times 9$ subfaults. Since we were expecting predominantly thrust events at the Manila trench, the rake angle was allowed to vary between 70 and 100 degrees. The two horizontal components were weighted twice as much as the vertical component, and a smoothing constraint was imposed to avoid unrealistic slip patterns.

Figure 4b,c present slip distributions as inverted from the simulated real-time co-seismic BDS (Figure 4b) and joint BDS/GPS (Figure 4c) displacements. Inverted slip distributions from single BDS and BDS/GPS show a high degree of consistency, which is obvious since the processing accuracy is very similar in both cases (Table 1). In contrast, comparison with the forward model (Figure 4a) shows some bias. To be exact, inverted slip distribution covers a somewhat larger area than the original model; the maximum slip of the forward model is 8.4 m while the inverted maximum slips are 7.31 and 7.32 m, respectively, and the derived moment magnitudes are both $M_w = 8.06$, slightly greater than the input value of $M_w = 8.0$. This insignificant discrepancy is explained, first of all, by the fact that our virtual observational network is located, due to natural reasons, only on one side of the Manila trench. Nevertheless, tsunami wave heights as predicted from the inverted source models are very similar to the forward computations (compare vertical bars on Figure 5 and refer to the supplementary Table S1 for more detailed information). Note that the difference in estimated wave heights is larger along the Luzon coast compared to the distant coasts of China and Taiwan. This demonstrates the known fact that, for reliable near-field tsunami early warnings, a trustworthy slip distribution is favorable.

From this synthetic inversion test, we conclude that a BDS (or a BDS/GPS) network with real-time processing, if deployed at the Luzon Island, will be able to contribute to a fast and reliable tsunami source inversion at the Manila trench.

4. Discussions

As we have noted before, precise real-time GNSS processing is a prerequisite for incorporation of coastal GNSS-arrays into tsunami early warnings. In this context, we assessed the potential capability of using real-time BDS to detect on-land co-seismic displacements caused by submarine rupture and addressed the special contribution from combined BDS/GPS solutions as well. According to our estimations from Section 2, horizontal accuracies of single BDS and GPS at the test site were 1.1 cm and 1.2 cm, respectively, while the joint BDS/GPS accuracy reached about 0.7 cm. For vertical component, BDS/GPS improved accuracy from about 2 cm to about 1.0 cm. In terms of current accuracy, BDS is close to GPS in the Asia-Pacific region, and the advantage of fusion of BDS and GPS is clear. Nonetheless, our simulation of the $M_w = 8.0$ earthquake did not show any notable difference in the case of BDS-only or BDS/GPS joint inversion (compare Figure 4b,c). For an earthquake with a magnitude of $M_w = 8.0$, near field deformation can be as large as tens of centimeters (Figure 4a); with respect to this scale of deformation, the signal-to-noise ratio (SNR) improvement by joint BDS/GPS processing is not clear in this case. However, joint BDS/GPS processing may show its advantage for smaller but still tsunamigenic earthquakes.

Consider a scenario of a smaller earthquake. Figure 6 presents a forward model as well as the two source inversion models for an $M_w = 7.5$ earthquake at the Manila trench (compare Figure 6 to Figure 4). Earthquakes of this magnitude will not pose a significant tsunami threat for the coasts of China and Taiwan but can still be dangerous in the near field (Figure 7). In this case, simulated co-seismic displacements are significantly smaller (Figure 6a), and the better signal-to-noise ratio of the joint BDS/GPS real-time processing results in better restoration of the original slip distribution (compare Figure 6b,c) and, therefore, in better tsunami forecasting for the near field (compare vertical bars on Figure 7 and refer to the supplementary Table S1 for more detailed information).
Figure 6. (a) Assumed slip distribution (colored dots) and correspondent co-seismic surface deformation for an event with a magnitude of $M_w = 7.5$ rupturing along the Northern Manila mega-thrust. Red arrows show horizontal displacements computed at the virtual BDS/GPS network (black triangles); (b) Colored dots show slip distribution as inverted from the simulated BDS-displacements (note, synthetic BDS-displacements include real-time processing noise). Additionally shown is the resulting vertical deformation; (c) Same as (b), but for the inversion of the joint BDS/GPS displacements.
Figure 7. Simulated tsunami scenarios. The color map shows the maximum wave height in the forward model (Figure 6a) after 6 h of tsunami propagation. Vertical bars are maximum wave heights as recorded at selected coastal locations for the three propagation models: the forward model (black); the model with the source as inverted from BDS-displacements (dark grey, Figure 6b); and the model with the source as inverted from joint BDS/GPS-displacements (light grey, Figure 6c).

Notice that our models are simple and generic enough to illustrate the main message of the present study: Tsunamigenic earthquakes at the Manila trench will trigger significant co-seismic deformation at the Luzon Island. Deformation amplitudes will be far above modern real-time GNSS resolution. That is why it makes sense to employ coastal GNSS-arrays for tsunami early warnings in the Southern China Sea.

At the same time, authors understand that there are many open questions regarding practical incorporation of GNSS component into any particular TEWS. For example, we opted for static inversion of final co-seismic displacements. Alternatively, some very recent studies [48,49] showed that static inversions have a tendency to over-smoothen slip distribution, which, in turn, can result in the under-prediction of tsunami heights. Kinematic (waveform) inversions demonstrate better spatial resolution and have a clear advantage when static displacements are small. On the other hand, static inversions are fast and robust, and use less data and a smaller number of control parameters.

The problem of optimal choice of methods and algorithms is not, of course, restricted to static vs. kinematic inversion. If the source alone is considered—the GNSS array distribution; the source forward model, which includes geometry, the elastic model, slip distribution; inversion algorithm itself—there is only a short list of issues requiring testing and optimization. It is the same for tsunami simulation and decision-making; each of them also have dozens of options. Authors believe that decisions about better methods and better practices should be taken on an individual basis, in harmonization within particular TEWS and its individual components.

5. Conclusions

In this study, we investigated the precise positioning performance of BDS and BDS/GPS in the Asia-Pacific region, and we specifically focused on their applications to tsunami early warnings in the
SCS region. By analyzing two MGEX stationary stations, we demonstrated that BDS and GPS display similar positioning accuracy in the Asia-Pacific region. Combined BDS/GPS solution not only shows a somewhat higher degree of precision but, what is even more important for operational implementation in decision-making applications like tsunami early warnings, also improves the robustness.

By considering a scenario of a devastating tsunami in the Southern China Sea triggered by a \( M_w = 8.0 \) earthquake along the Manila trench, we propose deploying a continuous BDS/GPS-network at the western coast of the Luzon Island, Philippines. This network, if complemented with the modern real-time processing technique (here TPP), will contribute to a very fast (2–3 min) earthquake evaluation and source inversion, moving toward a higher reliability of the local and regional tsunami early warning.

Using a scenario with a smaller but still tsunamigenic earthquake (\( M_w = 7.5 \)), we demonstrate that the advantage of combined BDS/GPS over single GNSS processing by source inversion grows with decreasing earthquake magnitudes. To conclude, BDS and BDS/GPS can become an important component of the future TEWS in the SCS, although numerous factors need to be optimized in future studies.

Considering that the BDS constellation is still under construction, with increasing numbers of BDS satellites and ground-tracking stations arising in the coming years, its visibility in other places in the world will improve significantly, and the orbit and clock bias products will be more accurate, which indicate that precision of BDS is expected to be equal on a global level. That is to say, the concept of BDS and BDS/GPS as a tsunami early warning component is not limited to the Asia-Pacific region, but will be able to be applied worldwide.

Acknowledgments: K.C. is supported by the China Scholarship Council (CSC) for his PhD study in the German Research Centre for Geoscience. N.Z. is supported by the Helmholtz Association through the GeoSim program. This study was also supported by the National Natural Science Foundation of China (Grant No. 41504010; 41574027; 41325015; 41231174). We appreciate valuable comments from the three reviewers which improve this manuscript greatly.

Author Contributions: Kejie Chen initialized the idea and conception; Kejie Chen, Natalia Zamora and Andrey Babeyko wrote the manuscript; All authors reviewed the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References


2. Japan Meteorological Agency (JMA). Lessons Learned from the Tsunami Disaster Caused by the 2011 Great East Japan Earthquake and Improvements in JMA’s Tsunami Warning System; Japan Meteorological Agency: Tokyo, Japan, 2013.


27. Li, X.; Ge, M.; Dai, X.; Ren, X.; Fritsche, M.; Wickert, J.; Schuh, H. Accuracy and reliability of multi-GNSS real-time precise positioning: GPS, GLONASS, BeiDou, and Galileo. J. Geod. 2015, 89, 607–635. [CrossRef]
33. Suppasri, A.; Imamura, F.; Koshimura, S. Tsunami hazard and casualty estimation in a coastal area the neighbors the Indian Ocean and South China Sea. J. Earthq. Tsunami 2012, 6, 1250010. [CrossRef]

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
Spatio-Temporal Error Sources Analysis and Accuracy Improvement in Landsat 8 Image Ground Displacement Measurements

Chao Ding 1, Guangcai Feng 1,*, Zhiwei Li 1, Xinjian Shan 2, Yanan Du 1 and Huiqiang Wang 1

1 School of Geosciences and Info-Physics, Central South University, Changsha 410083, China; dingchao_csu@csu.edu.cn (C.D.); zwli@csu.edu.cn (Z.L.); yndu_csu@csu.edu.cn (Y.D.); wanghaiqing@csu.edu.cn (H.W.)
2 State Key Laboratory of Earthquake Dynamics, Institute of Geology, China Earthquake Administration, Beijing 100029, China; xjshan@163.com
* Correspondence: fredgps@gmail.com; Tel.: +86-731-888-30573

Academic Editors: Zhenhong Li, Roberto Tomas and Prasad S. Thenkabail
Received: 25 August 2016; Accepted: 7 November 2016; Published: 10 November 2016

Abstract: Because of the advantages of low cost, large coverage and short revisit cycle, Landsat 8 images have been widely applied to monitor earth surface movements. However, there are few systematic studies considering the error source characteristics or the improvement of the deformation field accuracy obtained by Landsat 8 image. In this study, we utilize the 2013 Mw 7.7 Balochistan, Pakistan earthquake to analyze error spatio-temporal characteristics and elaborate how to mitigate error sources in the deformation field extracted from multi-temporal Landsat 8 images. We found that the stripe artifacts and the topographic shadowing artifacts are two major error components in the deformation field, which currently lack overall understanding and an effective mitigation strategy. For the stripe artifacts, we propose a small spatial baseline (<200 m) method to avoid the stripe artifacts effect on the deformation field. We also propose a small radiometric baseline method to reduce the topographic shadowing artifacts and radiometric decorrelation noises. Those performances and accuracy evaluation show that these two methods are effective in improving the precision of deformation field. This study provides the possibility to detect subtle ground movement with higher precision caused by earthquake, melting glaciers, landslides, etc., with Landsat 8 images. It is also a good reference for error source analysis and corrections in deformation field extracted from other optical satellite images.

Keywords: Landsat 8; cross-correlation; displacement monitoring; error analysis; spatial baseline; radiometric baseline

1. Introduction

With the continuous improvements of optical satellites in both hardware and software, moderate and high-resolution optical images are increasingly applied in ground surface displacement measurement [1]. Bindschadler et al. [2] took the lead in extracting the velocity field of an Antarctic ice stream by matching multi-temporal Landsat 5 TM images. Crippen et al. [3] preliminary investigated sand dunes migration in the southernmost California, USA with SPOT (Systeme Probatoire d’Observation de la Terre) images. The image matching theory was subsequently expanded as feature tracking in SAR (Synthetic Aperture Radar) imagery [4–6] and firstly applied to obtain coseismic deformation of the 1992 Mw 7.3 Landers earthquake [4]. Puymbroeck et al. [7] captured surface ruptures and horizontal coseismic deformation field of the 1992 Mw 7.3 Landers earthquake using the SPOT images with cross-correlation technique. After the software Coregistration of Optically Sensed Images and Correlation (COSI-Corr) had been released, the theory and algorithms of the optical
images cross-correlation technique were greatly improved and matured, especially in accuracy and computation efficiency [8]. The COSI-Corr is originally used to extract crustal deformation induced by earthquake released by Tectonics Observatory at Caltech in 2008 [8,9]. The technique is now widely used to monitor the earth surface changes and dynamics [8–15]. As richer and constantly updated satellite observation data become available, the application of optical images cross-correlation technique has been expanded to fields such as glacier flow velocity extraction [10–12], quantification of sand dune migration [13–15], terrain-deformation measurements of slow landslides [16], monitoring process of significant rift events [17], volcanic monitoring [18] and particularly extensive coseismic deformation extraction [19–21].

As a continuation of the NASA Landsat program, the Landsat 8 satellite was launched in February 2013 and is operating in a good condition now. Equipped with OLI (Operational Land Imager) and TIRS (Thermal Infrared Sensor) sensors, Landsat 8 can obtain image data of 11 bands, wherein the panchromatic band (Band 8) has a ground resolution of 15 m, equivalent to ASTER (Advanced Spaceborne Thermal Emission And Reflection Radiometer) VNIR bands but with a larger spatial coverage [22–25]. Since its operation, the Landsat 8 data have been applied to monitor earth surface movements. Many studies demonstrate that the Landsat 8 image cross-correlation technique has advantages in measuring surface deformation, such as mapping ice sheet velocity [26], coseismic deformation extraction [25,27–29] and volcano deformation monitoring [30]. Many researchers have done a lot of analyses and error source correction of displacement measurement, but they mainly focused on a few data sources, such as SPOT, ASTER, HiRISE (High Resolution Imaging Science Experiment) and aerial imagery [3,8,11,13–17,19–21,31,32]. In general, the deformation field obtained by different types of optical sensors with cross-correlation technique has disparities in error structure and magnitude level. As one of the few free available in-orbit operational data sources, Landsat 8 imagery has a relatively higher resolution, high imaging quality (12 bit) and short revisit cycle (16 days). The Landsat 8 surpasses most satellites in both scientific performances and economic cost. No doubt, earth surface deformation monitoring based on the Landsat 8 image will be more widely used in the future. Currently, there are few detailed studies on the error sources of deformation fields obtained by the Landsat 8 image. Thus, it is necessary and important to systematically study the error source components, characteristics and corresponding correction methods of the Landsat 8 image deformation field.

Based on this, we firstly analyze the error source compositions, distribution patterns and magnitude variation of the Landsat 8 image deformation field. Then, we give comprehensive explanations on the reason of error generation and propose corresponding correction methods. We also compare the results obtained by our method and traditional method, and validate the results’ precision. Finally, we discuss the potential application of our proposed methods in time series analysis with the Landsat 8 image.

2. Experimental Design and Data Processing

Balochistan, Pakistan is selected as the study area because it has dry climate and little cloud cover, which is favorable for optical images acquisition (see Figure 1). In addition, a $M_w$ 7.7 earthquake happened there on 23 September 2013, leaving obvious surface ruptures [25,27–29]. Totally, 25 images covering the study area (Path: 154, Row: 42, see Figure 1, Table A1) are selected, pairing freely with a temporal baseline of 16 days (see Figure A1) and calculating mutually image deformation fields.

Considering the computational efficiency, we uses a phase correlation algorithm in frequency domain in the experiments, which can quickly obtain the translation between correlated images and easily separate frequency-dependent noises in images. The Fast Fourier Transformation (FFT) correlation engine embedded in the algorithm of COSI-Corr [8–15] improves the traditional peak correlation algorithm, and its accuracy can reach 1/50 pixel [8]. Thus, we select COSI-Corr as the data processing platform. We set both the initial and final search window size as 32 pixel \(\times\) 32 pixel, step as 4 pixel \(\times\) 4 pixel and the corresponding ground resolution is 60 m \(\times\) 60 m. For mitigating the effects of
high-frequency noises on the accuracy of correlation results, we set the frequency mask threshold as 0.9 and robust iteration as 2. To maintain a good consistency of results correlated by different temporal images, these configuration parameters are used in cross-correlation calculation proceeding for the whole study.

Figure 1. Coverage of Landsat 8 images over the study area. The red star indicates the epicenter of the 2013 Mw 7.7 Balochistan, Pakistan earthquake. The black boxes are the footprint of the Landsat 8 images in this study. The red rectangle frame in the inset represents the approximate location of the study area in Pakistan.

In order to obtain sufficient correlation samples and weaken the long-time baseline influence, we select 16 days as the temporal baseline to pair freely the 25 images (Path: 154, Row: 42) in Table A1. To maintain the uniform standard of temporal baseline (16 days) of pair images, we acquire redundant observations containing enough variables, such as different orbit overlap and shadowing differences. This can facilitate the analysis of error source including stripe artifacts and topography-dependent artifacts relevant to these variables. We obtain 16 pairs of image deformation fields totally (East–West and North–South) with a ground resolution of 60 m × 60 m, as shown in Figure 2. It should be noted that these results have not been post-processed, so there are a variety of error sources in deformation fields.
3. Error Source Analysis and Correction Methods

As shown in Figure 2, by analyzing results in both E/W and N/S components, we can find that there are following types of error sources affecting the measurement accuracy, including temporal decorrelation noises, long wavelength orbital error, stripe artifacts (SA), attitude jitter distortions and topography-dependent artifacts. Figure 3 shows patterns of these errors, but excludes the attitude jitter distortions with a small magnitude. In the following, we will analyze these error sources, and propose some mitigation methods.

Figure 2. E/W and N/S components of image deformation fields. The serial numbers of all sub-graphs correspond to correlation image pairs in Table 1. Displacements are positive toward the East and North in this and following figures.

Figure 3. Error sources in E/W and N/S components of image deformation field, selecting the correlation image pair 29 November 2013–15 December 2013 as an example. We can see there are obvious orbital error, decorrelation noises, topography-dependent artifacts, and stripe artifacts in image deformation field.
3.1. Decorrelation Noises

Decorrelation noises are closely correlated to the radiometric properties of earth surface. Too high, too low or invariant radiation can lead to inconspicuous texture features, and drastic radiometric changes commonly can cause texture defect. Thus, it is difficult to achieve the best correlation based on window matching, showing inaccurate random measurements and low signal-to-noise ratio (SNR) values in corresponding regions. To reduce the effect of decorrelation noises, this study masks out measurements with low SNR (<0.9) and null values (NaN) in image deformation fields, and manually removes the large range of decorrelation noises caused by clouds, water regions, etc.

3.2. Long Wavelength Orbital Error

If Landsat 8 images downloaded from the USGS website are not orthorectified enough, they will cause significant signals of linear ramp in image deformation fields. In the image pre-processing, this type of error can be eliminated by rigorous geometric correction to original images if the state vector of the Landsat 8 image is available. In this study, we remove this error using the first-order polynomial fitness [25] because Landsat 8 images do not provide the state vector in these products. The results are detailed in Figure 4. The polynomial curve fitting method is also used to remove the orbital error in InSAR (Interferometric Synthetic Aperture Radar) interferogram [33–44].

![Figure 4. E/W and N/S components of image deformation field before and after removing the orbital error, taking the correlation image pair 29 November 2013–15 December 2013 as an example.](image)

3.3. Stripe Artifacts

Misalignment of Charge Coupled Device (CCD) arrays is a common problem in push-broom satellites with mosaic CCD imaging cameras. The general orthorectification cannot completely eliminate the CCD distortions, which perform as stripe artifacts (SA) in image deformation fields [31]. In Figure 5, there are many obvious evenly spaced bands along the flight direction in the crustal deformation field. We find that those SA are correlated with the corresponding spatial positions of CCD line-arrays imaging of pre- and post-event images. A method called “mean subtracting” was proposed to measure the patterns of SA in deformation fields [8,11,31]. This method can partly estimate and remove the SA using the mean value along the satellite flight direction in the stable area away from tectonic signals and serious decorrelation regions.
Figure 5. Relevance of the SA (stripe artifacts) and the CCD (Charge Coupled Device) arrays’ relative locations superimposed on the E/W component of the crustal deformation field. The yellow rectangles represent CCD arrays’ locations acquired on the images from: 10 September 2013 (top); and 26 September 2013 (bottom). The dislocation between the two CCD arrays presents inconsistency in scene center (spatial baseline), which is the main reason causing the SA. The misalignment between CCD sub-arrays in a single CCD chip is identified as the essential reason. Blue profile line represents the average along flight direction of gray block area with a magnitude range of $-1.5$ m–$2$ m ($-0.1$–$0.13$ pixels).

In this study, we try to understand the relationship between the SA and CCD arrays’ relative locations between correlated images. We acquire patterns of the SA for those 16 pairs of images in Figure A2. It can be found that most patterns of the SA jag significantly, and only three pairs (22 April 2014–8 May 2014, 16 November 2014–2 December 2014 and 20 February 2015–8 March 2015) present approximate straight lines with the mean value closing to zero. From Table 1, we find that those three pairs are different from the rest pairs in the spatial distances between the correlated image centers. Spatial distances of these three pairs are shorter than 180 m, while that of other pairs are up to several thousands (see Table 1). Here we use “spatial baseline” to define this spatial distance between two correlated images’ centers, which can be calculated based on the orthoimages centers’ coordinates in master and salve images. To quantify the magnitude level of the SA in different image pairs, we define a value $\gamma$ as a factor for quantifying the average magnitude of the SA.

$$\gamma = \frac{\sum_{i=1}^{n} |T_i|}{n}$$  \hspace{1cm} (1)

where, as shown in Figure A2, $i$ represents the transverse-axis parameter with a total number of $n$, and $T$ represents the measurement values on the vertical-axis in a single pattern of the SA. Using these data, we perform a regression analysis to examine the relationship between the spatial baseline and $\gamma$ in Figure 6. We find that variant $\gamma$ and spatial baseline have an approximately positive linear correlation, meaning the smaller the spatial baseline, the lower the magnitude level of the SA.
In fact, the SA are common in horizontal deformation fields extracted from the current mainstream optical images (e.g., ASTER, SPOT, and Landsat 8). The “mean subtracting” method is a conventional approach to mitigate similar linear artifacts [11,25,27]. In this study, we proposed a new method to reduce the SA by selecting correlation image pairs with small spatial baseline. In order to compare the performance of these two methods, we select two image pairs, 21 May 2013–11 May 2015 with a small spatial baseline (<100 m) and 10–26 September 2013 with a large spatial baseline 2324 m, for mapping coseismic deformation of this event. As shown in Figure 7, it has no significant SA in the E/W and N/S components of crustal deformation field produced by image pair 21 May 2013–11 May 2015. In contrast, the correlation results of image pair 10–26 September 2013 processed by the “mean subtracting” method still have some unmolded residuals, suggesting the “mean subtracting” method...
cannot completely remove the SA, especially when the stable regions are hard to choose for the calibration. Furthermore, the magnitude of the SA along the flight direction would be variable, not constant, if the topography along the satellite flight direction is changeable. In this study, the statistical results indicate that the image pair with the spatial baseline 200 m can lead to a $\gamma$ less than 0.095 m, which can be neglected in surface deformation. Thus, it can be a recommended threshold value for spatial baseline selection in other cases.

**Figure 7.** Comparison between the “mean subtracting” method and small spatial baseline method in terms of the SA reduction: (a,b) the E/W components of 10–26 September 2013 using the “mean subtracting” method and 21 May 2013–11 May 2015 using the small spatial baseline method; and (c,d) the N/S components of 10–26 September 2013 using the “mean subtracting” method and 21 May 2013–11 May 2015 using the small spatial baseline method. This comparison further demonstrates that the magnitude level of SA has a strong correlation with spatial baseline of the correlated image pair. In addition, the ellipse lines circle the topography-dependent artifacts area, in which the red arrows in (c) indicate the specific locations.

### 3.4. Satellite Attitude Distortions

As the satellite attitude variations are sparsely sampled during flying, deformation fields (see Figure 8a) have cyclical distortions along the satellite flight direction, which are roll variations on the East–West component and pitch variations on the North–South component. Such error is distributed approximately constant along the satellite across-track direction. Thus, for the deformation fields with a serious attitude jitter distortions, the destriping procedure can be used to reduce cyclical distortions by subtracting the average along azimuth vertical direction based on stable calibration regions [11].
Figure 8. E/W and N/S components before (a) and after (b) subtracting the attitude distortions (red line) along the across-track direction in correlation image pair 16 November 2014–2 December 2014.

3.5. Topographic Shadowing Artifacts

Solar position change, described by variation of sun azimuth and sun elevation angle, will cause shadowing difference in orientation and length (see Figure 9). The COSI-Corr algorithm will calculate related “offset” when the shadow texture dominates the correlation window. The experimental results in Figure A3 show that there are obvious topography-dependent biases in the Hinglaj Mountains area of the southeast image deformation fields, especially in the N/S component. Comprehensive analysis of Figure A3 and Table A2 shows that topography-dependent artifacts change with the sun angle disparity. The larger the difference in sun angle, the more obvious the artifacts. Generally, the artifacts are about 3–4 m, and the maximum is up to 10 m in the steep mountain area. Furthermore, the positive and negative variations of topography-dependent artifacts have a good correlation with the shadowing orientation changes (see Figure 9). Thus, we draw a conclusion that shadowing difference can be a source of the topography-dependent artifacts in image deformation fields, which is referred to “topographic shadowing artifacts” (TSA) in this study.

Based on above analysis, we can see the TSA is usually fixed in the spatial location and strongly related to the disparity of sun elevation and azimuth of the image. A factor “radiometric baseline” can be defined to evaluate the influence level of the TSA using the orientation and length of vector offset in Figure 9 and Table 2. The smaller is the radiometric baseline, the lower the level of the TSA and radiometric decorrelation noises becomes. As shown in Figure 7c,d, the radiometric baseline (0.14 h) in N/S component of image pair 10–26 September 2013 is much larger than that (0.05 h) of image pair 21 May 2013–11 May 2015 (h is the height of a reference datum in Figure 9). Thus, the TSA in the former pair is more significant than that of later pair in Figure 7. Moreover, the offset bias caused by the sun angle difference in 10 September 2013 and 26 September 2013 is mainly along N3° orientation (see Equation (3)), so the positive values TSA mainly distribute in the N/S component displacement map (see Figure 10a,d).
Table 2. Radiometric baseline list of the Landsat 8 correlation image pairs.

<table>
<thead>
<tr>
<th>Pairs Num.</th>
<th>Image Pair</th>
<th>$\Delta = \text{Pre.} - \text{Pos.}$</th>
<th>Radiometric Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\Delta$ Azi.</td>
<td>$\Delta$ Ele.</td>
</tr>
<tr>
<td>1</td>
<td>21 May 2013–6 June 2013</td>
<td>7.09</td>
<td>-0.25</td>
</tr>
<tr>
<td>2</td>
<td>10–26 September 2013</td>
<td>-9.46</td>
<td>3.98</td>
</tr>
<tr>
<td>3</td>
<td>26 September 2013–12 October 2013</td>
<td>-7.22</td>
<td>4.64</td>
</tr>
<tr>
<td>4</td>
<td>12–28 October 2013</td>
<td>-4.82</td>
<td>4.82</td>
</tr>
<tr>
<td>5</td>
<td>29 November 2013–15 December 2013</td>
<td>0.98</td>
<td>2.26</td>
</tr>
<tr>
<td>6</td>
<td>22 April 2014–8 May 2014</td>
<td>9.80</td>
<td>-2.93</td>
</tr>
<tr>
<td>7</td>
<td>8–24 May 2014</td>
<td>9.28</td>
<td>-1.36</td>
</tr>
<tr>
<td>8</td>
<td>24 May 2014–9 June 2014</td>
<td>6.24</td>
<td>-0.07</td>
</tr>
<tr>
<td>9</td>
<td>9–25 June 2014</td>
<td>1.68</td>
<td>0.64</td>
</tr>
<tr>
<td>10</td>
<td>13–29 September 2014</td>
<td>-9.14</td>
<td>4.07</td>
</tr>
<tr>
<td>11</td>
<td>29 September 2014–15 October 2014</td>
<td>-6.88</td>
<td>4.65</td>
</tr>
<tr>
<td>12</td>
<td>16 November 2014–2 December 2014</td>
<td>-0.36</td>
<td>3.39</td>
</tr>
<tr>
<td>13</td>
<td>3–19 January 2015</td>
<td>3.08</td>
<td>-1.47</td>
</tr>
<tr>
<td>14</td>
<td>20 February 2015–8 March 2015</td>
<td>4.46</td>
<td>-5.27</td>
</tr>
<tr>
<td>15</td>
<td>9–25 April 2015</td>
<td>8.33</td>
<td>-4.10</td>
</tr>
<tr>
<td>16</td>
<td>25 April 2015–11 May 2015</td>
<td>9.77</td>
<td>-2.68</td>
</tr>
</tbody>
</table>

Notes: The $\Delta$Azi. and $\Delta$Ele. represent the sun azimuth disparity and sun elevation disparity, respectively.

Figure 9. Mechanism of topographic shadowing artifacts. Equations below the figures show how to calculate the magnitude and orientation of the offset caused by shadowing variation. ($\alpha_1$, $\beta_1$) and ($\alpha_2$, $\beta_2$) represent the sun elevation and azimuth of pre- and post-event images, respectively. L1 and L2 represent the shadow lengths on a flattish surface of a reference datum named O with height $h$ at the sun elevation of $\alpha_1$ and $\alpha_2$, respectively. Equations (2) and (3) show how to calculate the magnitude and orientation $\theta$ of vector offset due to shadowing variation, in which $\Delta \alpha$ and $\Delta \beta$ denote the sun elevation and sun azimuth disparity, respectively. Equations (4) and (5) represent how to calculate the E/W offset and N/S offset decomposed from the vector offset, and the results are positive toward East and North.
Figure 10. Horizontal coseismic deformation of the 2013 $M_w$ 7.7 Balochistan, Pakistan earthquake obtained by our small radiometric baseline method, before and after masking out TSA and other non-targeted offsets: (a) E/W displacement field before masking; (b) E/W mask image, in which the white denotes the region that need to be masked; (c) E/W displacement field after masking; (d) N/S displacement field before masking; (e) N/S mask image; and (f) N/S displacement field after masking.

Here we can reduce the TSA in the deformation field by choosing the image pairs with a small radiometric baseline. The good performance of this strategy has been demonstrated by image pair 21 May 2013–11 May 2015 of Figure 7. We can also utilize the combined radiometric baseline to reduce TSA using multi-pair of images, if the single small radiometric baseline image pair is unavailable due to few archived images in the study area. Based on these analyses, we can summarize the strategy of reducing the TSA and radiometric decorrelation noises as the small radiometric baseline method corresponding to two situations of single and combined radiometric baselines. In the situation of single radiometric baseline, there is only one correlated image pair with a very small radiometric baseline. The TSA and radiometric decorrelation noises in the correlation result are very small in this situation, so it can be ignored. Furthermore, in the corresponding situation of combined radiometric baseline, there are two or more correlated image pairs with very similar radiometric baselines. The TSA and radiometric decorrelation noises in those correlation results are very close each other, so it can be cancelled by combination of image pairs. In the 2013 $M_w$ 7.7 Balochistan, Pakistan earthquake study, we find the radiometric baseline of image pair 13–29 September 2014 ($0.142$ h, N $3.1^\circ$) is very similar to that of image pair 10–26 September 2013 ($0.139$ h, N $2.9^\circ$) (see Table 2). Thus, we firstly acquire the image deformation field associated with the TSA and other seasonal errors except the earthquake event, using the correlation image pair 13–29 September 2014. The noise level of the TSA in the image...
deformation field extracted from image pair 13–29 September 2014 is very similar to the counterpart in the coseismic deformation field extracted from image pair 10–26 September 2013, due to the similar radiometric baseline in Table 2. After removing the decorrelation noises, orbital error, SA and attitude jitter distortions, we can obtain a mask image from image pair 13–29 September 2014 by setting its deformation field greater than a threshold both in E/W and N/S components (see Figure 10b,e). We choose the non-seismic deformation threshold value as 0.8 m because the surface displacements with 1/10–1/20 pixel magnitude accuracy can be usually detected by COSI-Corr [8]. Finally, we mask out corresponding TSA in the coseismic deformation field (10–26 September 2013) using the mask image obtained. This method can also be used to identify other non-targeted deformation regions with a high magnitude level, such as melting glaciers, migratory sand dunes and creeping landslides. Figure 10c, f indicate the E/W and N/S components of coseismic displacement field processed by our small radiometric baseline method in situation of combined radiometric baseline. Comparing between Figure 10d, f, we can find that TSA in the southeast and the moving sand dune area located in the northwest has been masked out effectively. We also use a non-local means filter tool in COSI-Corr to further reduce the noises in the deformation field.

4. Discussion

4.1. Error Analysis and Accuracy Assessment

Using the methods described in Section 3, we can remove most error sources in Landsat 8 image deformation fields. In order to analyze the error characteristics and validate the performance of our new methods in deformation fields, we have analyzed statistically the error level of a stable area far from the crustal deformation field in image pair 10–26 September 2013 (see the polygon area in Figure 11). The selected area contains a variety of land features, including rolling hills, flat plains and vegetation area. We divide the whole data process into four steps and calculate the means and standard deviations after those procedures in Table 3 and Figure 11. Those four steps’ results (R0, R1, R2, and R3) represent the process results of the original COSI-Corr process, after removing orbital error, after removing SA and attitude jitter distortions, and after removing TSA, respectively, in Figure 11. It should be noted that the procedures of reducing SA and attitude jitter distortions have been combined in R2, because they have similar error patterns.

Table 3. Accuracy evaluation results.

<table>
<thead>
<tr>
<th>Experimental Region</th>
<th>Mean/m</th>
<th>Standard Deviation/m</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R0</td>
<td>R1</td>
</tr>
<tr>
<td>E/W</td>
<td>0.251</td>
<td>0.083</td>
</tr>
<tr>
<td>N/S</td>
<td>−0.819</td>
<td>0.122</td>
</tr>
</tbody>
</table>

Figure 11. Flow chart of accuracy improvement in the Landsat 8 crustal deformation field. The black polygon line denotes the area used for accuracy assessment.

In Figures 11 and 12, we use image pair 10–26 September 2013 instead of the small spatial baseline pair 21 May 2013–11 May 2015 for accuracy analysis because the former is more effective for the 2013 Mw 7.7 Balochistan, Pakistan earthquake study. It has been widely used for geophysical
interpretation of this earthquake and also can be used for different error source validation in our study. The comparison in Figure 7 has explicitly proved the advantage of the small spatial baseline method. From Figure 12, both the mean and standard deviation of R1 and R2 become smaller compared to R0. It indicates that these two corresponding methods used in this paper are effective in reducing noises in coseismic deformation fields. After removing the TSA by the small radiometric baseline method, the noise level of R3 is further reduced, in which the standard deviation of E/W component reduces about 3.7% (around 0.02 m) and N/S component reduces about 15.2% (around 0.09 m). It is worth noting that, after subtracting the TSA from the N/S component of the coseismic deformation field (represented by R3), the absolute mean value increased from 0.012 m to 0.06 m (0.012 m vs. 0.06 m). As the TSA in the N/S component of the deformation field mainly have positive magnitudes, masking operation to reduce the TSA will mask amount of positive values. Thus, the mean of process result R3 shows a slight increase in the selected small area. We believe that the process result R3 would not increase the absolute mean value in large areas. As shown in Table 3, the final standard deviations of E/W and N/S components are 50.4 cm and 51.8 cm, respectively. From this perspective, the selected threshold (0.8 m) in the process of removing TSA is just about $\mu \pm 2\sigma$ ($\mu$—mean, $\sigma$—standard deviation, $[\mu - 2\sigma, \mu + 2\sigma]$ account for 95.45% measured values), which further validates the statistical reasonability of the threshold.

Figure 12. Histograms of the process results R0, R1, R2, and R3 in the selected area. The histograms are shown as below: original COSI-Corr process result (R0, red line), result after removing orbital error (R1, black line), result after removing SA and attitude jitter distortions (R2, blue line), and result after removing TSA (R3, pink line). The vertical dotted lines represent the relevant mean (location) of histogram curves in the same color.
4.2. Comparison with Current Researches

In the process of error analysis, we could know the magnitude level of every type of error sources in Landsat 8 image deformation field, such that the orbital error with around $-5\text{--}5$ m, the SA with around $-2\text{--}2$ m, the attitude distortions with around $-0.5\text{--}0.5$ m, and the TSA with around $-10\text{--}10$ m. In our study, we could easily remove the orbital error and attitude distortions. However, we can find little detailed studies with respect to the SA and TSA in Landsat 8 image deformation field. Thus, we endeavor to give comprehensive explanations for the reasons of error sources and propose corresponding weaken methods to reduce the SA and TSA from the perspective of time series analysis in our study. This is a work discriminating from other researches.

Due to differences in CCD sensors, the SA has different patterns and correction methods in different optical satellite images [31]. For example, the rigorous CCD calibration model proposed by Leprince et al. [31] has a good applicability for the SPOT4 image, but it is not suitable for the Landsat 8 image due to lacking reference images and auxiliary state vector. In this study, we propose the small spatial baseline method for the Landsat 8 image to reduce the SA. Compared with the traditional “mean subtracting” method, it can obtain much cleaner and more accurate results if suitable spatial baseline image pair is available. Basing on the experimental results, we suggest that the spatial baseline of 200 m can be used as a reference threshold to select correlation image pairs. However, the “mean subtracting” method can be a complementary approach when the archived Landsat 8 image is very limited in the study area. Furthermore, for sudden events study, such as earthquakes, the “mean subtracting” method can also be a complementary approach, because the larger temporal baseline may introduce much larger signals and decorrelation noises into the crustal deformation field.

Topographic shadow effect is a common problem in optical images, its orientation and length variations can lead to the TSA, especially in monitoring glacier and landslide [1,29]. Due to no effective reduction method, the TSA still remain in the coseismic deformation field of the 2013 Mwo 7.7 Balochistan, Pakistan earthquake [25,27]. In this study, we introduce a small radiometric baseline method to reduce the TSA and radiometric decorrelation noises. The TSA and other non-targeted displacements, crustal deformation excluding the coseismic deformation, are effectively masked out by our method, with much smaller mean and standard deviation both in E/W and N/S components. However, it is worth noting that although we could reduce the influence of TSA and radiometric decorrelation noises by our proposed method based on combined radiometric baselines, the crustal signals sometimes are also masked out when the signals and noises simultaneously occur in masked gaps. In addition, the TSA and stereoscopic parallax effects are easily confused, as they both having characteristics associated with topographic features. However, the latter also belongs to the geometric distortions and can be mitigated by a rigorous geometric model [1]. We can discriminate them by analyzing base-height ratio or the temporal variation characteristics of these two types of errors.

The Landsat 8 satellite has been continuously monitoring the surface change since the first half year of 2013 and it has acquired sufficient images worldwide. Thus, it enables exploring the time series surface deformation by the multi-temporal Landsat 8 images. The key issue is how to construct the optimal image pairs to accurately acquire the crustal deformation fields. Our small spatial baseline and small radiometric baseline methods based on various error analyses and interpretations can be verified as effective time series analysis methods. The richer the Landsat 8 archived images, the more effective these methods are able to extract the earth surface deformation, with higher precision, which is very helpful for investigating the slow but steady dynamic phenomenon of earth surface, such as the melting glaciers and migratory sand dunes. Furthermore, based on detailed studies on error sources in this paper, we can acquire more precise fault rupture geometry and coseismic displacement field, which may be helpful for us to strengthen the understanding of fault displacement mechanism and improve the fault slip modeling.
5. Conclusions

In this study, we summarize spatio-temporal distribution characteristics and completely systematic removal processes of error sources in correlation results of Landsat 8 images. Through error analysis, we propose two new conceptual terms of “spatial baseline” and “radiometric baseline”, as influential factors with relevance to the SA and TSA, respectively. Then, innovatively, we introduce the small spatial baseline method and small radiometric baseline method to reduce the SA and TSA in deformation field. Confirmatory experiments verify that the spatial baseline of correlated images with below 200 m is effective in reducing the influence of the SA in correlation results. Furthermore, the small radiometric baseline method, using combined radiometric baseline, improves the precision of the deformation field with apparent TSA by around 15% in our study. Meanwhile, the small radiometric baseline method, using single radiometric baseline, can also be helpful for reducing the radiometric decorrelation noises in deformation field. We also expect the content of this study can provide theoretical and experimental guidance for further utilizing the Landsat 8 image to monitor earthquakes, glaciers and so on.

Acknowledgments: This work is supported by Shenghua Yuying fund of Central South University, the National Natural Science Foundation of China (No. 41574005), and the State Key Laboratory of Earth Dynamics Open Fund (Grant Nos. LED 2014B02). We acknowledge Caltech for the COSI-Corr software. We thank the USGS for the free Landsat 8 optical images. We used GMT (Generic Mapping Tools) to prepare the figures. The authors would like to thank Marcello de Michele and other three anonymous reviewers for their constructive remarks and suggestions.

Author Contributions: Guangcai Feng supervised and designed the research work, as well as wrote the manuscript. Chao Ding performed the experiments, analyzed results and wrote the first draft of the manuscript. Zhiwei Li, Xinjian Shan, Yanan Du and Huiqiang Wang contributed to experiment implementation and result interpretation. All authors contributed to paper writing and revision.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

- **SPOT**: Systeme Probatoire d’Observation de la Terre
- **ASTER**: Advanced Spaceborne Thermal Emission and Reflection Radiometer
- **VNIR**: Visible and Near-InfraRed
- **HiRISE**: High Resolution Imaging Science Experiment
- **NASA**: National Aeronautics and Space Administration
- **Caltech**: California Institute of Technology
- **USGS**: United States Geological Survey
- **FFT**: Fast Fourier Transformation
- **COSI-Corr**: Coregistration of Optically Sensed Images and Correlation
- **OLI**: Operational Land Imager
- **TIRS**: Thermal Infrared Sensor
- **CCD**: Charge Coupled Device
- **SNR**: Signal-to-Noise Ratio
- **SAR**: Synthetic Aperture Radar
- **InSAR**: Interferometric Synthetic Aperture Radar
- **SA**: Stripe artifacts
- **TSA**: Topographic shadowing artifacts
Appendix A

Figure A1. Matching distribution of 16 correlation pairs using 25 Landsat 8 panchromatic images, following the uniform temporal standard of 16 days.

Table A1. Parameters catalog of Landsat 8 OLI Pan images used in the study.

<table>
<thead>
<tr>
<th>ID Codes</th>
<th>Acquired Time</th>
<th>Sun Azimuth</th>
<th>Sun Elevation</th>
<th>Cloud Cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>LC81540422013141LGN01</td>
<td>21 May 2013</td>
<td>101.360</td>
<td>69.407</td>
<td>0.74</td>
</tr>
<tr>
<td>LC81540422013157LGN00</td>
<td>6 June 2013</td>
<td>94.267</td>
<td>69.655</td>
<td>0.76</td>
</tr>
<tr>
<td>LC81540422013253LGN00</td>
<td>10 September 2013</td>
<td>131.684</td>
<td>60.502</td>
<td>1.37</td>
</tr>
<tr>
<td>LC81540422013269LGN00</td>
<td>26 September 2013</td>
<td>141.144</td>
<td>56.521</td>
<td>0.70</td>
</tr>
<tr>
<td>LC81540422013289LGN00</td>
<td>12 October 2013</td>
<td>148.360</td>
<td>51.883</td>
<td>3.08</td>
</tr>
<tr>
<td>LC81540422013301LGN01</td>
<td>28 October 2013</td>
<td>153.179</td>
<td>47.062</td>
<td>0.38</td>
</tr>
<tr>
<td>LC81540422013333LGN00</td>
<td>29 November 2013</td>
<td>156.360</td>
<td>38.990</td>
<td>1.32</td>
</tr>
<tr>
<td>LC81540422013349LGN00</td>
<td>15 December 2013</td>
<td>155.384</td>
<td>36.729</td>
<td>0.64</td>
</tr>
<tr>
<td>LC81540422014112LGN00</td>
<td>22 April 2014</td>
<td>118.468</td>
<td>64.727</td>
<td>0.76</td>
</tr>
<tr>
<td>LC81540422014128LGN00</td>
<td>8 May 2014</td>
<td>108.668</td>
<td>67.655</td>
<td>0.60</td>
</tr>
<tr>
<td>LC81540422014144LGN00</td>
<td>24 May 2014</td>
<td>99.387</td>
<td>69.013</td>
<td>1.99</td>
</tr>
<tr>
<td>LC81540422014160LGN00</td>
<td>9 June 2014</td>
<td>93.145</td>
<td>69.080</td>
<td>0.58</td>
</tr>
<tr>
<td>LC81540422014176LGN00</td>
<td>25 June 2014</td>
<td>91.468</td>
<td>68.439</td>
<td>8.62</td>
</tr>
<tr>
<td>LC81540422014256LGN01</td>
<td>13 September 2014</td>
<td>132.788</td>
<td>59.586</td>
<td>1.05</td>
</tr>
<tr>
<td>LC81540422014272LGN00</td>
<td>29 September 2014</td>
<td>141.927</td>
<td>55.520</td>
<td>1.34</td>
</tr>
<tr>
<td>LC81540422014288LGN00</td>
<td>15 October 2014</td>
<td>148.802</td>
<td>50.866</td>
<td>1.68</td>
</tr>
<tr>
<td>LC81540422014320LGN00</td>
<td>16 November 2014</td>
<td>155.558</td>
<td>41.757</td>
<td>0.54</td>
</tr>
<tr>
<td>LC81540422014336LGN00</td>
<td>2 December 2014</td>
<td>153.921</td>
<td>38.370</td>
<td>1.20</td>
</tr>
<tr>
<td>LC81540422015003LGN00</td>
<td>3 January 2015</td>
<td>152.443</td>
<td>36.075</td>
<td>0.48</td>
</tr>
<tr>
<td>LC81540422015019LGN00</td>
<td>19 January 2015</td>
<td>149.367</td>
<td>37.544</td>
<td>1.92</td>
</tr>
<tr>
<td>LC81540422015051LGN01</td>
<td>20 February 2015</td>
<td>141.845</td>
<td>45.076</td>
<td>1.07</td>
</tr>
<tr>
<td>LC81540422015067LGN00</td>
<td>8 March 2015</td>
<td>137.384</td>
<td>50.341</td>
<td>0.50</td>
</tr>
<tr>
<td>LC81540422015099LGN00</td>
<td>9 April 2015</td>
<td>125.252</td>
<td>61.051</td>
<td>7.07</td>
</tr>
<tr>
<td>LC81540422015115LGN00</td>
<td>25 April 2015</td>
<td>116.927</td>
<td>65.146</td>
<td>0.59</td>
</tr>
<tr>
<td>LC81540422015131LGN00</td>
<td>11 May 2015</td>
<td>107.159</td>
<td>67.828</td>
<td>1.59</td>
</tr>
<tr>
<td>LC81540422015253LGN00</td>
<td>10 September 2013</td>
<td>133.903</td>
<td>59.699</td>
<td>0.01</td>
</tr>
<tr>
<td>LC81540422015269LGN00</td>
<td>26 September 2013</td>
<td>142.841</td>
<td>55.534</td>
<td>0.03</td>
</tr>
<tr>
<td>LC81540422015285LGN00</td>
<td>13 September 2014</td>
<td>134.905</td>
<td>58.761</td>
<td>0.01</td>
</tr>
<tr>
<td>LC81540422015301LGN01</td>
<td>29 September 2014</td>
<td>143.547</td>
<td>54.520</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Notes: All images are acquired in Nadir and processed into L1T level.
Figure A2. Patterns of the SA in E/W and N/S deformation fields. Clearly, patterns of these three pairs (22 April 2014–8 May 2014, 16 November 2014–2 December 2014, and 20 February 2015–8 March 2015) with much less variation are different from the ones of the rest image pairs.
Figure A3. E/W (a) and N/S (b) deformation fields after removing the orbital error. The black ellipse lines circle the area of dominated TSA in E/W and N/S deformation fields. The symbol (+) indicates eastward and northward offsets of TSA and the symbol (−) indicates westward and southward offsets of TSA.
Table A2. Relationship between the TSA and Sun angles of Landsat 8 correlation images pairs.

<table>
<thead>
<tr>
<th>Pairs Num.</th>
<th>Image Pair</th>
<th>Pre. (°)</th>
<th>Pos. (°)</th>
<th>Δ = Pre. − Pos.</th>
<th>Measured TSA</th>
<th>Theoretical Calculated Values of Vector Offset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sun Azimuth</td>
<td>Sun Elevation</td>
<td>Sun Azimuth</td>
<td>Sun Elevation</td>
<td>N/S Offset Orientation (°)</td>
</tr>
<tr>
<td>1</td>
<td>21 May 2013–6 June 2013</td>
<td>101.360</td>
<td>69.407</td>
<td>94.267</td>
<td>69.635</td>
<td>7.093</td>
</tr>
<tr>
<td>2</td>
<td>10–26 September 2013</td>
<td>131.684</td>
<td>60.502</td>
<td>141.144</td>
<td>56.521</td>
<td>−9.460</td>
</tr>
<tr>
<td>3</td>
<td>26 September 2013–12 October 2013</td>
<td>141.144</td>
<td>56.521</td>
<td>148.360</td>
<td>51.863</td>
<td>−7.216</td>
</tr>
<tr>
<td>4</td>
<td>12–28 October 2013</td>
<td>148.360</td>
<td>51.863</td>
<td>153.179</td>
<td>47.062</td>
<td>−4.819</td>
</tr>
<tr>
<td>5</td>
<td>29 November 2013–15 December 2013</td>
<td>156.360</td>
<td>38.990</td>
<td>155.384</td>
<td>36.729</td>
<td>−0.976</td>
</tr>
<tr>
<td>6</td>
<td>22 April 2014–8 May 2014</td>
<td>118.468</td>
<td>64.727</td>
<td>108.668</td>
<td>67.655</td>
<td>9.799</td>
</tr>
<tr>
<td>7</td>
<td>8–24 May 2014</td>
<td>108.668</td>
<td>67.655</td>
<td>99.387</td>
<td>69.013</td>
<td>9.281</td>
</tr>
<tr>
<td>8</td>
<td>24 May 2014–9 June 2014</td>
<td>99.387</td>
<td>69.013</td>
<td>93.145</td>
<td>69.080</td>
<td>6.242</td>
</tr>
<tr>
<td>9</td>
<td>9–25 June 2014</td>
<td>93.145</td>
<td>69.080</td>
<td>91.468</td>
<td>68.439</td>
<td>1.677</td>
</tr>
<tr>
<td>10</td>
<td>13–29 September 2014</td>
<td>132.788</td>
<td>59.586</td>
<td>141.927</td>
<td>55.520</td>
<td>−9.139</td>
</tr>
<tr>
<td>11</td>
<td>29 September 2014–15 October 2014</td>
<td>141.927</td>
<td>55.520</td>
<td>148.802</td>
<td>50.866</td>
<td>−6.873</td>
</tr>
<tr>
<td>12</td>
<td>16 November 2014–2 December 2014</td>
<td>153.538</td>
<td>41.757</td>
<td>155.921</td>
<td>38.370</td>
<td>−0.364</td>
</tr>
<tr>
<td>13</td>
<td>3–19 January 2015</td>
<td>152.443</td>
<td>36.075</td>
<td>149.367</td>
<td>37.544</td>
<td>3.076</td>
</tr>
<tr>
<td>14</td>
<td>20 February 2015–8 March 2015</td>
<td>141.845</td>
<td>45.076</td>
<td>137.384</td>
<td>39.341</td>
<td>4.461</td>
</tr>
<tr>
<td>15</td>
<td>9–25 April 2015</td>
<td>125.252</td>
<td>65.146</td>
<td>116.927</td>
<td>65.146</td>
<td>8.324</td>
</tr>
</tbody>
</table>

Notes: It is difficult to quantify the magnitude of the TSA in the correlation results due to their distribution characteristics, so we only make an orientation statistics of the shadowing offsets. The symbol (+) indicates eastward and northward offsets, correspondingly the symbol (−) indicates westward and southward offsets. These blank represent that the positive or negative characteristics of the measured errors are so tiny to identify. The theoretical calculated values, in which the offset orientation and offset length are equal to the orientation and magnitude of the “radiometric baseline”, are calculated using a trigonometry based on the sun angle information as shown in Figure 9. We found that the measured orientations of TSA have a good consistency with the theoretical calculated offset orientations.
References


5. Xu, B.; Li, Z.; Feng, G.; Zhang, Z.; Wang, Q.; Hu, J.; Chen, X. Continent-wide 2-D co-seismic deformation of the 2015 Mw 8.3 Illapel, Chile earthquake derived from Sentinel-1A data: Correction of azimuth co-registration error. Remote Sens. 2016, 8. [CrossRef]


19. Dominguez, S.; Avouac, J.; Michel, R. Horizontal coseismic deformation of the 1999 Chi-Chi earthquake measured from SPOT satellite images: Implications for the seismic cycle along the western foothills of Central Taiwan. J. Geophys. Res. Solid Earth 2003, 108. [CrossRef]


38. Hu, J.; Wang, Q.; Li, Z.; Zhao, R.; Sun, Q. Investigating the ground deformation and source model of the yangbajing geothermal field in Tibet, China with the WLS InSAR technique. *Remote Sens.* **2016**, *8*. [CrossRef]


Article

Second-Order Polynomial Equation-Based Block Adjustment for Orthorectification of DISP Imagery

Guoqing Zhou 1,*, Tao Yue 1,*, Yujun Shi 2, Rongting Zhang 3 and Jingjin Huang 3

1 Guangxi Key Laboratory for Geospatial Informatics, Guilin University of Technology, No. 12, Jian’gan Road, Qixing District, Guilin 541004, China
2 Guangxi Institute of Surveying and Remote Sensing Information, No. 5, Jianzheng Road, Qingxiu District, Nanning 530023, China; shiyujun1988@163.com
3 The Center for Remote Sensing, Tianjin University, No. 92, Weijin Road, Nankai District, Tianjin 300072, China; zrt65@tju.edu.cn (R.Z.); jingjin_huang@163.com (J.H.)

* Correspondence: gzhou@glut.edu.cn (G.Z.); yuetao@glut.edu.cn (T.Y.);
Tel.: +86-773-589-6073 (G.Z.); +86-773-589-1963 (T.Y.)

Academic Editors: Zhenhong Li, Roberto Tomas, Zhong Lu, Richard Gloaguen and Prasad S. Thenkabail

Received: 4 July 2016; Accepted: 17 August 2016; Published: 22 August 2016

Abstract: Due to the lack of ground control points (GCPs) and parameters of satellite orbits, as well as the interior and exterior orientation parameters of cameras in historical declassified intelligence satellite photography (DISP) imagery, a second order polynomial equation-based block adjustment model is proposed for orthorectification of DISP imagery. With the proposed model, 355 DISP images from four missions and five orbits are orthorectified, with an approximate accuracy of 2.0–3.0 m. The 355 orthorectified images are assembled into a seamless, full-coverage mosaic image map of the karst area of Guangxi, China. The accuracy of the mosaicked image map is within 2.0–4.0 m when compared to 78 checkpoints measured by Real-Time Kinematic (RTK) GPS surveys. The assembled image map will be delivered to the Guangxi Geological Library and released to the public domain and the research community.

Keywords: declassified intelligence satellite photography (DISP); rectification; imagery distortion; second polynomial equation

1. Introduction

Rocky karstification in karst areas (also called karst rocky desertification (KRD)) is considered one of the major factors that contribute to the global carbon balance as a global CO2 sink [1–3]. With the increasing interest in global carbon emissions, studies and analyses have compared historical data with current data to discover how rocky karstification contributes to long-term environmental changes over decadal spans.

Guangxi is located in the southwestern karst area in China, and the KRD area is approximately 23,790.80 km², accounting for 19.8% of the total KRD area in China in 2005. It shrunk to 19,260.00 km² in 2011, accounting for 16.0% of the total KRD area in China. Although many researchers have investigated the Guangxi KRD area associated with its environmental evolution in recent decades, there have been no investigations or analyses of the KRD area that focused on the early 1960s. Fortunately, declassified intelligence satellite photography (DISP) released to the public domain in February 1995 has provided researchers with a unique opportunity to investigate the KRD in Guangxi in the 1960s. The DISP was collected by the first generation of United States photoreconnaissance satellites between 1960 and 1972 through the systems named CORONA, ARGON, and LANYARD. More than 860,000 images of the Earth’s surface were declassified with the issuance of this executive order and were contracted to the USGS for sale.
However, further processing and application of DISP has resulted in various problems:

(1) The USGS does not provide Chinese users with the parameters required to further process DISP. These parameters include satellite orbit parameters (e.g., inclination, flight height, descent time, etc.) and the camera’s interior orientation parameters (IOP) (e.g., focal length, principal point coordinates, fiducial marks, etc.). This implies that traditional bundle block adjustment based on the photogrammetric collinearity equation is not applicable [4,5].

(2) It is very difficult to obtain sufficient ground control points (GCPs) in the historical DISP imagery due to the time intervals of several decades and cloudy coverage in Southern China. Thus, it is almost impossible to rectify each DISP image on a frame-by-frame base.

For the two reasons above, this paper presents a second order polynomial equation-based rectification model for orthorectification of DISP images. The previous relevant studies on this topic are as follows: Kim et al. utilized a collinearity equation to rectify ARGON imagery from 1963 to study the seasonal variations of glaciers on the Queen Maud Land coast of Antarctica [6]. Zhou and Jezek proposed a collinearity equation-based self-calibration block bundle adjustment method that integrates the bundle adjustment method and satellite orbital parameters, solving interior orientation (including lens distortion) and exterior orientation parameters (EOPs) simultaneously to rectify ARGON images from 1962 and 1963 [4]. The rectified ARGON imagery was employed to mosaic Greenland ice sheets from the 1960s, which were then quantitatively compared to the ice sheet extent over a 30-year interval [5]. Kim and Jezek applied a state-of-the-art digital imaging technology based on an extended block adjustment to rectify ARGON imagery from 1963 that covered Antarctica [7]. They assembled all images into a quality mosaic of coastal Antarctica to study glaciers. In addition, due to the imaging model limitations of high-resolution satellites, such as IKONOS, rational polynomial-based block adjustment, also called rational polynomial coefficient (RPC), was proposed by multiple authors. For example, Tao et al. analyzed the accuracy of orthorectification of a Systeme Probatoire d’Observation de la Terre (SPOT) image and an aerial image using the RPC model [8,9]. Yang suggested that the RPC model can replace the rigorous sensor model orthorectification of SPOT images [10]. Liu developed a stereotaxic method of IKONOS images based on the RPC model [11]. Huang proposed a rational polynomial-based block adjustment for orthorectification of Synthetic Aperture Radar (SAR) images [12]. Grodecki and Gene Dial rectified IKONOS satellite imagery using the RPC method. The RPC model incorporates a priori constraints into the images described by RPC, and multiple independent images can be added in accordance with the needs of users [13]. However, the RPC model requires a number of GCPs, and the computation is very time consuming. Therefore, the RPC method is not applicable to DISP images that the USGS provides because the imaging model of DISP was not provided by the USGS. Additionally, few GCPs are available in the study area. Thus, this paper presents an effective and simple mathematical model for geometric rectification of DISP images, considerably improving the computational effectiveness.

2. The Second Order Polynomial Equation-Based Rectification Model Method

2.1. Polynomial Equation-Based Block Adjustment Model

The objective of polynomial equation-based block adjustment is to tie overlapping images together without the absolute need for ground control points in each image and obtain coordinates of tie points and conversion parameters for rectification. Since the study area is a karst landform with a large wavy terrain and large elevation differences, the relief displacement is large. For correction of relief displacement, relief displacement is introduced into the block adjustment model shown below.

Figure 1 shows the imaging geometry of DISP from the CORONA mission. \( S - W V U(W^V U') \) is a camera coordinate system, \( o - xy \) is an image plane system, and \( O - XYZ \) is a geographic coordinate system. There is relief displacement (\( \Delta h \)) in the imaging process; therefore, the relief displacement must be corrected in the rectification process. First, distortion caused by elevation differences should
be corrected. Then, other distortions should be corrected by utilizing a polynomial model, and the reverse is true in the resampling process.

Figure 1. The imaging geometry of declassified intelligence satellite photography (DISP) from the CORONA mission.

Since the relief displacement only occurs in the direction of scanning, CORONA images are panoramic camera images, and the panoramic projection scan direction is the $x$-direction. Therefore, as shown in the imaging equation above, there is no relief displacement in the $y$-direction. The relief displacement correction functions are as follows:

$$\Delta h = Z \cdot h / M$$  \hspace{1cm} (1)

$$\begin{cases} 
\Delta x = x \cdot Z / M \\
\Delta y = y \cdot Z / M = 0
\end{cases}$$  \hspace{1cm} (2)

where $x$ and $y$ are image coordinates; $\Delta x$ and $\Delta y$ are image distortions in the $x$- and $y$-directions, respectively, caused by elevation differences; $Z$ is the elevation; $h$ is the distance from the image point to the nadir point; and $M$ is the satellite flight altitude. Since the relief displacement occurs in the direction of scanning and the KH-4A/B’s images are panoramic camera images, the images can be rectified using the second-order polynomial equation-based model.

2.1.1. Traditional Second-Order Polynomial Equation

The traditional second order polynomial model has been widely applied for image rectification. This paper extends the traditional equation into a block situation by adding tie points, which tie overlapping images together. With the extended model, the 2D coordinates of tie points and the coefficients of second order polynomial equations are solved. Furthermore, these parameters are used for orthorectification of DISP imagery without the absolute requirement of at least six GCPs in each DISP image.

The traditional second order polynomial equations are expressed as follows [14]:

$$\begin{cases} 
x + \Delta x = a_0 + a_1 x + a_2 y + a_3 xy + a_4 x^2 + a_5 y^2 \\
y + \Delta y = b_0 + b_1 x + b_2 y + b_3 xy + b_4 x^2 + b_5 y^2
\end{cases}$$  \hspace{1cm} (3)

where $\alpha = (a_0, a_1, a_2, a_3, a_4, a_5)^T$ and $\beta = (b_0, b_1, b_2, b_3, b_4, b_5)^T$ are coefficients; $x$ and $y$ are image coordinates; $\Delta x$ and $\Delta y$ are image distortions in the $x$- and $y$-directions, respectively; and $X$ and $Y$ are 2D coordinates in a given map coordinate system.
For a given GCP, Equation (3) can be linearized using a Taylor series and is expressed as follows:

\[
\begin{align*}
\Delta v_x &= \Delta a_0 + X \Delta a_1 + Y \Delta a_2 + XY \Delta a_3 + X^2 \Delta a_4 + Y^2 \Delta a_5 - l_x \\
\Delta v_y &= \Delta b_0 + X \Delta b_1 + Y \Delta b_2 + XY \Delta b_3 + X^2 \Delta b_4 + Y^2 \Delta b_5 - l_y
\end{align*}
\]  

(4)

where \(\Delta a_i(i = 0, 1, \cdots, 5)\) and \(\Delta b_i(i = 0, 1, \cdots, 5)\) are correction terms of coefficients; \(\Delta v_x, \Delta v_y\) are residuals; X and Y are 2D coordinates of GCPs; and \(l_x, l_y\) are constants expressed by Equation (5).

\[
\begin{align*}
l_x &= x - (a_0 + a_1 X + a_2 Y + a_3 XY + a_4 X^2 + a_5 Y^2) \\
l_y &= y - (b_0 + b_1 X + b_2 Y + b_3 XY + b_4 X^2 + b_5 Y^2)
\end{align*}
\]  

(5)

As shown in Equation (4), one GCP only establishes two observations, but Equation (4) has 12 unknown parameters. Therefore, six GCPs, which establish 12 observation equations, are needed to solve the 12 coefficients that are used for to rectify a single image. Generally, more than six GCPs are observed in each image to establish more than 12 observation equations. The least-squares estimation is employed to calculate the 12 coefficients. Mathematically, the solution can be described as follows.

Assuming that N GCPs \((N \geq 6)\) are observed, the observation equations are expressed in matrix form as follows:

\[
V = A \cdot \alpha - L
\]  

(6)

where:

\[
A = \begin{bmatrix}
1 & X_1 & Y_1 & X_1 Y_1 & X_1^2 & Y_1^2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & X_1 & Y_1 & X_1 Y_1 & X_1^2 & Y_1^2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & X_N & Y_N & X_N Y_N & X_N^2 & Y_N^2 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix},
\]

\[
V = \begin{bmatrix}
V_{x_1} & V_{y_1} & \ldots & V_{x_N} & V_{y_N}
\end{bmatrix}^T,
\]

\[
\alpha = \begin{bmatrix}
a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & b_0 & b_1 & b_2 & b_3 & b_4 & b_5
\end{bmatrix}^T
\]

and:

\[
L = \begin{bmatrix}
l_{x_1} & l_{y_1} & \ldots & l_{x_N} & l_{y_N}
\end{bmatrix}^T.
\]

The least-squares estimation, i.e., \(V^T PV = \min\), gives the solutions of the coefficients of the second order polynomial equation below:

\[
\alpha = (A^T A)^{-1} A^T L
\]  

(7)

We can further obtain the following expressions from Equation (7):

\[
a_i = a_i^0 + \sum_{j=1}^{N_{ite}} \Delta a_i^j \quad (i = 1, \cdots, 5; \; j = 1, \cdots, N_{ite})
\]  

(8)

\[
b_i = b_i^0 + \sum_{j=1}^{N_{ite}} \Delta b_i^j \quad (i = 1, \cdots, 5; \; j = 1, \cdots, N_{ite})
\]  

(9)

where \(a_i^0, b_i^0\) are initial values; \(\Delta a_i^j, \Delta b_i^j\) are increases during each iteration; and \(N_{ite}\) is the number of iterations.
2.1.2. The Second-Order Polynomial Equation-Based Rectification Model

As mentioned above, due to the shortage of GCPs in each of the DISP images, the tie points (TPs) must be identified to tie images with the same overlapping areas. Under this condition, the TPs whose XY-coordinates are unknown are introduced into the traditional second order polynomial equation. This extended model is called the second-order polynomial equation-based rectification model (2OPE-RM) in this paper (see Figure 1). Equation (3) is extended with considering TPs as unknown parameters and linearized into the following form:

\[
\begin{align*}
\begin{cases}
    v_x = & \Delta a_0 + X \Delta a_1 + Y \Delta a_2 + XY \Delta a_3 + X^2 \Delta a_4 + Y^2 \Delta a_5 + (a_1 + a_3 Y + 2a_4 X) \Delta X + (a_2 + a_3 X + 2a_5 Y) \Delta Y - l_x \\
    v_y = & \Delta b_0 + X \Delta b_1 + Y \Delta b_2 + XY \Delta b_3 + X^2 \Delta b_4 + Y^2 \Delta b_5 + (b_1 + b_3 Y + 2b_4 X) \Delta X + (b_2 + b_3 X + 2b_5 Y) \Delta Y - l_y
\end{cases}
\end{align*}
\]

(10)

Then, Equation (10) can be rewritten as follows:

\[
\begin{align*}
\begin{cases}
    v_x = & \Delta a_0 + X \Delta a_1 + Y \Delta a_2 + XY \Delta a_3 + X^2 \Delta a_4 + Y^2 \Delta a_5 + f_1 \Delta X + f_2 \Delta Y - l_x \\
    v_y = & \Delta b_0 + X \Delta b_1 + Y \Delta b_2 + XY \Delta b_3 + X^2 \Delta b_4 + Y^2 \Delta b_5 + g_1 \Delta X + g_2 \Delta Y - l_y
\end{cases}
\end{align*}
\]

(11)

where:

\[
\begin{align*}
    f_1 &= a_1 + a_3 Y + 2a_4 X, \\
    f_2 &= a_2 + a_3 X + 2a_5 Y, \\
    g_1 &= b_1 + b_3 Y + 2b_4 X, \text{ and} \\
    g_2 &= b_2 + b_3 X + 2b_5 Y.
\end{align*}
\]

The symbols above are the same as those in Equation (10).

Additionally, assuming that there are \(N\) GCPs (\(N \geq 6\)), \(M\) TPs in \(t\) images are collected at the GCPs. Similarly, Equation (10) can be expressed in matrix form as follows:

\[
V = A \cdot \alpha + B \cdot \beta - L
\]

(12)

where:

\[
V = \begin{bmatrix}
    v_{GCP}^1 & v_{GCP}^2 & \ldots & v_{GCP}^M \\
    v_{TP}^1 & v_{TP}^2 & \ldots & v_{TP}^M
\end{bmatrix}^T,
\]

\[
\alpha = (\Delta a_0^1, \Delta a_1^1, \ldots, \Delta a_4^1, \Delta a_5^1, \Delta a_0^2, \Delta a_1^2, \ldots, \Delta a_4^2, \Delta a_5^2),
\]

\[
\beta = (\Delta X_1^1, \Delta Y_1^1, \ldots, \Delta X_M^1, \Delta Y_M^1),
\]

\[
A = \begin{bmatrix}
1 & X_1^1 & Y_1^1 & X_1^1 Y_1^1 & X_1^1 Y_1^2 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 1 & X_{N_1} & Y_{N_1} & X_{N_1} Y_{N_1} & X_{N_1}^2 & Y_{N_1}^2 & \ldots & \ldots & \ldots \\
1 & X_1^2 & Y_1^2 & X_1^2 Y_1^2 & X_1^2 Y_1^2 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 1 & X_{N_2} & Y_{N_2} & X_{N_2} Y_{N_2} & X_{N_2}^2 & Y_{N_2}^2 & \ldots & \ldots & \ldots \\
1 & X_1^N & Y_1^N & X_1^N Y_1^N & X_1^N Y_1^2 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 1 & X_{N_N} & Y_{N_N} & X_{N_N} Y_{N_N} & X_{N_N}^2 & Y_{N_N}^2 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 1 & X_{N_N} & Y_{N_N} & X_{N_N} Y_{N_N} & X_{N_N}^2 & Y_{N_N}^2 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots 
\end{bmatrix}
\]

The 1st image with \(N_1\) GCPs

\[
\begin{bmatrix}
X_1^1 & Y_1^1 & X_1^1 Y_1^1 & X_1^1 Y_1^2 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
X_{N_1} & Y_{N_1} & X_{N_1} Y_{N_1} & X_{N_1}^2 & Y_{N_1}^2 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
X_1^2 & Y_1^2 & X_1^2 Y_1^2 & X_1^2 Y_1^2 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
X_{N_2} & Y_{N_2} & X_{N_2} Y_{N_2} & X_{N_2}^2 & Y_{N_2}^2 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
X_1^N & Y_1^N & X_1^N Y_1^N & X_1^N Y_1^2 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
X_{N_N} & Y_{N_N} & X_{N_N} Y_{N_N} & X_{N_N}^2 & Y_{N_N}^2 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 1 & X_{N_N} & Y_{N_N} & X_{N_N} Y_{N_N} & X_{N_N}^2 & Y_{N_N}^2 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots 
\end{bmatrix}
\]

The \(m\)th image with \(N_m\) GCPs

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & \ldots 
\end{bmatrix}
\]

The \(N\)th image with \(N_N\) GCPs
and:

\[
B = \begin{pmatrix}
    f_1^1 & f_2^1 \\
    \vdots & \vdots \\
    f_{M_1}^1 & f_{M_1}^2 \\
    f_1^1 & f_2^1 \\
    \vdots & \vdots \\
    f_{M_i}^1 & f_{M_i}^2 \\
    \vdots & \vdots \\
    f_{M_n}^1 & f_{M_n}^2 \\
\end{pmatrix}
\]

The 1st image with \(M_1\) TPs

The \(i\)th image with \(M_i\) TPs

The \(M\)th image with \(M_n\) TPs

Equation (12) is the 2OPE-RM model derived in this paper. Relative to the traditional model in Equation (6), this model introduces TPs as unknown parameters.

Equation (12) is usually solved using least-squares estimation, which is expressed as follows:

\[
\Phi = \mathbf{V}^T \mathbf{V} = \text{min}
\]

With least-squares estimation, the normal equation matrix can be written as follows:

\[
\begin{pmatrix}
    \mathbf{A}^T \mathbf{A} & \mathbf{A}^T \mathbf{B} \\
    \mathbf{B}^T \mathbf{A} & \mathbf{B}^T \mathbf{B}
\end{pmatrix}
\begin{pmatrix}
    \delta \mathbf{a} \\
    \delta \mathbf{b}
\end{pmatrix}
= \begin{pmatrix}
    \mathbf{A}^T \mathbf{L} \\
    \mathbf{B}^T \mathbf{L}
\end{pmatrix}
\]

Thus, the solution of the unknown parameters is given by Equation (15):

\[
\begin{align*}
\Delta \mathbf{a} &= -(Q_{aa} \mathbf{A}^T \mathbf{L} + Q_{ab} \mathbf{B}^T \mathbf{L}) \\
\Delta \mathbf{b} &= -(Q_{ba} \mathbf{A}^T \mathbf{L} + Q_{bb} \mathbf{B}^T \mathbf{L})
\end{align*}
\]

where \(Q_{ij}(i, j = 1, 2)\) gives the components of the covariance matrix, which is the inverse of the normal matrix, as shown in Equation (16):

\[
Q_{ij} = \begin{pmatrix}
    \mathbf{A}^T \mathbf{A} & \mathbf{A}^T \mathbf{B} \\
    \mathbf{B}^T \mathbf{A} & \mathbf{B}^T \mathbf{B}
\end{pmatrix}^{-1} = \begin{pmatrix}
    Q_{aa} & Q_{ab} \\
    Q_{ba} & Q_{bb}
\end{pmatrix} \quad (i, j = 1, 2)
\]

The coefficients of the 2OPE-RM in each image and the 2D coordinates (XY) of each TP are as follows:

\[
a_i = a_i^0 + \sum_{j=1}^{N_{ite}} \Delta a_i^j \quad (i = 1, \cdots, 5; \ j = 1, \cdots, N_{ite})
\]

\[
b_i = b_i^0 + \sum_{j=1}^{N_{ite}} \Delta b_i^j \quad (i = 1, \cdots, 5; \ j = 1, \cdots, N_{ite})
\]

\[
X_{ti}^i = X_i^0 + \sum_{j=1}^{N_{ite}} \Delta X_i^j \quad (i = 1, \cdots, 5; \ j = 1, \cdots, N_{ite}; \ t_i = 1, \cdots, t)
\]

\[
Y_{ti}^i = Y_i^0 + \sum_{j=1}^{N_{ite}} \Delta Y_i^j \quad (i = 1, \cdots, 5; \ j = 1, \cdots, N_{ite}; \ t_i = 1, \cdots, t)
\]

where \(X_i, Y_i\) are coordinates of the \(i\)-th TP in image \(t_i\); \(\Delta X_i, \Delta Y_i\) are increases in \(X_i\) and \(Y_i\); \(a_i^0\) and \(b_i^0\) are initial values; and \(\Delta a_i\) and \(\Delta b_i\) are increases in the coefficients in each iteration.

As shown in Equation (12), each image has 12 unknown parameters \((a_i, b_i; i = 0, 1, \cdots, 5)\), and each TP has two unknown parameters (XY-coordinates). Two equations can be established
for each GCP or TP. Moreover, the TPs and/or GCPs should be well distributed in each image. For example, there are four images, 12 GCPs, and nine TPs in Figure 2. The four images imply that there are 48 unknown parameters. The 12 GCPs can be used to establish 42 observation equations (i.e., seven GCPs in Image #1 can be used to establish 14 observations, three GCPs in Image #2 can be used to establish six observations, six GCPs in Image #3 can be used to establish 12 observations, and five GCPs in Image #4 can be used to establish 10 observations). The nine TPs can be used to establish 34 observation equations (i.e., three TPs in Image #1 can be used to establish six observations, three TPs in Image #2 can be used to establish six observations, six TPs in Image #3 can be used to establish 12 observations, and five TPs in Image #4 can be used to establish 10 observations). With this model, we have 76 (76 = 42 + 34) observations and 66 (66 = 48 + 18) unknown parameters. Thus, 2OPE-RM does not require each DISP image to have more than six GCPs.

The accuracy of the adjustment computation is evaluated using Equation (21):

$$\delta_o = \sqrt{\frac{V^TV}{r}}$$

where $\delta_o$ is the standard deviation of the unit weight, $V$ is the matrix of residuals, and $r$ is the number of redundant observations. Thus, the standard deviations of individual unknown parameters can be calculated as follows:

$$\delta_X = \delta_o \sqrt{Q_X}$$

To evaluate the accuracies of TPs, assuming that there are $n$ TPs, the average of $\delta_X$ is as follows:

$$\mu_X = \frac{1}{n} \sum \delta_X$$

where $n$ is the number of TPs.

Figure 2. Illustration of the second-order polynomial equation-based rectification model (2OPE-RM).

2.2. Orthorectification of DISP Images

With the established model and the coefficients determined in Section 2.1, each original DISP image can be orthorectified. The steps are as follows:

1. Step 1: Determination of the Rectified Image Size

To properly establish the storage space of the orthorectified image, the size of the resulting image (upper left, lower left, upper right, and lower right) must be determined in advance. This procedure is proposed as follows.
Determination of the four corner coordinates: The four corner coordinates of the original image are projected into the UTM coordinate system. Then, eight coordinates are obtained: \((X_{ul}, Y_{ul}), (X_{ll}, Y_{ll}), (X_{ur}, Y_{ur}), (X_{lr}, Y_{lr})\).

The maximum and minimum values of \(X\) and \(Y\) (\(X_{min}, X_{max}, Y_{min}, \) and \(Y_{max}\)) are calculated from the eight coordinates above to constitute four coordinate pairs. These pairs are the map coordinates of the four boundaries of the resulting image’s scope.

\[
X_{min} = \min(X_{ul}, X_{ll}, X_{ur}, X_{lr}), X_{max} = \min(X_{ul}, X_{ll}, X_{ur}, X_{lr}) \\
Y_{min} = \min(Y_{ul}, Y_{ll}, Y_{ur}, Y_{lr}), Y_{max} = \min(Y_{ul}, Y_{ll}, Y_{ur}, Y_{lr})
\]  

(24)

Determination of the resulting image’s size: The size of the resulting image can be determined by \(M\) and \(N\) as follows:

\[
M = \frac{Y_{max}}{Y_{GSD}} + 1, \quad N = \frac{X_{min}}{X_{GSD}} + 1
\]  

(25)

where \(M = \text{row}, N = \text{col},\) and \(Y_{GSD}, X_{GSD}\) are the ground-sampled distances (GSD) in the resulting image.

2. Step 2: Coordinate Transformation

Because the orthorectification model only expresses the relationship between the original coordinates \((x_{ori}, y_{ori})\) and ground coordinates \((X_{gro}, Y_{gro})\), the ground coordinates should be transformed into the coordinates of the resulting image \((x_{re}, y_{re})\) as follows:

\[
x_{re} = \frac{Y_{max} - Y_{gro}}{Y_{GSD}} + 1; \quad y_{re} = \frac{X_{gro} - X_{min}}{X_{GSD}} + 1
\]  

(26)

where \(Y_{gro}, X_{gro}\) are the ground coordinates of the pixel after rectification.

3. Step 3: Orthorectification

The calculation of the geographic coordinates of individual pixels, resampling of the original image, and registration of the chosen map coordinates system are carried out as follows:

- The process can be applied to any point \(P(I, J)\) in the resulting image with image coordinates \((I, J)\).
- In accordance with image coordinates \((I, J)\) and GSD, calculate the geographic coordinates \((X, Y)\).
- Compute the image coordinates \((i, j)\) of point \(P\) in the original image using Equation (5).
- Calculate the gray value \(g_{ori}\) via bilinear resampling interpolation.
- Assign the gray value \(g_{ori}\) to point \(P\) as \(g_{res}\) in the resulting (rectified) image/pixel.

The above procedure is then repeated for each pixel that must be rectified until the entire image is completely rectified.

2.3. Data Set

- Study area

The study area is located in Guangxi, China, spanning from 20.54°N to 26.24°N latitudes and 104.26°E to 112.04°E longitudes (Figure 3) and encompassing 23,790.8 km². The study area is in the south central subtropics of China.

308
Figure 3. Study area, which is located in Guangxi, China, with the encompassing 23,790.8 km².

- Dataset

DISP imagery: In total, 444 DISP images from five orbits of different missions, including the CORONA 1035-1 Mission (24 images) on 25 September 1966, the CORONA 1102-2 Mission (48 images) on 18 December 1967, and the 1106-1/2 Mission (39 images) on 7 February 1969, were purchased from the USGS (Figure 4).

Aerial photos: Five aerial photos with film formats of 18 × 18 cm² from 1961 were acquired at a photographic scale of 1:14,000. Each photo covers approximately 6.35 km². Five aerial photos were purchased from the Guangxi Bureau of Geospatial Information, China.

Coordinate data of GCPs: The coordinate data associated with GCPs in the KRD area were collected from Google Earth.

Figure 4. DISP image dataset (There are 444 DISP images from five orbits of different missions covering the whole study area).
3. Results and Accuracy Analysis

3.1. Image Preprocessing

The DISP film was scanned into digital images, producing film-grain noise and resulting in image quality degradation. Many noise filters have been used in the public domain. However, most of these approaches are either time consuming, because of complex modelling, or they erroneously remove geophysical features because of noise in the overall image. The filter algorithm developed by Zhou et al. was used to remove noise in this study [5]. One of the advantages of the algorithm is that it avoids the problems noted above because this approach performs statistical calculations within variable-size and variable-shape sub-windows (see Figure 5) that are determined individually for every pixel in the image, rather than modelling the noise in the overall image. The algorithm is briefly described as follows:

1. Select a window of $5 \times 5$ pixels.
2. Calculate the mean $n_i (i = 1,2...9)$ and variance $\alpha_i (i = 1,2...9)$ of nine masks.
3. Select one mask with the lowest variance $\alpha_k$ and mean $n_k$, and calculate the weights of every pixel within the $k^{th}$ mask using the following equation:
   $$\omega_i = e^{\left| \Delta_i \right|}, \Delta_i = \text{gray}_i - n_k$$  \hspace{1cm} (27)

4. Calculate the output using Equation (28):
   $$\text{gray}_{\text{output}} = \left( \frac{\sum_{i=1}^{M} W_i \cdot \text{gray}_i}{\sum_{i=1}^{M} \omega_i} \right)$$  \hspace{1cm} (28)

where $M$ is the number of pixels in the $k^{th}$ mask and $\text{gray}_i (i = 1,2...9)$ is the intensity.

With the filter algorithm above, the results of removing the DISP image noise are depicted in Figure 6, which demonstrates the effectiveness of the proposed approach.

---

**Figure 5.** The adaptive filter algorithm.
Figure 6. The results of noise removing: (a) The original image, and (b) The filtered image.

3.2. DISP Image Orthorectification and Accuracy Analysis

3.2.1. DISP Image Orthorectification

Since sufficient numbers of GCPs are not observed in each DISP image, TPs are identified to tie overlapping images together and solve for the coefficients of the 2OPE-RM. The study area consists of 355 DISP images (there are 444 DISP images in total, but we only employed 355 high-quality images). Thus, it is impractical to construct a block in the entire study area and then solve for the orthorectification parameters of all DISP images simultaneously because such a large block will produce a significantly large number of observation equations, resulting in a huge computational burden during matrix inversion. Therefore, this paper divides the study area into 24 blocks consisting of various DISP images (see Figure 7a). Each block was rectified independently. For example, Block 1 consists of nine images in Figure 7b, in which 20 GCPs and 29 tie points were identified and measured. The 20 GCPs are employed 36 times to establish 72 observation equations. The 29 TPs are employed 60 times to establish 120 observation equations. Thus, 192 observation equations (72 + 120 = 192) are established in Block 1. There are 166 unknown parameters (9 × 12 + 29 × 2 = 166). There are 26 redundant observations (i.e., 192 − 166 = 26), meeting the requirements of least squares adjustment.

With the 192 observation equations established using Equation (12); the parameters used to rectify the nine DISP images are solved simultaneously using Equation (15). The 2D coordinates of TPs are obtained using Equations (19) and (20). With the solved coefficients and TP coordinates in each image, orthorectification is performed for each DISP image at a GSD of 2.0 m. Figure 8a is part of one orthorectified DISP image.

The computational accuracies of TPs using the 2OPE-RM are evaluated by Equation (23). The standard deviations of TPs ($\mu_X$ and $\mu_Y$) are averagely 0.34 m and 0.23 m, respectively. In addition,
the “absolute” accuracy of the orthorectified aerial photo created in 1961 is calculated using the following equations:

\[
\Delta X_{\text{RMSE}} = \sqrt{\frac{\sum_{k=1}^{n}(X_k - x_k)^2}{n-1}} \tag{29}
\]

\[
\Delta Y_{\text{RMSE}} = \sqrt{\frac{\sum_{k=1}^{n}(Y_k - y_k)^2}{n-1}} \tag{30}
\]

where \(X_k\) and \(Y_k\) are XY-coordinates of TPs in the orthorectified DISP image, \(x_k\) and \(y_k\) are XY-coordinates in the orthorectified aerial photo created in 1961, and \(n\) is the total number of TPs. Using Equations (29) and (30), \(\Delta X_{\text{RMSE}}\) and \(\Delta Y_{\text{RMSE}}\) are 2.0 m and 1.6 m, respectively. These values are equivalent to approximately 2.0 pixels in the orthorectified DISP imagery.

Figure 7. The polynomial block adjustment: (a) Design of polynomial block adjustment in the entire area, which is divided into 24 blocks; (b) One block, which is used to explain establishment of the observation equations.

Figure 8. Accuracy verification using orthorectified aerial photos from 1961: (a) DISP image orthorectified using the proposed method; and (b) orthorectified aerial photo from 1961.
3.2.2. Accuracy Comparison Analysis

Accuracy comparison between the DISP images orthorectified using the traditional second-order polynomial model and the 2OPE-RM was conducted. Two test fields, which are located in mountainous and flat areas, were selected for the accuracy comparison.

(1) The Bameng Field is a mountainous area located in Bameng County to the west of the city of Baise, Guangxi, China, at 23.671°N to 24.135°N and 106.941°W to 107.698°W. This test area covers the entire DS1106-2119DF107a image. The maximum and minimum elevations are 1128 m and 790 m, respectively, above mean sea level (MSL). Therefore, the relief displacement is significant. There are 12 GCPs and seven TPs scattered throughout the test field. The 12 GCPs are used for second order polynomial equations to solve for the 12 rectification coefficients, and the 12 GCPs and seven TPs are used in the 2OPE-RM to calculate the coefficients. Twenty-three checkpoints were chosen to evaluate the achievable accuracy. The orthorectified aerial photo provided by the Bureau of Guangxi Geomatics and Geographic Information is considered to represent the “true” values for validation. The results are listed in Table 1.

(2) The Longzhou field is a flat area located in Longzhou County to the west of the city of Chongzuo, Guangxi, China, at 22.105°N to 22.469°N and 106.593°W to 106.878°W. This test field completely covers the entire DS1106-2119DF110a image. In this test field, 11 GCPs and seven TPs are scattered throughout the DISP image. The same GCPs are employed in the traditional second order polynomial model and the 2OPE-RM. Twenty-three checkpoints were chosen to evaluate the accuracy. The planimetric accuracies of the two models relative to the orthorectified aerial image are shown in Table 1.

<table>
<thead>
<tr>
<th>Models</th>
<th>Test Area (Located in A Mountainous Area)</th>
<th>Test Area (Located in A Flat Area)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bameng Field</td>
<td>Longzhou Field</td>
</tr>
<tr>
<td></td>
<td>( \Delta X_{\text{RMSE}} ) (m)</td>
<td>( \Delta X_{\text{RMSE}} ) (m)</td>
</tr>
<tr>
<td></td>
<td>( \Delta Y_{\text{RMSE}} ) (m)</td>
<td>( \Delta Y_{\text{RMSE}} ) (m)</td>
</tr>
<tr>
<td>Second order polynomial model</td>
<td>1.96</td>
<td>1.85</td>
</tr>
<tr>
<td></td>
<td>1.84</td>
<td>1.57</td>
</tr>
<tr>
<td>Our model</td>
<td>1.85</td>
<td>1.67</td>
</tr>
<tr>
<td></td>
<td>1.69</td>
<td>1.49</td>
</tr>
</tbody>
</table>

3.3. Image Mosaicking

Based on the individual image orthorectification above, the next work is to mosaic the individual orthorectified DISP images into an image map. First, the characteristics of the study area and DISP images must be understood:

1. The study area covers 23,790.8 km² (between latitudes 20.54°N and 26.24°N and longitudes 104.26°E and 112.04°E), which consists of 355 DISP images that total 100 GB. A good mosaicking scheme may save computational time and computer storage;
2. The study area is located in a karst landscape, where mountainous and hilly terrain areas account for two-thirds of the total area;
3. The overlap between neighboring images must be less than 30%; and
4. The study area is covered by five strips of DISP images from four missions (Figure 9).

To minimize the influence of error propagation and avoid repeatedly sampling images, based on the characteristics above, the mosaicking is designed as follows (see Figure 9):

1. The 16 DISP images from Mission 1106 were first mosaicked, covering the western portion of the study area. The mosaicked map is depicted in Figure 10a. Twenty DISP images from Mission 1102-2 were mosaicked, and the mosaicked map is depicted in Figure 10b. Twenty-eight DISP
images from Mission 1106 were mosaicked, and the mosaicked map is depicted in Figure 10c. Twenty-three DISP images from Mission 1106 were mosaicked, as the mosaicked map is depicted in Figure 10d. Finally, 18 DISP images from Mission 1106 were mosaicked, and the mosaicked map is depicted in Figure 10d, covering the eastern portion of the area; and

2. With the five mosaicked maps above, a map image of the entire study area was assembled by merging the five mosaicked images. The order of mosaicking is from the east and west to the middle of the study area (see Figure 10).

3.4. Radiance Balance

Due to the differences in the imaging date/time and different imaging conditions during different missions, brightness differences between neighboring strips are unavoidable. In addition, patchwork lines are also unavoidable. To produce a seamless mosaic of the entire study area, this paper used a histogram equalization method to adjust the brightnesses of two neighboring strips. The boundary line was chosen along the center image, and overlapping areas were feathered. Figure 11 shows the result of the radiance balance.
3.5. Mosaic Result and Accuracy Evaluation

The entire study area has been mosaicked by 355 orthorectified DISP images (Figure 12a). A mountainous area located in Du’an County (Figure 12b) and a flat area located in Xingbin County (Figure 12c) are select as the samples for accurate validation. Seventy-eight GCPs, which were measured by RTK GPS measurements, are uniformly distributed in other countries throughout the entire study area. These include 25 check points (CPs) scattered throughout the two test fields. $\Delta X_{RMSE}$ and $\Delta Y_{RMSE}$ in Equations (26) and (27) are used to measure the accuracy. The results are listed in Table 2. As shown in Table 2, the accuracy in flat areas is better than that in mountainous areas, and the overall accuracies of the entire study area are 2.11 m and 1.74 m.

Table 2. Final accuracies of the assembled DISP image map in the study area.

<table>
<thead>
<tr>
<th>Area</th>
<th>$\Delta X_{RMSE}$ (m)</th>
<th>$\Delta Y_{RMSE}$ (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mountainous area</td>
<td>2.07</td>
<td>1.60</td>
</tr>
<tr>
<td>Flat area</td>
<td>1.86</td>
<td>1.79</td>
</tr>
<tr>
<td>The entire study area</td>
<td>2.11</td>
<td>1.74</td>
</tr>
</tbody>
</table>
4. Discussions

The results of the accuracy comparison for the DISP images orthorectified by the 2OPE-RM and by traditional second order polynomial model [14] are listed in Table 1. As observed from Table 1, it is demonstrated that the RMSEs of XY-coordinates in the DISP images orthorectified by the 2OPE-RM are smaller than those orthorectified by traditional second order polynomial model in both mountainous and flat areas. With the experimental result, it can be concluded that:

(1) The proposed 2OPE-RM method can successfully solve the problems below when orthorectifying the DISP images that:

(a) Each of the original DISP image has insufficient GCPs;
(b) The camera’s imaging model is unknown; and
(c) The camera’s interior orientation parameters (IOPs) including camera’s principal point coordinates, focal length, and lens distortion parameters are unknown.

(2) The proposed 2OPE-RM is capable of obtaining a higher accuracy than the traditional second order polynomial method does when orthorectifying the images under the above conditions (see Table 2).
Although the proposed 2OPE-RM method is experimented and validated on the DISP images with a satisfied accuracy, it should be suitable for the high-resolution of satellite images whose imaging model and whose camera IOPs are not released. The major limitations of the proposed 2OPE-RM method include:

(a) The proposed method needs a lot of tie points, which tie all images together. As observed in Table 1, the accuracy of images rectified by the proposed method can be increased with increasing the number of TPs. For example, if only the 12 GCPs are used in Bameng Field, the RMSEs of the XY-coordinates are 1.96 m and 1.84 m, respectively. If seven TPs are added in addition to 12 GCPs, the RMSEs of XY-coordinates reach 1.85 m and 1.69 m, respectively. The accuracy of the rectification result has been improved. Thus, the more the tie points, the higher the accuracy of orthorectification.

(b) The proposed method is time-consuming and labor-intensive, because a lot of tie points, which are usually feature points in images, are manually selected and measured. Although a semi-automation of measurement and selection of TPs are used in this paper, a zoom-in window operation for high-accuracy of location of the TPs is usually employed.

5. Conclusions

This paper presents a highly effective, simple, practical mathematical model for the orthorectification of CORONA DISP images from the 1960s, whose interior and exterior parameters are unknown and in which GCPs are lacking. The model is called the second order polynomial equation-based block rectification model (2OPE-RM). With the proposed model, all images can be orthorectified at an accuracy level of 2.0 pixels, corresponding to approximately 2.0–4.0 m with respect to the WGS 84 datum. All of the images covering the entire karst area of Guangxi, China, are assembled into a high-quality image map. The sampled distance of the assembled mosaicking map is 2.0 m. The proposed model can solve the problems associated with the traditional second order polynomial model, such as lack of GCPs, yielding acceptable and improved accuracy. The assembled image map of the entire rock desertification area in Guangxi, China, will be delivered to the Guangxi Geological Library for use by the research community.

Acknowledgments: This paper is financially supported by the National Natural Science of China under Grant numbers 41431179 and 41162011, Guangxi Natural Science Foundation under grant numbers 2015GXNSFDA139032, Guangxi Science & Technology Development Program under the contract number GuiKeHe 14123001-4, and GuangXi Key Laboratory of Spatial Information and Geomatics Program (Contract No. GuiKeNeng 14052401, 14052409, 151400701 and 151400712), the “BaGuiScholars” program of the provincial government of Guangxi.

Author Contributions: Guoqing Zhou contributes most to this manuscript. He has made the main contribution to the whole idea and writing the initial draft for this manuscript. Tao Yue contributed to re-organization of the framework and the multiple reviewing of the manuscript. Yujun Shi supported data processing and analysis. Rongting Zhang and Jingjin Huang contributed to editing and reviewing the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCP</td>
<td>Ground Control Point</td>
</tr>
<tr>
<td>DISP</td>
<td>Declassified Intelligence Satellite Photography</td>
</tr>
<tr>
<td>KRD</td>
<td>Karst Rocky Desertification</td>
</tr>
<tr>
<td>IOP</td>
<td>Interior Orientation Parameter</td>
</tr>
<tr>
<td>EOP</td>
<td>Exterior Orientation Parameter</td>
</tr>
<tr>
<td>RPC</td>
<td>Rational Polynomial Coefficient</td>
</tr>
<tr>
<td>2OPE-RM</td>
<td>Second Order Polynomial Equation-Based Rectification Model</td>
</tr>
<tr>
<td>GSD</td>
<td>Ground Sampled Distance</td>
</tr>
<tr>
<td>TP</td>
<td>Tie Point</td>
</tr>
</tbody>
</table>
References


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
Taking Advantage of the ESA G-POD Service to Study Ground Deformation Processes in High Mountain Areas: A Valle d’Aosta Case Study, Northern Italy

Martina Cignetti 1,*, Andrea Manconi 2, Michele Manunta 3, Daniele Giordan 1, Claudio De Luca 3, Paolo Allasia 4 and Francesca Ardizzone 4

1 National Research Council of Italy, Research Institute for Geo-Hydrological Protection (CNR-IRPI), Strada delle Cacce 73, 10135 Torino, Italy; daniele.giordan@irpi.cnr.it (D.G.); paolo.allasia@irpi.cnr.it (P.A.)
2 Departement of Earth Sciences, Engineering Geology, ETH Zürich, Sonneggstrasse 5, NO G65, CH-8092 Zürich, Switzerland; andrea.manconi@erdw.ethz.ch
3 National Research Council of Italy, Istituto per il Rilevamento Elettromagnetico dell’Ambiente (CNR-IREA), via Diocleziano 328, 80124 Napoli, Italy; manunta.m@irea.cnr.it (M.M.); deluca.c@irea.cnr.it (C.D.L.)
4 National Research Council of Italy, Research Institute for Geo-Hydrological Protection (CNR-IRPI), Via della Madonna Alta 126, 06128 Perugia, Italy; francesca.ardizzone@irpi.cnr.it

* Correspondence: martina.cignetti@irpi.cnr.it; Tel.: +39-011-397-7831

Academic Editors: Roberto Tomas, Zhenhong Li and Prasad S. Thenkabail

Received: 15 July 2016; Accepted: 11 October 2016; Published: 20 October 2016

Abstract: This paper presents a methodology taking advantage of the GPOD-SBAS service to study the surface deformation information over high mountain regions. Indeed, the application of the advanced DInSAR over the arduous regions represents a demanding task. We implemented an iterative selection procedure of the most suitable SAR images, aimed to preserve the largest number of SAR scenes, and the fine-tuning of several advanced configuration parameters. This method is aimed at minimizing the temporal decorrelation effects, principally due to snow cover, and maximizing the number of coherent targets and their spatial distribution. The methodology is applied to the Valle d’Aosta (VDA) region, Northern Italy, an alpine area characterized by high altitudes, complex morphology, and susceptibility to different mass wasting phenomena. The approach using GPOD-SBAS allows for the obtainment of mean deformation velocity maps and displacement time series relative to the time period from 1992 to 2000, relative to ESR-1/2, and from 2002 to 2010 for ASAR-Envisat. Our results demonstrate how the DInSAR application can obtain reliable information of ground displacement over time in these regions, and may represent a suitable instrument for natural hazards assessment.

Keywords: high mountain regions; ground surface deformation; DInSAR; GPOD-SBAS

1. Introduction

Glaciers and paraglacial processes, as landslides and rock glaciers, play an important role in the short- and long-term landscape evolution of high mountain regions [1–3]. Changes in climatic conditions, as well as human activity, can have a severe influence on such geomorphological processes, and increase the frequency and intensity of catastrophic mass wasting events [4,5]. In alpine regions, glaciation has been recognized as one of the most important agents influencing the regional geomorphology [6,7]. Indeed, several geological and geomorphological studies have outlined the relation between large slope instabilities and deep-seated gravitational slope deformation located in glacial valleys, likely due to the debuttressing caused by glacier retreat [8–11].

In this scenario, monitoring surface deformation in high mountain regions can help to better interpret the kinematic behavior of mass wasting processes, and to infer potential influences of
meteo-climate variables on their evolution [12–14]. This information is crucial to increase our capability to assess landslide hazards, and to manage the associated risks [15–18]. In recent decades, surface deformation in mountain environments has been measured by means of diverse monitoring approaches, ranging from classical in situ instruments [19–24] to more advanced remote sensing methods [25–28]. In situ instruments are typically installed to obtain a time series with high temporal resolution, which allows for a better analysis of local phenomena over time, while remote sensing techniques are more suitable for the evaluation of deformation trends at regional scales. Among different techniques, space-borne differential synthetic aperture radar (SAR) interferometry (DInSAR) has gained an important role in measuring ground displacements over the last few years [29]. DInSAR is a consolidated method used to provide data with extensive spatial coverage and to investigate different types of phenomena simultaneously, including those hindered by limited or difficult access. For example, DInSAR was successfully applied to detect and monitor slow movements of mountain slopes in the order of few centimeters/year in the Austrian and Swiss Alps, and the Pyrenees, confirming its operational potential in high mountain areas [30–32].

Advanced DInSAR methods have been developed to derive ground velocity maps and displacement time series [33–35]. The small baseline subset (SBAS) technique [34], which combines sets of interferograms with small orbital separation (baseline) and short revisiting time to reduce the temporal decorrelation and maximize the number of coherent SAR targets, has proven to be suitable in different deformation scenarios [36–39]. Accuracies are in the order of 1 mm/year for mean surface velocities and 5 mm for displacement measurements [40].

Nowadays, the increasing availability of multi-temporal satellite acquisitions allows for the generation of time series of ground deformation spanning periods as long as 20 years. This information is particularly suitable for studying the long-term behavior of landslides, rock glaciers, and glaciers. However, the analysis of surface deformation via DInSAR is challenging for several reasons, including (i) the high topographic gradients associated with the complex orography; (ii) abundant vegetation affecting the temporal correlation of the SAR signal; and (iii) unsuitable valley flank orientations relative to the SAR view angle [41–44]. Moreover, DInSAR ground deformation monitoring for systematic analyses is complicated due to the intrinsic limitations of the technique (i.e., coherence loss due to large revisit time, phase decorrelation due to large or rapid displacement, and line-of-site (LOS) measurements only) [33,45–47]. Additional issues are caused by the atmospheric phase screen (APS), which is amplified by high topographic gradient in high mountain regions. The APS, caused by the atmospheric pressure, temperature, and water vapor variations between two SAR acquisitions, may cause artifacts on the surface deformation results [48]. Moreover, SAR images acquired in winter periods are highly affected by snow cover, which causes temporal coherence loss [47,49,50].

All these factors have to be carefully taken into account for SAR data processing to thoughtfully interpret the ground deformation in high mountain regions. In several cases, however, the complexity of SAR data processing, as well as the large number of attempts to be performed before obtaining reliable results, hinders the achievement of surface deformation results. Currently, the application and the treatment of SAR data in high mountain regions requires specific background knowledge of the user and demand the application of complex processing algorithms and software. Moreover, the increasing amount of SAR data available from different satellites missions leads to rising needs of storage and computing resources. Recently, a relevant service was released within the ESA GRID-based operational environment [51], i.e., the unsupervised implementation of the Parallel-SBAS (P-SBAS) algorithm [52]. GPOD is coupled with high-performance and sizeable computing resources managed by GRID technologies, and it provides flexibility for building an application virtual environment with quick accessibility to data, computing resources, and results. The users access various services useful in the EO applications through a web interface, which guides the users from the creation of a new task until the results publication, passing through the data selection and the job monitoring. The P-SBAS algorithm has been implemented to exploit the GPOD resources, and to process the SAR data archived by ESA to perform the full SBAS-DInSAR processing chain in an unsupervised fashion from the
Level 0 (raw) SAR data to the generation of surface velocity maps and displacement time series [53]. Since the average processing time is smaller than 24 h, several processing attempts can be performed in a reasonable time-span to optimize the final results on a specific target area. The main advantage of this service, hereafter referred to as GPOD-SBAS, is that users with different backgrounds, more interested in the analysis of the surface deformation due to different natural and/or anthropic processes than in the details of the SAR data processing, can apply a reliable and validated state of the art algorithm without downloading large amounts of input data, nor owning costly processing hardware or learning specific software [53].

In this study, we define best practices to obtain surface displacement time series in high mountain regions by using the GPOD-SBAS service. Our work derives from a specific need: even though the unsupervised GPOD-SBAS service provides a great advance and allows a larger number of users to carry out advanced DInSAR analyses, obtaining reliable results in mountain areas is not a trivial task. The proposed methodology is applied to the Valle d’Aosta (VDA) region, Northern Italy, an alpine area characterized by high altitudes, complex morphology, and susceptibility to different mass wasting phenomena (e.g., landslides, rock glaciers, and glaciers), in some cases threatening urban areas and human activities, and it can be easily borrowed in similar areas.

The paper is organized as follows. We first introduce the area of study by highlighting the geomorphological, environmental, and climatic characteristics. Then, we describe the method applied to analyze, through the use of the GPOD-SBAS service, the data acquired by ERS-1/2 and ASAR-Envisat during the 1992–2000 and 2002–2010 periods, respectively, in ascending orbit for ERS-1/2, and in ascending and descending orbit for ASAR-Envisat. Finally, we present the results, discuss our findings, and propose a number of suggestions and best practices to be considered to retrieve effective and reliable information while approaching the analysis of SAR data in high mountain areas.

2. The Valle d’Aosta Region, Northern Italy

The Valle d’Aosta region (3200 km²) is located in the Northwestern Italian Alps (Figure 1). More than 50% of the Valle d’Aosta (VDA) territory has an elevation above 2000 m a.s.l., though topographic relief is highly variable, ranging from 300 m a.s.l. to peaks higher than 4000 m a.s.l. (e.g., Mt. Bianco, 4810 m a.s.l.; Mt. Rosa, 4635 m a.s.l.; Cervino, 4478 m a.s.l.; Gran Paradiso, 4061 m a.s.l.).

![Figure 1. Relief terrain of the study area (the Valle d’Aosta (VDA) region, Northwestern Italy). The map shows the high and complex morphology of this mountainous area; the light blue polygons correspond to the present glaciers of the region, modified from the Glacier map of Fondazione Montagna Sicura [54].](image-url)

From a geological point of view, the VDA region is part of the alpine chain [55–59], passing through the principal Europe-vergent Austroalpine-Penninic structural domain of the Western Alps. The Valley
presents a complete section of the orogenic prism including (i) the Austroalpine domain; (ii) the ophiolitic Piedmont zone; and (iii) the Pennidic domain. These tectonic-metamorphic units composed of a complex pile of nappes, which present a post-collisional tectonic activity and a neo-tectonic dislocation system activation (e.g., semi-graben Aosta-Ranzola fault) [57]. The morphological traits of the VDA are characterized by the main valley trending east-west (Dora Baltea Valley). This is a wide (3400 km²) glacial valley with several tributary valleys (trending north-south), which ends around the Pont Saint Martin municipality, at the Ivrea moraine amphitheater, the main evidence of Quaternary glaciations [60].

As typical in alpine valleys, VDA has a prevalence of highly vegetated areas, while human settlement distribution is located at the valley bottom. The land cover is prevalently represented by forest, natural grassland, and rock outcrops with little or no vegetation. In addition, the massifs dominating the VDA landscape act as morphological boundaries and deeply influence the local climatic conditions. Indeed, the high peaks limit the access of air masses from the Mediterranean Sea or the Atlantic Ocean, causing a clear difference in the rain and snow precipitation regimes [61]. The regional climate, in mountainous and marginal sector, is characterized by wide range of temperatures and rainfall/snowfall (min, max and average values of 1000–1100 mm/year, while, in correspondence of the principal valley bottom, by a temperate climate, characterized by relatively lower rainfall (lower than 600 mm/year) [62].

2.1. Mass Movement in the Valle d’Aosta Region

The slopes of VDA are characterized by high density of mass movements, which can be synthetically grouped in glaciers, rock glaciers, and landslides. These phenomena are mainly related to the long- and short-term evolution of glacial, periglacial, and paraglacial processes [63,64]. Glaciers nowadays occupy as much as 5% of the regional territory, which represent one-third of the Italian glaciers [65]. The rate of ground deformation vary significantly passing from glacier flows, which is characterized by rates of tens to hundreds of meters per year, to permafrost creep, in order of cm to m per year [63,66,67]. The alpine permafrost has a very variable and fragmented distribution, showing creep deformation depending on permafrost temperature [68]. Rock glaciers (RGs), i.e., periglacial landforms common in high alpine environments, occupy about 2% of the VDA territory [69,70]. RGs are defined as lobate to tongue-shaped structure of frozen rock debris, with steep front and side slopes, interstitial ice cement, ice lenses, a core of massive ice, or a combination thereof. These ice-rocky structures may creep depending on the mountain permafrost condition. Due to the ice content, the RGs are classified as active, inactive, or relict [70–72]. Active RGs have higher ice content and move downslope with velocities ranging from several cm/year to several m/year [73]. Inactive RGs contain some ice, but do not move. Finally, relict RGs present a front slope less steep gradient and are frequently covered by vegetation, do not contain ice, and do not move. In VDA, 937 rock glaciers were identified, and most of them (56%) were classified as relict [74–76]. The intact RGs are located at a higher altitude, presenting a lower limit variable from 2126 m and 3150 m a.s.l., and a prevalently northeast-to-northwest exposure. The relict RGs are located not below 1600 m a.s.l. altitude, and reach on average at the upper limit of 2700 m a.s.l., presenting variable exposure.

Another important gravitational process that affects 18% of the VDA territory [77,78] are landslides. These phenomena vary in size and typology, from shallow landslides (soil slips, planar, and rotational sliding) to large slope instabilities. In the VDA region, deep-seated gravitational slope deformations (DsGSDs) affect at least 13.5% of the regional territory [77], for a total of 280 phenomena inventoried. These large slope instabilities are widespread phenomena in mountain areas of the entire European Alps [11,79] and represent a large mass movement of an entire high-relief flank valley [80]. Their evolution is controlled by the interaction of different factors: lithology, geology, climate weathering, seismicity, and deglaciation [11]. In most cases, DsGSDs are located in initially glacial valleys and related to Lateglacial trigger [80]; for these phenomena, the major cause of instability is ascribed to glacial debuttressing. These phenomena generally present very slow–slow deformation rates, variable from a few millimeters per year to a maximum of a few centimeters per year in uncommon cases [81].
2.2. Challenges for Retrieving Surface Deformation with DInSAR of VDA

The geomorphological and meteo-climatic characteristics make VDA a challenging environment for the analysis of surface deformation with DInSAR, including (i) a high topography gradient, (ii) complex orography, (iii) a predominance of forest or abundant vegetated areas, and (iv) unsuitable valley flank orientations, specifically due to the alternatively west- and east-facing orientation of the tributary valleys.

By assuming ERS-1/2 and ASAR-Envisat acquisitions, we aim at understanding the theoretical distribution of scatters in the VDA region. This analysis has been carried out by computing the R-Index [82], which considers slope and aspect angles of the ground surface, as well as the satellite’s incident acquisition geometry (incident angle, i.e., 23°, and an azimuth angle of the satellite orbit from north (clockwise), i.e., 345° in ascending orbit and 195° in descending orbit, respectively). The resulting map reveals, for the north–south oriented valleys, layover (thus less exploitable scatters) on the west-facing slope in ascending orbit (Figure 2a) and layover on the east-facing slope in descending orbit (Figure 2b). On the other end, east–west oriented valleys experience layover or foreshortening on the north-facing slopes in ascending orbit and the south-facing slope in descending orbit.

Figure 2. R-Index map (a) of the ascending orbit and (b) of the descending orbit for the VDA region. Layover/foreshortening cases are present for values lower than zero (blue portion).
Furthermore, land use is also an important factor that has effects on the SAR result distribution (Figure 3). The prevalence of natural and high-vegetated areas for the entire VDA regional territory generates temporal decorrelation effects [47]. In addition, temporal coherence loss may occur due to the meteo-climatic conditions of the investigated area, characterized by a prevalent snow cover during the winter season.

Figure 3. Land cover map of the VDA region (redraw from CORINE 2006).

3. Methods

By taking into account the challenges outlined above, here we present a methodology to exploit the capabilities and potentialities supplied by the GPOD-SBAS service to retrieve the surface deformation information over high mountains. The VDA region is a representative area of investigation where we apply GPOD-SBAS processing.

We consider SAR images collected by ERS-1/2, in ascending orbit, and ASAR Envisat missions from ascending and descending orbits (see Table 1). The tracks cover most of the VDA regional territory, excluding only about 10% of the entire region localized in the eastern portion (see Figure 4). The reference point is selected corresponding to an area assumed stable, on the basis of geological and geomorphological background analysis, during the investigation period, and where DInSAR coherence values are expected to be high (i.e., within the Aosta municipality, see Figure 4).

Table 1. Summary of the ERS-1/2 and ASAR Envisat images considered for the VDA study area.

<table>
<thead>
<tr>
<th>Satellite</th>
<th>Observation Period</th>
<th>Track</th>
<th>Frame</th>
<th>Orbit Acquisition</th>
<th>Number of Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERS-1/2</td>
<td>April 1995–May 2000</td>
<td>301</td>
<td>909</td>
<td>Ascending</td>
<td>32</td>
</tr>
<tr>
<td>ENVISAT</td>
<td>June 2004–October 2010</td>
<td>301</td>
<td>909</td>
<td>Ascending</td>
<td>50</td>
</tr>
<tr>
<td>ENVISAT</td>
<td>September 2004–October 2010</td>
<td>294</td>
<td>2673</td>
<td>Descending</td>
<td>46</td>
</tr>
</tbody>
</table>
Figure 4. Map of the VDA region study area (in light grey). The red box represents the extent of ERS-1/2 and ASAR Envisat track in ascending orbit, while the blue box represents the track in descending orbit. The reference point (REF) is located in the Aosta municipality.

Though the SBAS approach can be jointly applied to ERS-1/2 and ASAR Envisat datasets to retrieve a continuous time series covering almost 20 years [83], at the moment of the processing, the implementation available in GPOD allows us to process the datasets only separately. To retrieve reliable results and to maximize the spatial coverage of the achieved coherent points, we perform an iterative procedure in order to choose the most suitable SAR dataset, to minimize the temporal decorrelation effects mainly related to the presence of snow cover in the areas of interest, and to preserve the processing of the largest number of SAR scenes. We perform four tasks (named from T1 to T4) for each orbit, summarized in Table 2. It is worth noting that such an analysis phase, addressed to select the most suitable dataset, plays a key role when analyzing mountain areas. Indeed, the presence of several decorrelation sources, i.e., vegetation and snow cover, requires proper identification of the SAR acquisitions less affected by noise through an iterative mechanism. This analysis can be easily and efficiently carried out within the GPOD environment thanks to its capability to perform SBAS analyses in a reasonably short time (about 24 h). As a final remark, the following analysis is based on the two available Envisat datasets because they have a good number of SAR acquisitions (more than 50 acquisitions) uniformly distributed during the observed time period, and permit a better investigation of the impact of winter acquisition on the final results. As regards ERS-1/2 dataset, due to the limited number of available SAR acquisitions, it is directly processed by selecting the images and the processing parameters on the basis of the Envisat analyses. Finally, we advise that GPOD-SBAS operates with the three-arcsec DEM free available from the SRTM archive.
Table 2. Summary of the features of the attempts processed by the GPOD-SBAS service for the ASAR Envisat database in the VDA study area, also reporting the different values of the GPOD-SBAS processing parameters.

<table>
<thead>
<tr>
<th>Attempt</th>
<th>Satellite</th>
<th>Orbit</th>
<th>Images</th>
<th>Results (Number of Targets)</th>
<th>Increment from T1 (%)</th>
<th>Ground Pixel Size (m)</th>
<th>Coherence Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>ENV</td>
<td>ASC</td>
<td>50/50</td>
<td>4598</td>
<td>80</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DESC</td>
<td>46/46</td>
<td>2143</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>ENV</td>
<td>ASC</td>
<td>38/50</td>
<td>13343</td>
<td>193%</td>
<td>80</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DESC</td>
<td>38/46</td>
<td>5987</td>
<td>179%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>ENV</td>
<td>ASC</td>
<td>37/50</td>
<td>19011</td>
<td>313%</td>
<td>80</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DESC</td>
<td>37/46</td>
<td>9748</td>
<td>355%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>ENV</td>
<td>ASC</td>
<td>37/50</td>
<td>33086</td>
<td>620%</td>
<td>40</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DESC</td>
<td>37/46</td>
<td>20317</td>
<td>648%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In T1, we processed all the available images for each Envisat dataset by keeping the default parameters of the GPOD-SBAS unmodified; this task is very important because it allows us to produce a large number of interferograms involving all the available SAR scenes and to retrieve a complete overview of the whole SAR dataset. Moreover, the results achieved represent a backbone analysis carried out with the maximum number of available acquisitions; thus, they provide a deformation time series with the highest temporal sampling. However, the spatial distribution of resulting coherent targets is limited.

In order to increase our capacity for image surface deformation phenomena, in T2 we reprocessed only the images acquired from April to November without changing any other input or parameter, assuming that the snow season in VDA corresponds to the December–March period [84]. The approach followed in T2 can be particularly critical because it introduces very long temporal gaps in the SAR dataset during the winter periods and can strongly reduce the number of available acquisitions. Moreover, it does not guarantee the removal of all acquisitions affected by snow cover because the climatic conditions in mountain areas can significantly change year-by-year and zone-by-zone. However, the second analysis significantly improved the number of targets with respect to the first one: if we consider, for example, the Envisat ascending data, we registered an increase in the number of targets of 193% (Table 2).

As stated above, snow seasons in VDA mainly range from December to March; however, there are exceptions. For example, the 2008–2009 winter registered extraordinary snow season from November to May [84]. According to this experience, in T3 we improved the SAR image selection method throughout the analysis of meteorological records coeval to SAR observation. Considering several weather stations of the regional meteorological network, a day-by-day analysis in correspondence to the satellite acquisition epochs is carried out. All the considered weather stations [85] are located above 2000 m a.s.l. to guarantee the significance of the meteo-climatic measurements for high mountain environments. This analysis is performed in two phases: (i) snow coverage identification and (ii) interferogram evaluation. In particular, we examined the daily rainfall (in mm) and the snow depth (in cm) parameters. The images characterized by abundant rain or snow precipitation were discarded from further analyses. Successively, an interferogram analysis was performed to identify SAR acquisition strongly affected by snow cover. During this analysis, arbitrary thresholds were considered. We observed that over 50 cm in snow depth could generate noisy interferograms. Regarding rainfall, only a single image was discarded from the analysis (28 May 2007), associated with a rainfall event of more than 30 mm/day, which has also generated, in this case, noisy interferograms. According to the analysis of the meteo-climatic parameters, we discarded only the images associated with the most critical values of daily rainfall and snow depth to ensure the SBAS processed the largest number of acquisitions. Using this fine-tuning, in T3 we still improved the number of targets: considering the Envisat ascending orbit again as a reference, the increment of coherent points with respect to the first analysis is of 313% (Table 2); with respect to T2, an increase in the number of targets to 5668 was revealed.
It is worth noting that, for the reliability of the SBAS-DInSAR analysis, it is very important to keep in the processed dataset a number of SAR acquisitions as large as possible to preserve a good temporal sampling and to limit as much as possible the presence of temporal gaps [86]. Accordingly, we only removed the SAR acquisitions affected by a widespread snow cover that clear impacts on the spatial coherence of the interferograms. To this aim, it was crucial to crosscheck the meteo-climatic conditions that provide local information and with the quality of the interferograms that supply large-scale information.

Finally, in T4, we performed a final tuning of the processing parameters [53], such as the ground pixel dimension (from 80 m to 40 m) and the coherence value of the analyzed pixels (from 0.7 to 0.8), while the default values of the maximum perpendicular (400 m) and temporal baseline (1500 days) were maintained. The increment of the number of targets (Envisat ascending orbit) for this last processing corresponds to about 620%.

Similar or even larger increments were observed for the Envisat descending analysis (Table 2). The significant increment of coherent targets due to the fine-tuning of the processing parameters in T4 is worth noting. This can be mainly explained by the coherence value used in this task. This parameter allows users to define the selection of the pixels to be considered in the SBAS analysis by changing the spatial network of the pixels to be unwrapped [87]. The default value in the GPOD-SBAS service is 0.7, and it is suitable for a very large number of test sites. However, in some particular sites, the presence of a very large number of target with poor coherence can hinder the phase unwrapping (PhU) operation [87] within the GPOD-SBAS service processing chain. In these cases, the PhU results can be significantly improved if only the targets with very high coherence are selected. This generally happens in mountain areas, where the pixels are characterized by lower coherence values and, moreover, have a sparse distribution over the scene, thus making PhU step a challenging operation.

4. Results

By exploiting the GPOD-SBAS, we obtained mean deformation velocity maps and displacement time series for the VDA regional territory, covering a time interval of more than 4 years (ERS-1/2) and of more than 6 years (ASAR-Envisat), respectively (see also Figure S1). Table 2 summarizes the testing activity done to perform the iterative procedure implemented for the high mountain region, operating on the Envisat dataset, and later directly applied to the ERS-1/2 dataset. All of the processed tasks were successfully completed by the unsupervised GPOD-SBAS service in less than 24 h. Depending on the SAR images considered, and on the processing parameters applied, we obtained a different amount of coherent targets (see Table 2). In Figure 5, we show the distribution of coherent targets subdivided according to slope class, obtained by the slope calculation from a DTM with 10 m square grid spacing for each processed task of the Envisat datasets. We note that, for all the datasets, we have a general increase in the number of coherent targets for all slope classes. In the first task T1, the prevalent distribution corresponds to the lowest degree slope (0°–20°), coincident to the most urbanized areas, corresponding to the most coherent sectors. On the contrary, on the high to very high classes (40°–60° and >60°), we note a general lack of target distribution. However, the number of targets obtained in T4 is as large as one order of magnitude with respect to T1.

In Figures 6 and 7, we present the best results obtained for both Envisat (ascending and descending orbits) and ERS (ascending orbit) datasets (see also Figure S1). Coherent targets are mainly located at the main valley bottom, where urbanized areas (e.g., streets, railway, main towns, hamlets, and isolated building) are located. Within the observed period, LOS velocities in these areas are rather small, ranging from about ±0.2 cm/year and comparable between Envisat and ERS processes, confirming an overall ground stability. Relative to the ERS-1/2 dataset, directly processed on the basis of the results achieved with the Envisat analysis, a significantly lower data coverage respect to the Envisat resulted, presenting a total amount of targets of 8574. As expected from the SAR visibility analysis performed (see Figure 2), coherent target distribution is mainly on east-facing slopes, in ascending orbit, and on west-facing slopes, in descending orbit. Moreover, in the valleys, targets are mainly localized in hamlets and villages, as well as in areas of little or no vegetation (e.g., talus, moraine, and bare rock).
From the comparison between the GPOD-SBAS process results and the known distribution of the VDA gravitational process considered in this work, we notice that, in correspondence to glaciated sectors, a decorrelation effect due to their large displacement, and a consequently total absence of targets, was revealed.

**Figure 5.** Graph of the number of targets of each dataset available for the Envisat satellite for each task processed. The targets are subdivided and classified with respect to four slope classes, based on the slope calculation of the VDA region.

**Figure 6.** Surface velocities measured along the satellite LOS for the Envisat dataset: (a) in ascending orbit for the time period from June 2004 to October 2010 and (b) in descending orbit for the time period from September 2004 to October 2010. The light blue polygons correspond to the present glaciers.
Figure 7. Surface velocities measured along the satellite LOS for the ERS-1/2 dataset in ascending orbit for the time period from June 1995 to May 2000. The light blue polygons correspond to the present glaciers.

4.1. Deep-Seated Gravitational Slope Deformations

We compared the ground velocity map obtained through the GPOD-SBAS processing to the DsGSD phenomena present in the area of investigation, according to the IFFI (Italian Landslides Inventory) catalogue [88], and found a good correlation between deforming areas observed and the occurrence of DsGSD phenomena. We performed this comparison making use of Envisat results (both ascending and descending orbits), which present the best coverage of the VDA study area and a higher number of targets with respect to the ERS-1/2 results. Considering the ascending orbit, the inventoried DsGSDs included within the track considered are 179, and only 44% present targets inside, of which 25% present more than ten targets. Considering the descending orbit, 191 DsGSDs inventoried are included within the track considered; in this case, 35% of these phenomena present targets inside, of which only 19% present more than ten targets (Table 3). Moreover, the highest coherent target density for the east face-oriented DsGSD cases in ascending orbit, and for the west face-oriented DsGSD cases in descending orbit was observed.

Table 3. Summary of the number of deep-seated gravitational slope deformations (DsGSDs) inventoried in the IFFI project covered by targets generated by the GPOD-SBAS process.

<table>
<thead>
<tr>
<th>DsGSD (from IFFI)</th>
<th>Number of DsGSD (within the Track)</th>
<th>Number of DsGSD Covered by Targets</th>
<th>Number of DsGSD with &gt;10 Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENVISAT Ascending</td>
<td>280</td>
<td>179</td>
<td>79</td>
</tr>
<tr>
<td>ENVISAT Descending</td>
<td>191</td>
<td>66</td>
<td>36</td>
</tr>
</tbody>
</table>

Two specific examples are reported and illustrated in Figures 8 and 9. The first example is the Beauregard DsGSD, located in the southwest portion of the region, while the second is the Quart DsGSD, located in the middle portion of the principal valley.

The Beauregard DsGSD is an extended phenomena, going from 1700 m a.s.l., close to the valley bottom, up to 2900 m a.s.l. in the northern sector, corresponding to its main scarp. In [89], a non-homogeneous deformation pattern of this phenomenon is described, with the upper part
presenting the highest displacement rates, and a slow and constant deformation along the basal portion of the phenomenon.

Exploiting our results (Figure 8a), we noted a good density of the target distribution, higher than that in other DsGSDs. The target distribution mainly covers the upper-northern and the central sector, while no targets were revealed in the southern part and in general at the DsGSD toe, where only a small basal portion of the landslide close to the dam structure is covered. The distribution of the targets is influenced by the distribution of bedrock outcrops: in the area covered by grassland or woods, the number of targets is very poor.

Figure 8. (a) Mean deformation velocity map of the Beauregard DsGSD case study, for Envisat ascending dataset. (b,c) Time series corresponding to the targets signed by the white dots in the map.
In general, we notice a good agreement between our results and the hypothesized kinematic domains presence. In fact, it is possible to identify two distinct portions presenting different LOS velocity values for the observed period: (i) the upper sector, corresponding to the main scarp and constitute by two sub-domains belonging to the DsGSDs (reported in the IFFI project), present velocities ranging from $-1.04$ up to $-0.5$ cm/year, while (ii) the basal sector present the lowest LOS velocity, with a range variable from $-0.4$ to $-0.2$ cm/year. Moreover, observing Figure 9a, additional kinematic domains can be identified, i.e., a central sector, presenting ground LOS velocities from $-0.8$ to $-0.5$ cm/year, and a sector in the west portion of the gravitational phenomenon, characterized by LOS velocities from $-0.57$ to $-0.25$ cm/year.

Analyzing the ground deformation time series available for several targets within the phenomenon, we observe a substantially constant deformation rate, represented by a linear trend; however, several accelerations, related to the later spring–summer period, are observed. Specifically, five targets distributed within different Beauregard domains, are presented (Figure 8b,c). The higher displacement of $-3.7$ cm (away from the satellite), for the observed period from June 2004 and October 2010, was registered by the target B1 located in correspondence to the main scarp of the DsGSD, while the target B5 located in correspondence to the toe of the DsGSD, close to the dam, registered a maximum displacement of $-1.07$ cm. Instead, a displacement comparable to what was observed for the B1 target, is registered by target B4, with a maximum of $-3.5$ cm.
The DsGSD of Quart is located on the north side of the broad middle portion of the principal valley. This phenomenon covers an area of about 13 km$^2$ from 530 m a.s.l. at the valley bottom up to 2200 m a.s.l., in the terminal part of the main DsGSD scarp. In [10], the complex evolution of this phenomenon is described, and the ongoing deformation was attributed to an interaction between gravitational movements with glacial and fluvial processes operating over a long time period.

The mean deformation velocity map obtained with GPOD-SBAS processing confirms the ground deformation trend outlined in the literature [10] (Figure 9a) and allows for the identification of ground LOS velocities in the upper and central portions of about $-0.8$ and $-0.6$ cm/year, respectively, while lower sectors appear relatively stable over time. Furthermore, in this case, the ground deformation time series generally present a constant deformation rate and show local acceleration during the summer months; the maximum displacement of $-5.25$ cm for the observed period was registered by the upper sector of the DsGSD (Figure 9b); going down to the valley bottom, the central sector presents a maximum displacement of $-1.9$ cm, up to $-0.2$ cm in the lower sector of the slope.

4.2. Periglacial Ground Deformations

We compared the ground velocity map obtained by the GPOD-SBAS and the VDA rock glaciers (RGs) inventory [76]. In general, despite the widespread lack of targets coverage associated with the rock glaciers, we found a fairly good correlation. Considering the ascending orbit, 759 inventoried RGs were observed within the track considered (Table 4); in this case, just the 16% present targets inside, and only 4% present more than ten targets. For the descending orbit, a total of 811 phenomena inventoried are included within the task considered, of which 8% present targets, and only 2% present more than ten targets.

<table>
<thead>
<tr>
<th>Rock Glaciers (from Glaciers Inventory)</th>
<th>Rock Glaciers (within the Track)</th>
<th>Rock Glaciers Covered by Targets</th>
<th>Rock Glaciers with &gt;10 Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENVISAT Ascending</td>
<td>937</td>
<td>759</td>
<td>123</td>
</tr>
<tr>
<td>ENVISAT Descending</td>
<td>811</td>
<td>65</td>
<td>13</td>
</tr>
</tbody>
</table>

The mean deformation velocity maps obtained for both ascending and descending orbits of the Envisat processing results reveal a generally low coverage of the rock glaciers by the targets. Figure 10 shows the mean deformation velocity map detail for the descending orbit acquisition, reporting an example of an intact rock glacier and a relict one. The Petit Emilius (PE in Figure 10a) RG corresponds to the intact one and presents only one target within, with a LOS velocity of $-0.8$ cm/year, while the Becca di Nona (BN in Figure 10a) RG corresponds to the relict one and presents more than ten targets within, with a mean LOS velocity of $-0.5$ cm/year, for the entire observed period.

For these two cases, the analysis of the ground deformation time series of the targets (Figure 10b) reveals a constant deformation rate, with a maximum displacement of $-1.4$ cm for the observed period from September 2004 to October 2010. Analyzing the time series in detail, they seem to display a sort of acceleration in correspondence to the summer and the mid-autumn (e.g., summer 2009). It is worth noting that such phenomena present a seasonal fluctuation [68], where the highest velocities appear during the summer to early autumn months; looking at the results from year to year, comparing the late summer with respect to the early subsequent summer, a moderate acceleration can be observed.

Finally, considering the rock glaciers included within each track (Table 4), we observed in general that the intact rock glaciers present the largest deformation, while, for the rock glaciers identified as relict, a modest displacement, a smaller deformation with respect to the intact ones, was disclosed. In Figure 11, we present the Clapey Gerbioz relict rock glacier example: in ascending orbit (Figure 11a), this RG presents mean LOS velocities from $-1.1$ cm/year up to $-0.26$ cm/year; in descending orbit (Figure 11b), a range from $-0.7$ cm/year up to $-0.3$ cm/year.
Figure 10. (a) Mean deformation velocity map of the Becca de Nona (BN) and Petit Emilius (PE) rock glaciers in descending orbit; (b) time series corresponding to the targets signed by the white dots in the map.

Figure 11. Cont.
5. Discussion

We have shown that the GPOD-SBAS service may allow for the revelation of information about the ground displacements occurring in difficult environments such as high mountain regions.

The GPOD-SBAS service, which is a completely free web-tool, operating on a grid infrastructure, supplies the resources to process an enormous amount of data through the distribution of computing environments via the Parallel-SBAS (P-SBAS) technique [52]. The web-tool operates in an unsupervised manner, processing the P-SBAS chain step-by-step: (i) raw data focusing; (ii) SAR image co-registration; (iii) interferogram generation; (iv) the unwrapping of the computed phases; (v) the generation of the final results. However, in most cases all over the world, the unsupervised manner is not sufficient. For this reason, in order to maximize the number of coherent targets and their distribution over those areas, a number of best practices have to be performed. First, geomorphological and meteo-climatic settings of the study area have to be carefully considered. High topography gradients and complex orography, the predominance of abundant vegetation (e.g., forested areas), and unsuitable valley flank orientation are recognized as challenging environments for DInSAR techniques. The computation of the R-Index [43], considering the satellite’s acquisition geometries to analyze the potential distribution of the targets, is important preliminary information. In our example, the prevalent distribution has been observed not only on the south-facing slope in the ERS and Envisat ascending orbits and on the north-facing slope for the descending orbits, for the east–west oriented valleys, but also on the east-facing slope in the ascending orbits and on the west-facing slope in the descending orbits, for the north–south oriented valleys. Moreover, our results have shown the importance of a day-by-day analysis of the meteo-climatic parameters in correspondence to the satellite acquisition epoch, removing the SAR acquisitions affected by snow, causing temporal decorrelation effects. It is worth noting that the crosscheck of the local information obtained by the weather stations parameters available, with the large-scale information provided by the interferograms elaborated by the GPOD-SBAS service, has been crucial. In addition, by fine-tuning the advanced parameters of the GPOD-SBAS service, we have additionally improved the number of coherent targets and their distribution over the study area (Table 2). In particular, we have proved that of the parameter “coherent threshold,” which influences the spatial network of the pixels to be unwrapped, mountainous environments have to be increased with respect to the default values, and only very high-quality targets guarantee a significant improvement in the PhU operation. Our study allowed for the obtainment of ground deformation...
maps and the associated time series of ERS-1/2 and ASAR-Envisat data, over the entire VDA territory, for the time period respectively from June 1995 to May 2000 and from June 2004 to October 2010.

The comparison of the results obtained for the VDA region to the regional inventory including glaciers, rock glaciers, and DsGSDs shows good general agreement describing the deformation processes of already known phenomena. On glaciers, no displacement has been retrieved due to decorrelation effects caused by their fast evolution. On rock glaciers, a general lack of targets has been observed, limiting the deformation degree establishment (e.g., active, inactive, and relict). In general, an increment in LOS deformation during summer and mid-autumn could be revealed for the intact rock glaciers, while for the relict rock glaciers a slight LOS displacement has been detected, despite the absence of ice within. Finally, on DsGSDs already known and inventoried, a good coverage and distribution of targets have been observed on a regional scale. A downslope velocity deformation comparable with the DsGSDs of very slow moving phenomena [90] has been detected. Moreover, the results obtained showed different activity domains within DsGSD phenomena, in agreement with studies [10,89] of such phenomena in this region, presenting LOS deformation with a linear trend characterized by several acceleration related to the later spring–summer period.

Since in situ measurements suitable for a regional data comparison were not available, we compared our results to the data available on the Portale Cartografico Nazionale [91]. In this web-portal, the ERS-1/2 and ASAR-Envisat acquisitions are processed by the PSInSAR technique, over the entire Italian national territory. For the comparison, we selected a common stable area (Figure 12) and a number of targets presenting significant ground deformation (Figures 13 and 14).

![Figure 12. Cont.](image-url)
Figure 12. Comparison between the Envisat ascending (A) GPOD-SBAS results and (B) the PSInSAR results (from the Portale Cartografico Nazionale) over the Aosta municipality, considered as common stable area; (C) time series corresponding to the SBAS target and the Permanent Scatterers (PS), signified respectively by the black square in (A) and by the red square in (B).

Figure 13. Cont.
Figure 13. Comparison between the Envisat ascending (A) GPOD-SBAS results and the (B) PSInSAR results (from the Portale Cartografico Nazionale) over the Quart DsGSD; (C) time series corresponding to the SBAS target and the PS, signified respectively by the black square in (A) and by the red square in (B) within the Quart DsGSD phenomenon.

Figure 14. Cont.
Figure 14. Comparison between the Envisat ascending (A) GPOD-SBAS results and the (B) PSInSAR results (from the Portale Cartografico Nazionale) over the Valtournenche Valley corresponding to the Mont Meabè rock glacier; (C) time series corresponding to the SBAS target and the PS, signified respectively by the black square in (A) and by the red square in (B), within the Mont Meabè RG phenomenon.

The comparison shows good agreement between the measurements performed with different DInSAR approaches. Furthermore, the number of SBAS coherent targets and their distribution are comparable with the PS of the Portale Cartografico Nazionale.

6. Conclusions

The possibility to investigate surface deformation over wide areas with sub-centimeter accuracy can be very useful for the interpretation of the kinematic behavior of mass wasting phenomena typical of high mountain region (e.g., landslides, rock glaciers, and glaciers). However, the high complex relief, together with the land use and the snow cover, typical of mountainous areas, constitute important limitations for the DInSAR processing applications [42–44,86], also associated to the well-known intrinsic limitations of these techniques [33,45,46].

The proposed iterative procedure, applied through the P-SBAS web-tool within the GPOD environment, allowed us to improve SAR data selection and processing, over high mountain regions. The G-POD user-friendly interface permits to everyone—not just experienced users—quick accessibility to the SAR data and an analysis of SAR data using validated process, without tinkering with the different SBAS-DInSAR processing steps of specific and costly software.

Our processing strategy shows that it is possible to obtain long-term deformation analysis of diverse phenomena affecting mountain areas, and demonstrates how the DInSAR application could obtain reliable information about the ground displacement in the area without in situ instruments. This will be even more evident with the new Sentinel-1 satellite constellation that provides SAR acquisition characterized by significantly reduced revisit time (6 days), which represents an important improvement for the application and exploitation of the DInSAR results.

Supplementary Materials: The following are available online at www.mdpi.com/2072-4292/8/10/852/s1; Figure S1: Maps of surface velocities measured along the satellite LOS, for the Envisat dataset in ascending and descending orbit. The four images represent the results of the performed tasks for each orbit using the GPOD-SBAS service.

Acknowledgments: The research has been partially supported by HAMMER project, “Relationships between meteo-climatic parameters and ground surface deformation time series in mountain environment,” in the framework of the NextData project. The ERS-1/2 and ASAR Envisat data have been provided by ESA. The authors acknowledge Roberto Cuccu of the GPOD support team for providing assistance and help on processing the SAR dataset. The authors are grateful to the four anonymous referees for their comments, which substantially improved the manuscript.

Author Contributions: Martina Cignetti processed the dataset by the GPOD-SBAS service, performed the analysis, and wrote the paper. Andrea Manconi and Michele Manunta contributed to process the dataset and supervised the analysis. Daniele Giordan supervised the interpretation of the results from the geomorphological point of view. Claudio De Luca contributed to process the dataset. Francesca Ardizzone has coordinated the work. All authors contributed to paper writing and revision.
Conflicts of Interest: The authors declare no conflict of interest.

References


47. Cascini, L.; Fornaro, G.; Peduto, D. Advanced low- and full-resolution DInSAR map generation for slow-moving landslide analysis at different scales. *Eng. Geol.* **2010**, *112*, 29–42. [CrossRef]
64. Millar, C.I.; Westfall, R.D. Rock glaciers and related periglacial landforms in the Sierra Nevada, CA, USA; inventory, distribution and climatic relationships. *Quat. Int.* **2008**, *188*, 90–104. [CrossRef]


79. Tibaldi, A.; Rovida, A.; Corazza, C. A giant deep-seated slope deformation in the Italian Alps studied by paleoseismological and morphometric techniques. Geomorphology 2004, 58, 27–47. [CrossRef]


85. Centro Funzionale Regione Autonoma Valle d’Aosta Weather Station VDA. Available online: http://cf.regione.vda.it/il_centro_funzionale.php (accessed on 1 September 2015).


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
A 3D Shape Descriptor Based on Contour Clusters for Damaged Roof Detection Using Airborne LiDAR Point Clouds

Meizhang He1,2, Qing Zhu1,2,3,4, Zhiqiang Du1,2,*, Han Hu2,3,4, Yulin Ding2,3,4 and Min Chen2,3,4

1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China; vip_hmz@163.com (M.H.); zhuq66@263.net (Q.Z.)
2 Collaborative Innovation Center for Geospatial Technology, 129 Luoyu Road, Wuhan 430079, China; huhan19880715@163.com (H.H.); yulin.ding@whu.edu.cn (Y.D.); minchen@home.swjtu.edu.cn (M.C.)
3 State-Province Joint Engineering Laboratory of Spatial Information Technology for High-Speed Railway Safety, Southwest Jiaotong University, Chengdu 611756, China
4 Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 611756, China
* Correspondence: duzhiqiang@whu.edu.cn; Tel.: +86-15827108512

Abstract: The rapid and accurate assessment of building damage states using only post-event remote sensing data is critical when performing loss estimation in earthquake emergency response. Damaged roof detection is one of the most efficient methods of assessing building damage. In particular, airborne LiDAR is often used to detect roofs damaged by earthquakes, especially for certain damage types, due to its ability to rapidly acquire accurate 3D information on individual roofs. Earthquake-induced roof damages are categorized into surface damages and structural damages based on the geometry features of the debris and the roof structure. However, recent studies have mainly focused on surface damage; little research has been conducted on structural damage. This paper presents an original 3D shape descriptor of individual roofs for detecting roofs with surface damage and roofs exhibiting structural damage by identifying spatial patterns of compact and regular contours for intact roofs, as well as jagged and irregular contours for damaged roofs. The 3D shape descriptor is extracted from building contours derived from airborne LiDAR point clouds. First, contour clusters are extracted from contours that are generated from a dense DSM of individual buildings derived from point clouds. Second, the shape chaos indexes of contour clusters are computed as the information entropy through a contour shape similarity measurement between two contours in a contour cluster. Finally, the 3D shape descriptor is calculated as the weighted sum of the shape chaos index of each contour cluster corresponding to an individual roof. Damaged roofs are detected solely using the 3D shape descriptor with the maximum entropy threshold. Experiments using post-event airborne LiDAR point clouds of the 2010 Haiti earthquake suggest that the proposed damaged roof detection technique using the proposed 3D shape descriptor can detect both roofs exhibiting surface damage and roofs exhibiting structural damage with a high accuracy.

Keywords: contour cluster; 3D shape descriptor; shape similarity; entropy; building damage detection

1. Introduction

Building collapse is one of the primary causes of heavy human casualties in destructive earthquakes [1]. Rapid and reliable damage assessment on the individual building level following
earthquakes has become imperative for the optimal utilization of available resources for rescue [2,3]. Damage to roofs is an important feature for distinguishing extreme damage states, i.e., collapsed buildings, from lesser damaged or undamaged buildings [4]. Therefore, vertical remote sensing, including optical, SAR and LiDAR, represents an efficient tool for rapid damage assessment due to its low cost, high availability, minimal corresponding fieldworks, large coverage, digital processing and quantitative results. The effectiveness of remote sensing has also been proven following earthquakes worldwide [5–8]. Numerous methods have been reported for building damage detection using 2D features, such as gray scale, spectra, texture, edge and morphological features, and amplitude and phase information derived from optical or SAR imagery [8–17]. However, certain damage types (e.g., pancake collapse) cannot be identified using vertical remote sensing data due to the absence of precise height data [18]. Airborne laser scanning systems are particularly suitable for damaged roof detection because precise 3D point clouds can be rapidly obtained at all times and under most weather conditions without entering the quake-stricken area [19,20], and the elevation accuracy is higher compared to point clouds derived from vertical optical or SAR imagery [18,21]. Change detection using both pre- and post-event remote sensing data is a popular method of acquiring building damage information because detailed pre-event data are invaluable in reconnaissance [18,22–24]. However, the major limitation concerning this method is the lack of homogeneous pre-event reference data in many situations [18]. The method of damage interpretation using post-event remote sensing data can be applied even in the absence of homogeneous reference data, which is an alternative to rapid damage assessment during earthquake emergency response when pre-event data are limited.

According to the essential clue that damaged buildings, unlike organized manmade patterns on intact buildings, usually manifest themselves as disturbed spatial or spectral patterns [25], various methods have been developed to infer damage patterns from post-event data. The 2D features, including edge, texture and spectra, have been assessed by numerous studies [17,26–28] as important cues for damage detection because damaged regions tend to exhibit disturbed spatial or textural patterns, in contrast to intact buildings [29]. On the other hand, 3D features have been found to be useful for identifying specific damage types based on geometric reasoning, as highlighted by some studies [30–33]. Therefore, this paper focuses on severely damaged buildings and proposes a damaged roof detection approach based on the roof’s 3D features extracted from only post-event airborne LiDAR point clouds. In the following subsection, we do not provide an exhaustive review of all these methods; instead, we highlight only the 3D-feature-based approaches using only post-event airborne point clouds that are directly relevant to our work in the next subsection.

1.1. Damage Types

Prior knowledge about damaged buildings is necessary when performing building damage detection using airborne LiDAR point clouds [34]. A building damage catalog, as shown in Figure 1, including typical damage types of buildings, is used to identify buildings exhibiting different damage types based on geometry features such as reductions in volume and height, changes in the inclination of building surfaces and surface structures as well as the size of debris [35]. However, extracting geometry features such as reductions in volume and height and changes in the inclination of building surfaces using only post-event airborne LiDAR point clouds is difficult. Therefore, a majority of the building damage detection approaches using only post-event airborne LiDAR point clouds utilize geometry features including the surface structure and debris.
This paper categorizes the building damage catalog into two categories based on geometry features of roofs including debris and the roof structure. The first category, named surface damages, contains damaged building with debris surfaces such as multilayer collapse (2), top story pancake collapse (4c, 5, 5c), heap of debris (6, 7a, 7c) and heap of debris with planes (3, 7b). The second category, named structural damages, contains damaged buildings with relatively intact surfaces and includes inclined plane (1), middle or lower story pancake collapse (4a, 4b, 5a, 5b) and inclination (9a). Pancake collapse is a damage type of concern. The top story pancake collapse types, including damage types 4c, 5 and 5c, can be detected using post-event airborne LiDAR point clouds because the roofs are collapsed or damaged, whereas middle or lower story pancake collapses, including damage types 4a, 4b, 5a and 5b, are difficult to detect because the roofs are nearly intact. However, few perfect middle or lower story pancake collapses wherein the building entirely maintains its surface and structure during destructive earthquakes occur; the majority of buildings exhibiting this type of pancake collapse usually exhibit an inclination as well [36]. Therefore, this form of pancake collapse can be detected through the structure analysis of inclination.

1.2. Building Damage Detection Approaches

The majority of approaches to damage detection using only post-event airborne LiDAR data can be categorized into surface damage detection approaches and structural damage detection approaches, according to the damage categories defined in the above section.

The majority of methods for surface damage detection detect the surface damages of collapsed buildings based on the planarity of the roof surface because airborne LiDAR point clouds are particularly suited to extracting planar roof surfaces [37–39]. Rehor et al. [40] produced a 2.5D planar Delaunay-based triangulated irregular network (TIN) using planar segments and non-segmented points segmented from roof points and extracted debris triangles from the TIN as the damaged building parts. Rehor et al. [41] compared the random sample consensus (RANSAC) and region-growing algorithms applied to Digital Surface Models (DSMs) for building damage detection and suggested that the region-growing algorithm is more suitable for building damage detection than the RANSAC algorithm. Labiak et al. [42] presented a line-based slope threshold method for evaluating and identifying the damaged points of each roof based on the idea that points in intact roof planes have constant slopes, whereas points in damaged roof surfaces have varying slopes. Segment-based classification methods have been presented to detect collapsed buildings using only post-event airborne LiDAR point clouds based on the assumption that damaged buildings are represented as many small planar segments or unsegmented points, whereas intact buildings are represented as large planar segments [43–45].
Structural damage detection approaches detect damaged buildings via structure analysis based on prior knowledge about intact buildings. Shen et al. [31] extracted the geometric axis line of a flat or symmetric roof and identified the inclined roofs based on the assumption that the inclined roof’s angle between its geometric axis line and plumb line is greater than an empirical threshold angle. Gerke and Kerle [32] presented a graph-based approach for structural seismic damage assessment based on oblique airborne images. The structural integrity of a building is inferred based on the spatial relation between observable features such as vegetation, façade, intact roofs and destroyed roofs. The relations are represented through a directed graph and are trained by a graph-based learning algorithm. However, this approach is difficult to apply directly to airborne LiDAR point clouds because it is difficult to extract building façades from airborne LiDAR point clouds. Vetrivel et al. [29] developed a gap-based classification method for building structural damage assessment using post-event image-based 3D points. In this method, 3D points of building elements are voxelized based on a pre-defined voxel size; then, gaps are identified as the voxels that are visible from a sufficient number of cameras but that are not occupied by 3D points. Finally, using radiometric features, the gaps caused by damage are detected based on surrounding damage patterns such as spalling or debris. However, it is difficult to extract such evidence from airborne LiDAR point clouds because of their limited radiometric information. As a result, damage-related gaps cannot be reliably classified using airborne LiDAR point clouds. Fernandez Galarreta et al. [4] presented an UAV-based method for urban structural damage assessment using object-based image analysis and semantic reasoning. In this method, the detailed 3D point clouds were generated from multi-view imagery obtained by unmanned aerial vehicles; then, the z component of the normal of a local tangent plane of each point was computed from the co-variance matrix of the neighborhood points and was used to visually assess the D4-D5 damage elements in terms of the European Macroseismic Scale 1998 (EMS-98). Finally, the D1-D3 damage features were extracted via a more detailed façade and roof analysis using object-based image analysis and semantic reasoning. However, the D4-D5 damages were identified not by automatic analysis but by visual analysis.

The above-mentioned methods using airborne LiDAR point clouds all mainly focus on surface damage and pay minimal attention to structural damage. Thus, the spatial relations between the components of a building must be analyzed to infer structural damage [32]. The topological relationships between adjacent planar segments can be described and reconstructed using a roof topology in building modeling [46–48]. However, it is challenging to mathematically describe and explain the topological relationships of complex, damaged roofs based on a scattered point cloud at a low level, including at the point level and segment level, because the topology of random and irregularly damaged roofs is uncertain.

The 3D shape of a complex, damaged roof can be quickly represented by contours, which avoids the problem of topological analysis and reconstruction based on planar segments [49]. By analyzing the characteristics of a building’s contours, some of their features, such as closed and regular shapes, simple topology and density, are leveraged to extract and reconstruct buildings [49–52]. Our previous studies have shown that it is possible to detect roof damage using a contour-based method [53]. Damaged roofs with confusing contours were detected by a shape similarity analysis algorithm applied to building contours derived from airborne LiDAR point clouds. However, the damage feature based on contours was not explicitly described, and the automation of the algorithm was poor.

Focusing on surface and structural damages on roofs, this paper defines a 3D shape descriptor based on shape analysis of the contour clusters of buildings for detecting severely damaged buildings from post-event airborne LiDAR point clouds. The contribution of the paper lies in the presentation of a 3D shape descriptor that provides a comprehensive description of both surface and structure features of roofs based on the shapes and spatial relations of building contours. Compared to other 3D features of roofs, the 3D shape descriptor can more reliably and completely detect damaged roofs.

The remainder of this paper is organized as follows: Section 2 details the key procedures of damaged roof detection, including the data preprocessing, the definition and calculation algorithm of
the proposed 3D shape descriptor, and damaged roof detection based on the 3D shape descriptor using a maximum entropy threshold. Section 3 introduces the study area and the data source. Section 4 presents experimental results and discussion about damaged roof detection on the airborne LiDAR dataset of the Haiti earthquake in 2010. The selection of values for algorithmic parameters is also discussed in Section 4. The conclusions are presented in Section 5.

2. Methodology

This study proposes a method of damaged roof detection for both roofs suffering surface damage and roofs suffering structural damage based on a 3D shape descriptor using post-earthquake airborne LiDAR point clouds. The 3D shape descriptor is used to quantitatively describe both surface and structure features of roofs based on the shapes and spatial relations of a building’s contours. The descriptor is a more comprehensive description of 3D shapes of damaged roofs compared to other 3D features. Therefore, damaged roof detection based on the 3D shape descriptor can be more reliably and effectively performed.

There are three key procedures that constitute this method: data preprocessing, feature extraction and damaged roof detection. In data preprocessing, the DSM of each individual building is extracted from post-earthquake airborne LiDAR point clouds with guidance from 2D GIS vector data of building footprints and a digital elevation model (DEM). In feature extraction, the 3D shape descriptor of each building is calculated based on contour clusters generated from the DSM of each individual building. In damaged roof detection, damaged roofs are detected based on the 3D shape descriptor. Figure 2 illustrates the detailed procedures.

Figure 2. Flow chart of damaged roof detection.
2.1. Data Preprocessing

With the guidance of the 2D GIS vector data on building footprints, a dense DSM of an individual building is constructed using the airborne LiDAR point cloud and the DEM. The individual building points are extracted from the point cloud according to the building footprint, as observed in Figure 3b,f. A thin plate spline (TPS) interpolation method is used to interpolate the dense DSM with a grid cell size of $\lambda$ to remove the impact of noise and to create a continuous and smooth surface [54,55], as shown in Figure 3c,g. The ground points are set as having the same height, which is acquired from the DEM.

![Figure 3. The data processing samples: (a) The image of an intact building with an arched roof; (b) The airborne LiDAR point cloud; (c) The dense DSM; (d) Contour clusters colored with different colors; (e) The image of a completely collapsed building with intact walls; (f) The airborne LiDAR point cloud; (g) The dense DSM; (h) Contour clusters colored with different colors.](image)

2.2. Feature Extraction

The shapes and spatial relations of building contours can represent the roof’s 3D shape feature integrated with both surface and structure features more effectively than can planar segments. A 3D shape descriptor is defined to quantitatively describe the 3D shape feature using the shapes and spatial relations of building contours. Therefore, the 3D shape descriptor can potentially be used to identify surface or structural damages on roofs.

2.2.1. Feature Definition

The 3D shape of a roof’s surface and structure can be represented by a contour cluster, which is a set of contours extracted from a contour tree based on a containment relationship. A contour tree is a data structure addressing the description of surface topologies and is built upon the containment relationship between contours [56]. A contour cluster is defined as a subtree in the contour tree according to the following three conditions: (1) the subtree’s root node has no parent node or has a brother node in the contour tree; (2) the subtree’s terminal node has no child node or has at least two child nodes in the contour tree; and (3) the subtree’s other node has only one child node in the contour tree. The building’s surface regions belonging to the same structure are segmented from the complex building surface by the contour cluster based on a homogeneous spatial topology relationship, and each region is represented by a set of contours in the contour cluster, as shown in Figure 4.
Figure 4. Three building surface regions belonging to different structures are segmented by three contour clusters and are different colors: (a) Contours of the building; (b) Contour clusters of different structures that are represented as nodes colored with the same color in a contour tree.

The 3D shape difference between damaged and intact roof surface regions can be reliably distinguished using the full 3D geometric relationships among the contour clusters. One of the essential differences between damaged and intact buildings is that the former usually manifest as disturbed spatial patterns such as irregular surfaces and structures, unlike the organized, manmade patterns of the latter [25], as demonstrated in Figure 5. For an intact roof, the organized manmade pattern of the regular surface and structure is represented by a contour cluster containing similar shape contours. For a damaged roof, the disturbed spatial pattern of the irregular surface and structure is represented by a contour cluster containing chaotic shape contours. Therefore, the shape chaos of a contour cluster can be used to describe the irregular 3D shape of a damaged roof’s surface and structure and thus represents a potential feature for reliably distinguishing between intact roofs and damaged roofs due to surface or structural damages.

Figure 5. The spatial patterns of complex roof surfaces are represented by contour clusters: (a) The complex surface of Haiti’s National Palace, which was destroyed in the 2010 Haiti earthquake (obtained from Google Earth); (b) Contour clusters of the roof are rendered in different colors according to different surface or structural morphologies, such as contours of surfaces with debris (red), twisted surfaces (purple) and inclined roof structures (blue), which are jagged or irregular, and contours of intact surfaces and structures (green) are compact and regular.
According to the concept of information entropy developed by Shannon [57], the shape chaos index is defined to quantitatively describe the shape chaos based on the contour shape similarity measurement. The principle of information entropy is to use uncertainty as a measure to describe the information contained in a source [58]. For example, entropy is used to quantitatively measure the spatial information of a map such as metric information, topological information and thematic information [59]. In this paper, entropy is used to quantitatively measure the shape chaos of the entire contour cluster based on the shape similarities between two contours.

This can be achieved using the ratio between the number of contours with similar shapes and the total number of contours in the contour cluster as the probability $P_i$ used in the definition of the entropy $E$, as shown in Equation (1). Let $N$ be the total amount of contours, clustered into $n$ groups with similar shapes, and let the number of contours in each group be $N_i$. Thus, the probability $P_i$ can then be defined as shown in Equation (2). Specifically, if each contour is sufficiently similar, the shape chaos index will equal zero. If each contour is very different, which leads to each group having only one contour, the contour cluster will be clustered into $N$ groups, and all probabilities will be $1/N$; therefore, the shape chaos index will reach a maximal value: $\ln(N)$. Thus, the shape chaos index $C$ is the entropy $E$ normalized by the maximal value, as shown in Equation (3).

$$E = -\sum_i^nP_i\ln(P_i)$$  \hspace{1cm} (1)

$$P_i = \frac{N_i}{N}$$  \hspace{1cm} (2)

$$C = \frac{-\sum_i^nP_i\ln(P_i)}{\ln(N)}$$  \hspace{1cm} (3)

The 3D shape descriptor is the weighted sum of the shape chaos indexes of contour clusters corresponding to a single building, which describes the chaotic 3D shape of the entire building surface. The 3D shape descriptor is defined as shown in Equation (4).

$$S = \sum_1^m(Q_i \times C_i)$$  \hspace{1cm} (4)

where $m$ is the number of contour clusters corresponding to a single building, $Q_i$ is the weight of each contour cluster, and $C_i$ is the shape chaos index of each contour cluster. In this paper, $Q_i$ can be determined using the ratio between the area of the building surface region segmented by the contour cluster and the total area of the building. Let $A$ be the total area, and let $A_i$ be the area of each region. Such a weight can then be defined as shown in Equation (5). Because the shape chaos index and the weight are both normalized, the 3D shape descriptor is also normalized.

$$Q_i = \frac{A_i}{A}$$  \hspace{1cm} (5)

The damaged roof can be explicitly described by the 3D shape descriptor based on contour clusters. The 3D shape descriptor measures the chaotic degree of the entire roof surface’s 3D shape based on the shape chaos index. According to the Shannon entropy [57], the greater the difference between the contour shapes within the contour cluster, the larger the shape chaos index of the contour cluster, which means that the 3D shape of the surface region segmented using the contour cluster is more irregular. If the area of the irregular surface region is greater, the 3D shape descriptor will be larger. This means that the probability that the roof was damaged is higher. Accordingly, we assume that damaged roofs can be distinguished from intact roofs by a threshold $\delta$ of the 3D shape descriptor. Furthermore, the damaged roof is defined using the 3D shape descriptor, as shown in Equation (6).

$$D = \{S \mid S > \delta\}$$  \hspace{1cm} (6)
2.2.2. Feature Extraction Algorithm

This paper proposes a feature extraction algorithm for a 3D shape descriptor based on the shape analysis of contour clusters. There are three key procedures in the feature extraction algorithm: the contour cluster extraction, the shape chaos index calculation and the 3D shape descriptor calculation. Contour clusters are extracted from the contour tree, which is built using relationships based on the containment method for contours generated from the DSM. The shape chaos index of the contour cluster is calculated based on the information entropy of shape similarities, which are computed based on the shape similarity measurements among contours. The 3D shape descriptor is calculated based on the shape chaos indexes. Figure 6 shows the complete workflow of the feature extraction algorithm.

![Figure 6: The complete workflow of the feature extraction algorithm.](image)

(1) Contour Cluster Extraction

Contour clusters are extracted from the contour tree based on the three conditions discussed in Section 2.2.1. Dense contours with a contour interval of $\varepsilon$ are generated from a dense DSM of an individual-building-based grid-tracking method that interpolates points at given heights between grid cells and connects them sequentially [60,61], as shown Figure 3d,h. The contour tree is built using relationships based on the containment method [62,63]. The containment method suffers from the limitation that all contours are considered closed, and the unclosed contours are thus initially excluded. In a contour tree, each node represents a different contour, and every node may have a list of descendants. If one contour is contained by another, then that contour is a descendant. The containment method begins from the root node with the lowest elevation and recursively creates the contour tree in a depth-first manner.

(2) Shape Chaos Index Calculation

The shape chaos index of the contour cluster is used to quantitate the chaotic spatial pattern as the information entropy of shape similarities among contours. The shape similarity measurement is a crucial step in calculating the shape chaos index. The shape similarity is measured using the Euclidean distance between the normalized Fourier descriptors (FDs) of two contours. FDs are typically used in image retrieval and pattern recognition because of their insensitivity to geometric translation, rotation, and scaling [64–67]. We employ the normalized Fourier descriptor (nFD) to measure the approximate shape similarity between two contours based on the assumption that the shape differences among the
intact contour clusters are only produced by geometric translations and scaling, whereas the damaged contour clusters do not follow this assumption. Thus, the shape similarities based on nFDs among the intact contour clusters will remain small, and the shape similarities among the damaged contour clusters will be random. This results in the shape chaos indexes of contour clusters of intact roof surface regions having small values, whereas such values will be large in damaged roof surface regions.

Suppose a contour is composed of \( M \) points as a discrete complex function; then, the function can be transformed into the frequency domain without any loss of information by the Discrete Fourier Transformation (DFT), defined as \( a(k) \) in Equation (7). The coefficients \( a(k) \) are FDs that represent the discrete contour of a contour line in the Fourier domain [64].

\[
a(k) = \frac{1}{M} \sum_{j=0}^{M-1} [x(j) + iy(j)]e^{-\frac{ijk2\pi}{M}}, k = 0, 1, ..., M - 1
\]  

(7)

The Fourier coefficients \( a(k) \) are normalized to make them invariant to the translation, rotation, and scaling of contours [67]. The normalized FDs are defined as in Equation (8). Translation invariance can be achieved by omitting the Fourier coefficient \( a(0) \) and using the other Fourier coefficients because translation affects only the first Fourier coefficient \( a(0) \), and the other FDs retain their values [64]. The FDs are made invariant against rotation by taking the magnitude of each Fourier coefficient, and they are made scale invariant by dividing all Fourier coefficients by the magnitude of \( a(1) \) [64,67]. The normalized FD \( na(k) \), defined in Equation (8), also omits the second Fourier coefficient because it is a constant value of 1.

\[
na(k) = \left| \frac{a(k + 2)}{a(1)} \right|, k = 0, 1, ..., M - 3
\]  

(8)

The Euclidean distance is based on the normalized FDs of two contours following Equation (9).

\[
s = \sqrt{\sum_{i=0}^{L} (na_\alpha(i) - na_\beta(i))^2}
\]  

(9)

where \( na_\alpha \) and \( na_\beta \) are the normalized FDs of contour \( \alpha \) and contour \( \beta \), respectively, and \( L \) is the number of normalized Fourier coefficients. In most cases, these two contours contain a different number of points, and their normalized Fourier coefficients are not same; therefore, the Euclidean distance cannot be directly calculated. To solve this problem, both contours are first sampled such that they have the same number of points before FDs are applied to the two contours [67].

The shape chaos index is calculated based on an entropy definition whereby the probability is defined as the ratio between the number of contours with similar shapes and the total number of contours in the contour cluster, as detailed in Section 2.2.1. The first step is to cluster contours into several groups following the clustering rule that the shape similarity between two contours in the group must be less than a threshold \( \omega \). Based on the results of contour clustering, the probabilities of the shape distributions are computed, and the shape chaos index is calculated following Equation (3).

(3) 3D Shape Descriptor Calculation

The 3D shape descriptor calculation is dependent on the shape chaos indexes of contour clusters and their area weights. The area weight is defined as the ratio between the area of the roof surface region segmented by the contour cluster and the total area of the roof. Thus, it is crucial to estimate the area of the roof surface region for the weight calculation. The contour cluster is used to estimate the area of the corresponding roof surface region. Based on the containment relationship between contour clusters, there are two types of contour clusters. The first type of contour cluster does not contain any other contour cluster and has the same area as its outermost contour. The other type of contour cluster fully contains some contour clusters. The area of the cluster is the area of the outermost contour.
reduced by the areas of the outermost contours in contour clusters contained by the contour cluster. The total area of the roof’s surface is the sum of the areas of all contour clusters. Finally, the weight is computed following Equation (5).

2.3. Damaged Roof Detection

The damaged roof detection problem is transformed into a roof classification problem by using a maximum entropy threshold for the 3D shape descriptor. The maximum entropy threshold is implemented to select the optimal threshold $\delta$ from a group of buildings that includes intact and damaged buildings. As an optimal criterion, the maximum entropy was initially used for image thresholding by Pun [68] and was later corrected and improved by Kapur [69]. The maximum entropy threshold considers an image histogram as a probability distribution, and later, one obtains an optimal threshold value by maximizing the upper bound of the a posteriori entropy [58,70]. This paper adopts Kapur’s method based on the Shannon entropy to select an optimal threshold $\delta$ that yields the maximum entropy between damaged and intact buildings.

Suppose that all roofs are classified as intact or damaged roofs using a threshold of the 3D shape descriptor, which is denoted as $t$. Thus, the numbers of intact and damaged roofs are determined by the threshold $t$ and are defined as functions of $t$, as shown in Equations (10) and (11).

$$N_I = \text{Intact}(t)$$

$$N_D = \text{Damage}(t)$$

where $N_I$ is the number of intact roofs and $N_D$ is the number of damaged roofs. Intact and damaged roof histograms are built based on the 3D shape descriptor with a bin width $\varphi$. Because the 3D shape descriptor is normalized, the number of bins is denoted as $B$, which is determined using Equation (12). The frequency of each bin of the intact roof histogram is $FI_i$, as given in Equation (13), and the frequency of each bin of the damaged roof histogram is $FD_i$, as given in Equation (14).

$$B = \frac{1}{\varphi}$$

$$\sum_{i=1}^{B} FI_i = N_I$$

$$\sum_{i=1}^{B} FD_i = N_D$$

The entropy of the intact roofs is $E_I$, as given in Equation (15), and the entropy of the damaged roofs is $E_D$, as given in Equation (16). The a posteriori entropy is defined as a function of the threshold $t$ according to Kapur’s method [69], as in Equation (17). The optimal threshold $\delta$ is obtained by adjusting the threshold $t$ iteratively to maximize the upper bound of the a posteriori entropy. When the threshold $\delta$ is determined, damaged roofs will be identified following Equation (6).

$$E_I = -\sum_{i=1}^{B} \frac{FI_i}{N_I} \ln \left( \frac{FI_i}{N_I} \right) = -\sum_{i=1}^{B} \frac{FI_i}{\text{Intact}(t)} \ln \left( \frac{FI_i}{\text{Intact}(t)} \right)$$

$$E_D = -\sum_{i=1}^{B} \frac{FD_i}{N_D} \ln \left( \frac{FD_i}{N_D} \right) = -\sum_{i=1}^{B} \frac{FD_i}{\text{Damage}(t)} \ln \left( \frac{FD_i}{\text{Damage}(t)} \right)$$

$$E(t) = -\sum_{i=1}^{B} \frac{FI_i}{\text{Intact}(t)} \ln \left( \frac{FI_i}{\text{Intact}(t)} \right) - \sum_{i=1}^{B} \frac{FD_i}{\text{Damage}(t)} \ln \left( \frac{FD_i}{\text{Damage}(t)} \right)$$
3. Study Area and Data

3.1. Study Area

The study area (Figure 7) is located in the area surrounding Haiti’s National Palace, Port-au-Prince, Haiti. The site is approximately 776,660 square meters of flat terrain with dense buildings. On Tuesday, 12 January 2010, Haiti was hit by an earthquake that struck nearly 15 miles from Haiti’s capital, Port-au-Prince. According to official estimates, 316,000 people were killed, 300,000 were injured, 1.3 million were displaced, 97,294 houses were destroyed, and 188,383 houses were damaged in the Port-au-Prince area and in much of southern Haiti [71]. Many buildings in the study area were destroyed by the earthquake.

![Figure 7](image1.png)

Figure 7. The location of the study area.

3.2. Data Source

The airborne LiDAR data were acquired between 21 January and 27 January 2010, and had an average point cloud density of approximately 3.4 points per square meter [72], as shown in Figure 8a. A 2D GIS vector data of building footprints of the area, which contains 1875 buildings, as shown in Figure 8b, were provided by a third party.

![Figure 8](image2.png)

Figure 8. Data sets of the study area; (a) The airborne LiDAR point cloud; (b) The 2D GIS vector data of building footprints.
4. Results and Discussion

4.1. Damaged Roof Detection

In this paper, the detection of damaged roofs is based on damage interpretation using post-event airborne LiDAR data. The process includes three key procedures: data preprocessing, feature extraction and damaged roof detection. Five algorithm parameters, as shown in Table 1, are estimated experimentally or automatically during the workflow, which will be discussed in the Section 4.3.

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Symbol</th>
<th>Value</th>
<th>Set Mode</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data preprocessing</td>
<td>λ</td>
<td>0.1 m</td>
<td>Experimentally</td>
<td>The grid cell size of DSM</td>
</tr>
<tr>
<td>Feature extraction</td>
<td>ε</td>
<td>0.08 m</td>
<td>Experimentally</td>
<td>The contour interval</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>0.02</td>
<td>Experimentally</td>
<td>The shape similarity threshold</td>
</tr>
<tr>
<td>Damaged roof detection</td>
<td>ϕ</td>
<td>0.05</td>
<td>Experimentally</td>
<td>The bin width of the histogram</td>
</tr>
<tr>
<td></td>
<td>δ</td>
<td>0.4269</td>
<td>Automatically</td>
<td>The threshold of the 3D shape descriptor</td>
</tr>
</tbody>
</table>

During data preprocessing, 1875 individual buildings are extracted from the raw experimental airborne LiDAR data guided by building footprints, and a dense DSM of each building is generated with a grid cell size of λ from the building points. During feature extraction, first, contour clusters are extracted from each individual building’s contours, which are generated based on the contour interval ε from the dense DSM. Two examples of the procedure are shown in Figure 3. Second, the shape chaos index of each contour cluster is calculated using the shape similarity measurement based on the normalized FDs. Contours of the contour cluster are clustered into some groups using the shape similarity threshold ω. Entropy is calculated based on the probability of each contour group and is normalized as the shape chaos index. Third, the 3D shape descriptor of each roof is calculated based on the shape chaos indexes and the area weights of the contour clusters. In damaged roof detection, damaged roofs are discriminated from intact roofs using the optimal threshold δ, which is automatically selected using the maximum entropy threshold based on the roof histogram with a bin width of ϕ, as shown in Figure 9. The results of damaged roof detection are given as building footprints in a polygon shape file, as shown in Figure 10.

![Figure 9](image_url)  
*Figure 9.* The distribution of all intact and damaged roofs based on the 3D shape descriptor and the shape similarity threshold δ. The damage information of roofs is from the reference data.
Figure 10. The validated results of damaged roof detection, where correctly detected intact roofs (identified intact roofs) are colored green, correctly detected damaged roofs (identified damaged roofs) are colored red, undetected damaged roofs (misidentified intact roofs) are colored yellow, and undetected intact roofs (misidentified damaged roofs) are colored blue.

4.2. Accuracy Evaluation

For validation, the results are compared to manually labeled reference data based on remote-sensing-based building damage assessment data on the 2010 Haiti earthquake [73], where buildings were classified using the EMS-98 criteria. In these criteria, safe buildings with intact roofs and damaged walls are given grades of 1–3, and heavily or completely collapsed buildings with damaged roofs are given grades of 4–5 [74]. To evaluate the accuracy of damaged roof detection in this paper, buildings with grades of 1–3 are re-categorized as intact roofs, and buildings with grades of 4–5 are re-categorized as damaged roofs.

The results of the validation were compared to the reference data, and the accuracy indices, including the overall accuracy (OA), Kappa accuracy (KA), completeness (Compl.) rate and correctness (Corr.) rate, are listed in the confusion matrix in Table 2. The results show that the damaged roof
detection technique performed well at classifying both intact and damaged roofs, as shown in Figure 10: the detection technique correctly identified 652 out of 767 (85.01%) damaged roofs and 985 out of 1108 (88.90%) intact roofs. The overall accuracy is 87.31%, and the Kappa accuracy is 73.79%.

Table 2. Confusion matrix assessing the accuracy of damaged roof detection based on the 3D shape descriptor.

<table>
<thead>
<tr>
<th>Roof Status</th>
<th>Reference Data</th>
<th>Intact</th>
<th>Damage</th>
<th>Row Total</th>
<th>Corr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intact</td>
<td>985</td>
<td>123</td>
<td>1108</td>
<td>88.90%</td>
<td></td>
</tr>
<tr>
<td>Damage</td>
<td>115</td>
<td>652</td>
<td>767</td>
<td>85.01%</td>
<td></td>
</tr>
<tr>
<td>Column Total</td>
<td>1100</td>
<td>775</td>
<td>1875</td>
<td>89.55%</td>
<td>84.13%</td>
</tr>
<tr>
<td>Compl.</td>
<td>89.55%</td>
<td>84.13%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OA</td>
<td>87.31%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KA</td>
<td>73.79%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Some examples of typical buildings, including intact, surface-damaged and structurally damaged roofs, are shown in Figure 11. A surface-damaged roof with debris (Figure 11a) and intact roofs including a flat (Figure 11b), hipped (Figure 11c), gabled (Figure 11d) and pyramid (Figure 11e) roof are correctly identified; this result can also be reliably obtained using other surface damage detection methods using features based on points or planar segments. A completely collapsed building with intact walls as shown in Figure 3e is also detected as a roof with surface damage. Structurally damaged roofs with large planar segments are correctly identified; identifying these structural damages using planar-segment-based features can be difficult. The inclined roof shown in Figure 11f perhaps is a combination of middle and lower story pancake collapse. The inclined roof can be easily confused with an intact single sloping roof; therefore, this structural damage is difficult to identify using segment-based detection methods but is easily identified by the method based on the 3D shape descriptor. Top story pancake collapse, as shown in Figure 11g, with a twisted roof is distinctly represented by contour clusters and is correctly identified. Some intact regular arched roofs are also correctly identified, as shown in Figure 11h, i.

Figure 9 presents the distributions of intact and damaged roofs using a histogram based on the 3D shape descriptor. The undetected damaged roofs mainly show distributions in the critical region below the threshold, whereas undetected intact roofs are widely distributed above the threshold. The main cause of damaged roofs being undetected is that the number and extent of contours affected by damages in some partially damaged roofs are relatively low, which results in the 3D shape descriptors having small values. For example, there is a gap caused by damage on the roof, as shown in Figure 12a. The gap containing points is represented as a contour cluster with various shapes, where the outermost contour has the maximum elevation and is easily detected by the shape chaos index based on contour clusters. However, the area of the gap is very small compared to the whole roof; therefore, the 3D shape descriptor is understated. There are three main reasons for some intact roofs being undetected. The first reason is that the 3D shape characteristics of many shanties exhibit similarities with those of damaged shanties. The second reason is that small components or sundries on intact roofs lead to various irregular contours, as shown in Figure 12b. The third reason is that some intact roofs have regular but dissimilar contours, such as L roofs, as shown in Figure 12c. The shape transformation among these contours is not a translation or scaling transformation; rather, it is an approximate distance transformation. As a result, the normalized FD with invariance under shifting and scaling produces large values of the corresponding shape similarities as well as the 3D shape descriptors. Therefore, the normalized FD is not suitable for measuring the shape similarities among these contours. Some intact roofs are misidentified as damaged because the 3D shape descriptors are overstated for the above reasons.
Figure 11. Contour clusters of typical buildings: (a) Completely collapsed building with debris; (b) Intact flat roof; (c) Intact hipped roof; (d) Intact gabled roof; (e) Intact complex building consisting of pyramid roof; (f) Inclined roof; (g) Completely collapsed building with twisted roof; (h) Intact ruled surface roof; (i) Intact arched roof.

Figure 12. Contour clusters of typical buildings that are difficult to identify using the proposed 3D shape descriptor: (a) Partially damaged roof with gaps; (b) Intact roof exhibiting complex features; (c) Intact L-shaped roof.

In addition, many roofs had already experienced substantial decay prior to the earthquake and were often unfinished [75], thus incorrectly appearing as earthquake-induced damage in post-disaster data. Nevertheless, in the context of earthquake emergency response, undetected damaged roofs are considerably more problematic than misidentified damaged roofs [42].
4.3. Parameter Selection and Sensitivity Analysis

In damaged roof detection based on the proposed 3D shape descriptor, four parameters are experimentally estimated, as shown in Table 1. Several key parameters, such as the grid cell size $\lambda$, the contour interval $\epsilon$ and the shape similarity threshold $\omega$ for contour clustering, are determined using the characteristics of the airborne LiDAR data. It is generally impossible to have each point of an unorganized point cloud associated to one grid height. Consequently, dense DSM is used to approximate the complex building surface represented by the unorganized point cloud. Considering the size of the roof structure and the computational efficiency, the grid cell size of each individual building’s DSM is set as 0.1 m in the experiments. The quality of contours derived from a DSM is associated with the grid cell size of the DSM and the average slope in the topography [76]. To achieve a comparable quality of contours derived from a DSM, the average horizontal distance between contour lines of the same vertical interval should be approximately equal to the grid cell size. Therefore, we suppose that the optimal average horizontal distance is the grid cell size $\lambda$, and the contour interval can be determined as Equation (18).

$$\epsilon = \tan(\theta) \times \lambda$$  \hspace{1cm} (18)

where $\epsilon$ is the contour interval, $\theta$ is the average roof slope, and $g$ is the grid cell size of the DSM. According to general building structure knowledge, the average slope is usually less than 45°; therefore, the optimal contour interval should be no more than the grid cell size of the DSM. To select the key parameters $\epsilon$ and $\omega$, a sensitivity analysis is conducted using the variations in each parameter in a reasonable range while the other parameter is fixed at the value that leads to an optimum accuracy. The overall accuracy (OA), completeness (Compl.) rate and correctness (Corr.) rate are used to evaluate $\epsilon$ and $\omega$.

The quantitative analysis results for the two parameters are shown in Figures 13 and 14 which compare several reasonable values of $\epsilon$ and $\omega$. As the parameter $\epsilon$ increases, the overall accuracy (OA) and the completeness (Compl.) rate increase at first and then decrease; the correctness (Corr.) rate follows the opposite trend. When the parameter $\epsilon$ is 0.08 m, which is less than the DSM’s grid cell size of 0.1 m, the OA reaches the local maximum. The result verifies that a relatively reliable result can be achieved when the contour interval is no more than the grid cell size of the DSM. The exact roof slope estimated based on the point cloud can be used to select an ideal contour interval for each building. The OA, Compl. and Corr. also increase at first and then decrease as the parameter $\omega$ increases. When the parameter $\omega$ is 0.02, the OA reaches a local maximum. It is difficult to quantitatively analyze the parameter $\omega$ because it is affected by many factors such as the roof structure, the quality of the contours and the characteristics of the shape similarity measurement based on the normalized FDs. The sample training of contour clusters represents a good option toward selecting a proper parameter $\omega$. Based on the experiments, the values of $\epsilon$ and $\omega$ are set as 0.08 m and 0.02, which produce the optimum accuracy in damaged roof detection, as shown in Table 2.

![Figure 13. Sensitivity test of the contour interval $\epsilon$.](image-url)
4.4. Comparison

The main objective of this paper is for reliable and complete damaged roof detection using post-earthquake airborne LiDAR point clouds. The relevant methods in the literature usually only focused on one or several damage types of roofs. Consequently, in this section, the damage types and the accuracy of the proposed method are compared to those of other methods of damaged roof detection using 3D geometric features extracted from post-earthquake data.

Buildings are usually characterized by various damage types in a serious earthquake zone. In this situation, if a method can detect more damage types of buildings, it will have wide applicability, especially for earthquake emergency response. According to the damage types mentioned in Section 1.1, the damage types of roofs that can be detected by building damage detection methods using 3D geometric features are listed in Table 3.

Table 3. The applicability comparison results and damage types that can be detected by the building damage detection methods using 3D geometric features.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Damage Types</th>
<th>Surface Damages</th>
<th>Structure Damages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Multilayer Collapse (2)</td>
<td>Top story Pancake Collapse (4c, 5, 5c)</td>
<td>Middle or Lower Story Pancake Collapse (4a, 4b, 5a, 5b)</td>
</tr>
<tr>
<td></td>
<td>Heap of Debris (6, 7a, 7c)</td>
<td>Heap of Debris with Planes (3, 7b)</td>
<td>Inclined Plane (1)</td>
</tr>
<tr>
<td>Rehor et al. [40]</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Rehor et al. [41]</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Labiak et al. [42]</td>
<td>√</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td>Oude Elberink et al. [44]</td>
<td>√</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td>Shen et al. [31]</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Gerke and Kerle [32]</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Vetrivel et al. [29]</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Fernandez Galarreta et al. [4]</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>The proposed method</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

Many methods [40–42,44] extract debris for damaged roof detection based on the planarity of the roof. These methods are suitable for detecting surface-damaged roofs but present difficulties in detecting structurally damaged roofs because there is minimal debris in a structurally damaged roof. Vetrivel et al. [29] only focused on the gaps caused by damage and could not detect these building damage types. Gerke and Kerle [32] detected surface-damaged and inclined buildings based on the spatial relation between observable features such as façades, roofs and rubble piles. However, their method cannot detect inclined planes (1) and middle and lower story pancake collapse (4a, 4b, 5a, 5b)
because façades are intact and because rubble piles do not exist for these damage types. Shen et al. [31] was able to detect structurally damaged roofs using the geometric axis lines of the roofs. However, it was difficult to detect inclined flat roofs because this building damage type is easily confused with an intact single sloping roof. Fernandez Galarreta et al. [4] was able to detect both surface- and structurally damaged roofs via visual analysis.

A fair comparison between the precision of these methods depends on the type and accuracy of the remote sensing datasets, the number of damage classes, and the type and accuracy of the reference data. The methods in [42,44] and the proposed method detected intact and damaged roofs from different study areas using airborne LiDAR point clouds of the Haiti earthquake, and the results were evaluated using a damage assessment performed by the Global Earth Observation—Catastrophe Assessment Network (GEO-CAN). Labiak et al. [42] obtained an overall accuracy of 73.40% and a Kappa accuracy of 27.51%. Oude Elberink et al. [44] achieved an overall accuracy of 60%. In contrast, the proposed method achieved a high overall accuracy and Kappa accuracy of 87.31% and 73.79%, respectively.

The comparison suggests that the proposed method based on the 3D shape descriptor automatically detects both surface- and structurally damaged roofs and achieves a higher accuracy in damaged roof detection using the 3D geometric features extracted from post-earthquake airborne LiDAR point clouds. This is because the 3D shape descriptor provides a comprehensive description of both surface and structure features based on the shapes and spatial relations of building contours, thus significantly improving the completeness and accuracy of damaged roof detection. A major drawback of the proposed method is that it is difficult to automatically identify specific damage types using the 3D shape descriptor. However, the first priority is to reliably and completely detect damaged roofs in earthquake emergency response.

5. Conclusions

This paper focuses on both surface-damaged and structurally damaged roof detection using post-event data and proposes a 3D shape descriptor based on building contour clusters derived from airborne LiDAR point clouds for practical application.

The novelty of the 3D shape descriptor is that the 3D shapes of complex roof surfaces are quantitatively described through the spatial patterns of contours. The significant 3D features of surface and structural damages are characterized as a chaotic spatial pattern of contours and are represented as a group of contours with chaotic shapes. The shape chaos index of contour clusters quantitates the chaotic spatial pattern as the information entropy of shape similarities among contours by integrating the spatial attributes and spatial relationships of complex roof surfaces at the regional level. The 3D shape descriptor quantitatively describes the characteristics of chaotic 3D shapes of surface and structural damages at the whole roof level by combining the shape chaos indexes using the area weight.

A significant performance improvement is achieved through the use of the novel 3D shape descriptor based on contour clusters compared to other geometric features directly extracted from point clouds. The experiments produce good results for the airborne LiDAR point cloud used in Haiti earthquake damaged roof detection, which classifies roofs solely using the 3D shape descriptor with the maximum entropy threshold. The overall accuracy and Kappa accuracy are 87.31% and 73.79%, respectively. This damaged roof detection approach is quite valuable for rescue efforts in earthquake-struck areas, especially when pre-event 3D data are difficult to obtain.

Future work will be devoted to identifying specific damage types of damaged roofs detected using the proposed 3D feature, as well as to improving the shape analysis between contours to increase the robustness of the 3D shape descriptor.

Acknowledgments: This work was supported by the National High Resolution Earth Observation System (the Civil Part) Technology Projects of China, the National Natural Science Foundation of China (No. 41571390, 41571392, 41471320) and the Open Research Fund of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (No. 15101). The Haiti LiDAR data were supplied through an RIT partnership with ImageCat Inc. and Kucera International, which was sponsored by the World Bank. ImageCat also provided the GEO-CAN damage assessment validation data. The imagery of the study area was obtained from Google Earth.
We would like to thank the anonymous reviewers for their constructive comments, which greatly improved the quality of our manuscript.

Author Contributions: Meizhang He designed the study, performed the experiments, analyzed the data, and wrote the manuscript. Qing Zhu and Zhiqiang Du supervised Meizhang He. Meizhang He, Qing Zhu, Zhiqiang Du, Han Hu, Yulin Ding and Min Chen reviewed the paper for organization, clarification and English corrections.

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References


10. Vu, T.T.; Matsuoka, M.; Yamazaki, F. Detection and animation of damage using very high-resolution satellite data following the 2003 Bam, Iran, earthquake. Earthq. Spectra 2005, 21, 319–327. [CrossRef]


24. Awrangjeb, M. Effective generation and update of a building map database through automatic building change detection from LiDAR point cloud data. Remote Sens. 2015, 7, 14119–14150. [CrossRef]
48. Li, Y.; Ma, H.; Wu, J. Planar segmentation and topological reconstruction for urban buildings with LiDAR point clouds. Proc. SPIE 2011, 8286. [CrossRef]
49. Song, J.; Wu, J.; Jiang, Y. Extraction and reconstruction of curved surface buildings by contour clustering using airborne LiDAR data. Optik 2015, 126, 513–521. [CrossRef]
56. Guilbert, E. Multi-level representation of terrain features on a contour map. Geoinformatica 2013, 17, 301–324. [CrossRef]


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
Identification of Structurally Damaged Areas in Airborne Oblique Images Using a Visual-Bag-of-Words Approach

Anand Vetrivel *, Markus Gerke, Norman Kerle and George Vosselman

Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede 7500 AE, The Netherlands; m.gerke@utwente.nl (M.G.); n.kerle@utwente.nl (N.K.); george.vosselman@utwente.nl (G.V.)

* Correspondence: a.vetrivel@utwente.nl; Tel.: +31-654-257-742

Academic Editors: Zhenhong Li, Roberto Tomas and Prasad S. Thenkabail

Received: 22 December 2015; Accepted: 4 March 2016; Published: 11 March 2016

Abstract: Automatic post-disaster mapping of building damage using remote sensing images is an important and time-critical element of disaster management. The characteristics of remote sensing images available immediately after the disaster are not certain, since they may vary in terms of capturing platform, sensor-view, image scale, and scene complexity. Therefore, a generalized method for damage detection that is impervious to the mentioned image characteristics is desirable. This study aims to develop a method to perform grid-level damage classification of remote sensing images by detecting the damage corresponding to debris, rubble piles, and heavy spalling within a defined grid, regardless of the aforementioned image characteristics. The Visual-Bag-of-Words (BoW) is one of the most widely used and proven frameworks for image classification in the field of computer vision. The framework adopts a kind of feature representation strategy that has been shown to be more efficient for image classification—regardless of the scale and clutter—than conventional global feature representations. In this study supervised models using various radiometric descriptors (histogram of gradient orientations (HoG) and Gabor wavelets) and classifiers (SVM, Random Forests, and Adaboost) were developed for damage classification based on both BoW and conventional global feature representations, and tested with four datasets. Those vary according to the aforementioned image characteristics. The BoW framework outperformed conventional global feature representation approaches in all scenarios (i.e., for all combinations of feature descriptors, classifiers, and datasets), and produced an average accuracy of approximately 90%. Particularly encouraging was an accuracy improvement by 14% (from 77% to 91%) produced by BoW over global representation for the most complex dataset, which was used to test the generalization capability.

Keywords: damage detection; feature representation; oblique airborne images; supervised learning; texture; UAV; Visual-Bag-of-Words

1. Introduction

Rapid damage assessment after a disaster event such as an earthquake is critical for efficient response and recovery actions. Direct manual field inspection is labor intensive, time consuming, and cannot assess the damages in inaccessible areas. Remote sensing technology is the most predominant and early source to provide data for performing such assessments, either manually or using automated image analysis procedures [1,2]. Various kinds of remote sensing data such as optical, synthetic aperture radar (SAR), and LiDAR are being used for the damage assessment process [1]. However, optical data are often preferred as they are relatively easy to interpret [1]. Moreover, optical remote sensing provides very high resolution images ranging from the decimeter to the centimeter scale through various platforms such as satellites, manned aircrafts, and unmanned aerial vehicles (UAVs) [3–5]. This allows for performing comprehensive damage assessment through identifying
different levels of damage evidence, ranging from complete collapse to cracks on the building roof or façades, by choosing images at appropriate scales. Particularly oblique airborne images are recognized as the most suitable source, as they facilitate the damage assessment on both roofs and lateral elements [6,7]. For example, even extensive building damage such as inter-story collapse or pancake collapse can be identified reliably only with oblique view images, while conventional nadir views at best provide damage proxies such as external debris [1,7,8]. Although current remote sensing yields images at a vast range of views and scales, automatic recognition of even heavy damages to buildings is still challenging [1]. This is due to various reasons such as the complexity of the scene, uncertain characteristics of damage patterns, and the varying scale problem in oblique view images.

Generally, the regions corresponding to heavy damage are determined through the identification of damage patterns corresponding to rubbles piles, debris, and spalling in an image region (refer to Figure 1) [8]. Those damage evidences have a specific meaning and play a major role in damage classification. For example, the presence of significant amounts of debris/rubble piles around the building is the strong indication of (partial) building collapse. Spalling is an indicator of minor damage or partially broken structural elements. The recognition process of those damage patterns can be performed by analyzing features extracted either at the pixel or the region level [1,9,10]. However, the pixel level analysis is not meaningful for very high spatial resolution images, particularly in the context of damage assessment, as the evidences are identified based on the characteristics of their radiometric distribution pattern, which can be captured more precisely at a region level. However, in region-level classification the critical step is to define the region that is appropriate to identify the specific damage patterns. Generally, image regions are obtained either through a gridding approach or though image segmentation [11]. The most simple, efficient, and straightforward strategy is the gridding approach, where the image is split into uniform rectangular cells. However, the regions derived from gridding are often cluttered, as they may comprise different kinds of objects. For example, a single cell may contain trees, building elements, cars, road sections, debris, etc. Moreover, oblique images are more complex compared to nadir images, since they also capture façades that frequently comprise various elements, such as windows, balconies, staircases, etc. They generally also look more cluttered than nadir images containing largely roofs, and reveal façade information only at the image border, depending on the lens opening angle. It is quite challenging to identify the damage patterns in such a cluttered region. This can be alleviated by using a segmentation approach, which segments the damaged portions and other objects in the scene as separate regions. However, the selection of appropriate features and a segmentation algorithm that is suitable for a given damaged and cluttered environment is a challenging problem, one that requires substantial semantic analysis. Apart from clutter, the regions obtained from oblique images vary in scale. Nevertheless, the identification of damage patterns regardless of image scale is an important prerequisite in damage assessment. For example, damage at a building level such as inter-story collapse can be captured better at coarser scales (e.g., 100 × 100 pixel region in an image of the decimeter scale), while minor damages such as spalling at a building element level require finer scales (e.g., 100 × 100 pixel region in an image of the centimeter scale). Therefore, a robust method is required to recognize the damage pattern in a defined region irrespective of the scale and clutter. This is an analogue of the human visual pattern recognition system, which is extremely proficient at identifying the damage patterns regardless of the scale and complexity of the scene.
In the field of computer vision, various methods have been reported for pattern recognition tasks in various applications, such as object categorization, face recognition, and natural scene classification [12–14]. These methods are mostly based on supervised learning approaches, which work well for conventional image classification applications. However, the overall performance of the learning approach completely depends on the discriminative power of the image descriptors (features) considered for the classification [15]. Generally, images are described through either global (e.g., textures) or local features, like point descriptors such as Scale Invariant Feature Transform (SIFT) [13,16]. However, most global features are very sensitive to scale and clutter [17]. In contrast, the local descriptors are robust to clutter but cannot capture the global characteristics of the image [18,19]. An alternate feature representation strategy, such as Visual-Bag-of-Words (BoW), captures the global characteristics of the image through encoding a set of local features, which makes them robust to scale and clutter [20,21]. For example, in texture-based classification, the global texture pattern of the image is captured by the frequencies of the co-occurrence of the local texture patterns. This kind of feature representation outperforms the conventional global feature representation approaches in image classification [22]. Apart from general image classification, the Bag-of-Words framework has been demonstrated as a potential approach in many image-based domain specific applications including image retrieval [23], human action and facial expression recognition [24,25], image quality assessment [26], and medical image annotation [27]. Conceptually, thus, the BoW approach seems to be appropriate for identifying the damaged regions in airborne oblique images, which generally look cluttered and vary in scale.

Pattern recognition methods including BoW are based on a supervised learning approach that attempts to learn the underlying relationship between the image-derived features and the pattern of a specific category, in this case the damage pattern. Therefore, apart from a feature representation strategy, the choice of features that best discriminate the damaged and non-damaged regions is also a key element. Numerous studies reported that textures are the most influential feature for damage pattern recognition, as the damaged regions tend to show uneven and peculiar texture patterns, in contrast to non-damaged regions [28–30]. Many damage classification studies used statistical textures such as grey level co-occurrence matrix (GLCM)-based features for the damage pattern recognition [10,31,32]. However, other texture measures such as wavelets have been recognized as superior to GLCM in many pattern recognition problems, including land cover classification [33]. Particularly for region-level pattern classification problems, descriptors such as Histogram of Gradient Orientation (HoG), Gabor wavelets, SIFT and Speeded-Up Robust Features (SURF) have led to good results [34–36]. All these features describe the pattern of the given region in a unique way, based on the magnitude of gradient along various orientations and scales. Vetrivel et al. [37] demonstrated the potential of HoG and Gabor features to classify the damaged regions in very high resolution UAV images. However, they found limitations with the conventional global representation of HoG and Gabor features, especially with respect to generalization. So far, however, to our knowledge no work exists that combines the named features in a BoW fashion for damage mapping.
The objective of this paper is thus to develop a robust method based on the BoW approach that is suitable especially (but not only) for oblique images to identify the damage patterns related to rubble piles, debris, and spalling, regardless of the scale and the clutter of the defined region in an image. Following the above argumentation, a grid-based region definition is pursued. The robustness of the developed method based on this BoW approach is analyzed by comparing the performance of various learning algorithms and image descriptors (Gabor and HoG) under both the conventional and the BoW approach. Also, the generalization capability of the developed method is analyzed, by testing it on a variety of images corresponding to various scales, camera views, capturing platforms, and levels of scene complexity.

2. Methods

For the identification of damaged regions in an image, as a preparation step we provide reference data. That is, the given image is split into \( M \times N \) regions, which are termed image patches. The image patches are manually labeled as damaged if any kind of damage pattern related to debris, spalling, and rubble piles is observed in them. The automatic detection of those damage patterns within the patches is carried out using two different feature representation approaches: global and BoW representation. The feature descriptors and learning algorithms considered for both the global and BoW-based damage classification process are described in the respective sub-sections.

2.1. Damage Classification Based on Global Representation of Features

This process includes two steps: (1) extraction of image descriptors that provide the global description of the given image patch; and (2) classification of the given image patch as damaged or non-damaged, based on the extracted feature descriptors using a supervised learning algorithm.

2.1.1. Extraction of Feature Descriptors

The HoG and Gabor wavelets-based feature descriptors are considered for the global feature representation-based damage classification process.

(a) Histogram of Gradient Orientation (HoG)

The standard approach is used to extract the HoG features (e.g., Dalal Triggs [38]), where the given image patch is split into a number of overlapping blocks, and histograms of gradient orientation derived for each block are concatenated to form a feature vector. This gives the global representation of the image patch.

Procedure:

1. Derive gradient magnitude and its orientation for each pixel in the image patch.
2. Split the gradient image into \( A \times B \) cells.
3. Again split the gradient image into a number of overlapping blocks, where each block contains \( C \times D \) cells with 50% of overlapping cells between the blocks.
4. Define the bin size for the histogram of gradient orientation, where each bin corresponds to a specific orientation (the bin size remains fixed for all experiments later).
5. For each cell, compute the histogram of gradient orientation by adding the magnitude of the gradient to its corresponding orientation bin. Therefore, the size of the feature description of each cell is equal to the number of bins.
6. Concatenate the histograms of gradient orientation of all cells within each block to get the block level description. Normalize the histograms’ magnitude within the block to compensate for the local illumination variations [39].
7. Concatenate all block level descriptors to form the global descriptor of the patch.
(b) Gabor wavelets descriptors

The Gabor wavelets descriptors are obtained by convolving the image with a set of Gabor wavelet filters. These filters are derived by appropriate rotation and scaling of the mother Gabor wavelet function, where each filter is tuned to capture the pixel information at a specific orientation and frequency. The detailed procedure for Gabor wavelets filter generation can be found in Arivazhagan et al. [40]. After obtaining the Gabor filter responses for each pixel in the image patch, the region-level Gabor wavelet descriptor is represented by the histogram of magnitude of filter responses for all combinations of orientations and frequencies (cf. Jun Fei [41]). This histogram is computed for three consecutive pyramid levels of image patches, in order to capture the variation across scales, in addition to the variation across frequencies and orientations. The procedure used for extracting the global Gabor feature descriptors for an image patch is described below.

Procedure:

1. Generate $I \times J$ number of 2D Gabor wavelet filters, where $I$ and $J$ are the number of frequencies and number of orientations used to generate the Gabor wavelet filters, respectively.
2. Convolve the image patch with the generated filter banks, which results in $I \times J$ number of feature images.
3. Normalize each feature image using $l^2$ normalization.
4. Compute the histogram of Gabor filter responses, where each histogram bin corresponds to a specific frequency and orientation. Therefore, the number of histogram bins is equal to $I \times J$, which is the size of the final feature vector.
5. Also, extract the Gabor wavelet features for the other two pyramid levels of the image patch, by subsampling it to $\frac{1}{2}$ and $\frac{1}{4}$ of the image patch size.
6. Feature vectors derived at different scales are concatenated to form the final feature vector. Therefore, this final feature vector will comprise features extracted at multiple scales, multiple frequencies, and multiple orientations.

2.1.2. Damage Classification Using the Derived Global Feature Descriptors

Supervised learning approaches are adopted to classify the given image patch as damaged or non-damaged, based on the global feature descriptors. Three state-of-the-art and widely used supervised learning algorithms, Support Vector Machines (SVM) [42], Random Forests (RF) [43], and Adaboost [44], are considered for the damage classification process. These learning algorithms belong to the families of different learning paradigms, which learn the underlying relationship between the input features and the output label in a unique way. Three different learning paradigms are considered in order to analyze whether the considered feature descriptors are independent of the supervised algorithm, i.e., how the classification task is solved independently of the applied learning strategy. Also, each learning algorithm has a number of tunable parameters, referred to as hyper-parameters, which have a significant impact on the performance of the learning model [45]. Therefore, the hyper-parameters are tuned for the best model by searching the parameter space using the grid space search approach [46]. This approach constructs a number of learning models for different settings of the hyper-parameters, using the training set. The performance of each model is assessed using a cross-validation procedure. The best performing model is selected as the final model with tuned hyper-parameters, and then evaluated using the testing set.

2.2. Damage Detection Using Visual-Bag-of-Words

The standard BoW framework is adopted for the damage classification process. The BoW framework comprises different components, such as feature point detection, feature descriptors, visual word dictionary, and a classifier. The algorithms used for each component and the overall procedure are described below.
Overall, the BoW-based damage classification process is carried out in two stages: (1) construction of visual word dictionary; and (2) representation of the image in terms of BoW (histogram of visual words), and training the classifier based on them.

Stage 1:

(a) **Feature point detection**

The basic idea behind this step is that an image can be described by a small number of significant pixels (salient points). For example, pixels corresponding to edges and corners contain the most significant information compared to pixels of homogenous regions. Salient point descriptors that are invariant to scale and orientation are most appropriate to build an image classification model that is robust to scale and rotation. Numerous such salient point detection methods are available, with SIFT and SURF commonly being used in the BoW context [19]. In this study, SURF was used since it is faster than SIFT and its descriptor is suitable to be used as the feature in the BoW framework, as discussed in the following sub-section. A description of the SURF points detection process can be found in Bay et al. [47].

(b) **Feature extraction**

The purpose of this step is to extract the local feature descriptor for each salient point in the given image patch. The feature descriptors HoG and Gabor wavelets, which are used in the global representation-based damage classification, are also considered here for the local description of salient points in the BoW-based damage classification. This allows us to compare the potential of BoW and global feature representation irrespective of the features. In the BoW approach the SURF descriptor is additionally used to describe the salient points. This is because SURF is a well-proven point descriptor (local descriptor), widely used in BoW-based image classification processes [48]. Furthermore, SURF descriptors are based on wavelet responses, which also describe the image region in terms of textures, similar to HOG and Gabor feature descriptors. Therefore, the three feature descriptors HoG, Gabor wavelets, and SURF are used independently to describe each salient point in the given image patch for the BoW-based damage classification. The local pattern description for each salient point is derived by considering a local neighborhood of $P \times Q$ pixels around the salient point. The same procedure as described in Section 2.1.1 is followed to extract the Gabor and HoG features. The standard procedure is used to extract the SURF feature descriptor (cf. Bay et al. [47]).

(c) **Visual words dictionary construction**

The feature descriptors of salient points from all image patches (regardless of their class) are concatenated into a single feature vector. Numerous feature encoding methods have been reported for visual word dictionary construction [49]. We adopted the most commonly used iterative $k$-means clustering algorithm [48]. The obtained feature vector is clustered into $k$ clusters using the iterative $k$-means clustering [50]. Each cluster center is considered as the visual word, and the cluster centers are collectively referred to as a visual word dictionary.

Stage 2:

(a) **Image description based on visual words**

To represent the given image patch in terms of BoW (histogram of visual words), the salient points in the image patch are detected and feature descriptors are obtained for each point. The detected points in the image are assigned to their closest visual word in the dictionary. Subsequently, the frequency of occurrence of the visual words in the image is represented as a single histogram, which is referred to as the BoW representation of the image, which will be fed into the classifier in the next step.
(b) Classification of visual words using machine learning algorithms

Again, the three learning algorithms SVM, RF, and Adaboost are used as classifiers for classifying the damage and non-damaged image patches based on BoW. The procedure described in Section 2.1.2 is followed to develop the supervised learning models based on the BoW features.

The overall workflow of BoW-based damage classification process is depicted in Figure 2.

3. Experiments and Results

The damage classification method was evaluated using four different datasets, with each differing in its image characteristics such as scale, camera view, capturing platform, and scene complexity. Each data set was independently analyzed for the damage classification process based on the three feature descriptors HoG, Gabor wavelets, and SURF. The performances of HoG and Gabor wavelets for damage classification were analyzed by representing them in both a conventional and BoW framework. Also the potential of the SURF descriptor was analyzed for damage classification by representing it in a BoW framework and comparing it with BoW-based Gabor and HoG.
Three supervised learning algorithms, SVM, RF, and Adaboost, were used for analyzing the performance of the feature descriptors. Therefore, each dataset was tested with different combinations of feature descriptors and supervised learning algorithms, as depicted in Figure 3.

The conducted experiments for the damage classification process include a number of algorithms, and each algorithm was associated with a number of parameters. The values used for the parameters of the algorithm are shown in Table 1. The hyper-parameters considered for tuning the learning algorithms (cf. Section 2.1.2) are described in Table 2.

![Figure 3. Combinations of feature descriptors and learning algorithms tested for each dataset.](image)

**Table 1.** Definition of parameters associated with each algorithm/method used in the experiment.

<table>
<thead>
<tr>
<th>Algorithm/Method</th>
<th>Parameter Values</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image patch generation</td>
<td>M = 100; N = 100</td>
<td>To generate $100 \times 100$ image patches</td>
<td>Section 2</td>
</tr>
<tr>
<td>HoG procedure</td>
<td>A = 25; B = 25</td>
<td>Cell size $A \times B$ -- $25 \times 25$ pixels</td>
<td>Section 2.1.1</td>
</tr>
<tr>
<td></td>
<td>C = 4; D = 4</td>
<td>Block size $C \times D$ -- $4 \times 4$ cells</td>
<td></td>
</tr>
<tr>
<td></td>
<td>bin size = 9</td>
<td>Bin size of histogram of the gradient orientations</td>
<td></td>
</tr>
<tr>
<td>Gabor wavelet descriptor</td>
<td>I = 5; J = 8</td>
<td>$I, J$ are the number of frequencies and orientations respectively to generate the Gabor wavelet filters</td>
<td>Section 2.1.1</td>
</tr>
<tr>
<td>Feature extraction</td>
<td>P = 10; Q = 10</td>
<td>$10 \times 10$ local neighborhood is considered for deriving descriptor for each salient point</td>
<td>Section 2.2</td>
</tr>
<tr>
<td>Visual dictionary construction</td>
<td>$k = 500$</td>
<td>$k$ value for $k$-means clustering</td>
<td>Section 2.2</td>
</tr>
<tr>
<td>Supervised model for damage classification</td>
<td>10-fold cross validation</td>
<td>Cross-validation to identify the optimal hyper-parameters for a learning model based on the grid search approach</td>
<td>Section 2.1.2</td>
</tr>
<tr>
<td></td>
<td>The dataset is split into 70% and 30% for training and testing, respectively</td>
<td>Training set is used to train the model and also for cross-validation for tuning the hyper-parameters. Testing set is used for evaluating the trained model.</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Definition of grid search space for tuning the hyper-parameters of the classifiers.

<table>
<thead>
<tr>
<th>Supervised Classifier</th>
<th>Hyper-Parameter</th>
<th>Grid Search Space</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>C</td>
<td>0.001 to 100, step size—multiples of 10</td>
<td>Regularization parameter which has a significant effect on the generalization performance of the classifier.</td>
</tr>
<tr>
<td></td>
<td>Kernel</td>
<td>Linear, radial basis function (RBF) and histogram intersection</td>
<td>The function used to compute the kernel matrix for classification.</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>0.0001 to 1.0, step size—multiples of 10</td>
<td>Regularization parameter used in RBF kernel (Gaussian kernel function) which has significant impact in the performance of the kernel.</td>
</tr>
<tr>
<td>RF</td>
<td>N_estimators</td>
<td>3 to 20, step size 2</td>
<td>Number of trees in the forest.</td>
</tr>
<tr>
<td></td>
<td>Max_depth</td>
<td>1 to 5, step size 1</td>
<td>Maximum depth of the tree.</td>
</tr>
<tr>
<td></td>
<td>Min_samples_split</td>
<td>1 to 4, step size 1</td>
<td>Minimum number of samples required to split a node.</td>
</tr>
<tr>
<td></td>
<td>Min_samples_leaf</td>
<td>1 to 3, step size 1</td>
<td>Minimum number of samples required in newly created leaf after the split.</td>
</tr>
<tr>
<td>Adaboost</td>
<td>N_estimators</td>
<td>100 to 1000, step size 100</td>
<td>The maximum number of estimators that can be used to build the ensemble learning model.</td>
</tr>
<tr>
<td></td>
<td>Learning rate</td>
<td>0.01 to 0.1, step size 0.01</td>
<td>Regularization parameter that scales the contribution of each weak estimator.</td>
</tr>
</tbody>
</table>

3.1. Dataset 1: UAV Images

UAV images captured over two different areas were considered: (1) a small region around a church (“Church of Saint Paul”) in Mirabello, Italy, damaged by a 2012 earthquake; (2) a small region around a partly demolished industrial facility in Gronau, Germany. Both regions possess similar characteristics, and they contain only a few buildings that are largely disconnected. One building in each region was partially collapsed and produced a significant quantity of debris and rubble piles (cf. top left image in Figure 4—UAV image-subset of the Mirabello church). The UAV images were captured at different heights, positions, and views (nadir and oblique) with a spatial resolution of 1–2 cm. The images of both regions corresponding to various orientations and heights were split into 100 × 100 pixel rectangular image patches for framing of the training and testing datasets for the damage classification process. The patches are labeled as damage if at least 25% of their area represents damage evidences (debris/rubble or spalling) that are unambiguously recognizable by human analyst. For example, the samples of damaged and undamaged image patches of each dataset are shown to provide better insights (refer to Sections 3.1–3). Since the image resolution is very high, the defined rectangular patches cover only a small region (approximately 1 m²) and, therefore, most of them contain only either damage evidences or single homogenous object, i.e., the defined regions are mostly uncluttered; refer to the image training samples in Figure 4. In total 966 samples (482 damaged, 484 non-damaged) each of size 100 × 100 pixels, were considered. The dataset was constructed by selecting the specific samples across different regions within the scene that highly vary in their characteristics to avoid a large number of repetitive samples. The damage classification was performed for this dataset based on different combinations of feature descriptors and learning algorithm as described above, and the results are reported in Table 3.
Figure 4. Samples of image patches in dataset 1—UAV images.

Table 3. Performance of feature descriptors when associated with different learning algorithms for dataset 1 comprising patches from the UAV images (training samples = 676, testing samples = 290)—bold numbers indicate best performance per indicator.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>SVM</th>
<th>RF</th>
<th>Adaboost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision</td>
<td>Recall</td>
<td>Accuracy</td>
</tr>
<tr>
<td>Gabor</td>
<td>0.91</td>
<td>0.87</td>
<td>0.90</td>
</tr>
<tr>
<td>HoG</td>
<td>0.87</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>BoW-Gabor</td>
<td>0.96</td>
<td>0.93</td>
<td>0.95</td>
</tr>
<tr>
<td>BoW-HOG</td>
<td>0.98</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>BoW-SURF</td>
<td>0.97</td>
<td>0.92</td>
<td>0.94</td>
</tr>
</tbody>
</table>

3.2. Dataset 2: Oblique View Manned Aircraft Images

The airborne oblique images (Pictometry) with a Ground Sampling Distance (GSD) between 10 cm (foreground) and 16 cm (background) captured over Port-au-Prince after the 2010 Haiti earthquake were considered. The images cover almost the entire city, and contain numerous buildings ranging from simple to complex. Most of the buildings are densely clustered in such a way that it is difficult to differentiate each building even visually from the images. Numerous buildings are partially covered with densely leafed tall trees, adding to the clutter of the scene. A significant number of buildings are damaged, ranging from complete/partial collapse to heavy spalling on the intact elements of the building (cf. Figure 5). The images are split into 100 × 100 pixel images/patches to frame the training and testing datasets for the damage classification process. The defined image patches are highly cluttered as they cover a large area (at least 10 m²) and comprise different kinds of objects, such as trees, building elements, cars, road sections, and debris (cf. Figure 5). The dataset was constructed by selecting the specific samples across different regions within the city that vary highly in terms of their...
characteristics. Again, the selection of samples was driven by the idea of covering different damage characteristics rather than piling up redundant information. In total 1256 samples (698 damaged, 558 non-damaged) were selected and tested for the damage classification based on the developed approach. The patches cover an area of approximately 13,000 m². The results are reported in Table 4.

Table 4. Performance of feature descriptors when associated with different learning algorithms for dataset 2, comprising patches from Pictometry images (training samples = 879, testing samples = 377).

<table>
<thead>
<tr>
<th>Dataset2</th>
<th>SVM</th>
<th></th>
<th></th>
<th>RF</th>
<th></th>
<th></th>
<th>Adaboost</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision</td>
<td>Recall</td>
<td>Accuracy</td>
<td>Precision</td>
<td>Recall</td>
<td>Accuracy</td>
<td>Precision</td>
<td>Recall</td>
</tr>
<tr>
<td>Gabor</td>
<td>0.81</td>
<td>0.76</td>
<td>0.79</td>
<td>0.82</td>
<td>0.76</td>
<td>0.79</td>
<td>0.78</td>
<td>0.61</td>
</tr>
<tr>
<td>HoG</td>
<td>0.78</td>
<td>0.61</td>
<td>0.72</td>
<td>0.67</td>
<td>0.61</td>
<td>0.66</td>
<td>0.62</td>
<td>0.58</td>
</tr>
<tr>
<td>BoW-Gabor</td>
<td>0.89</td>
<td>0.86</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
<td>0.80</td>
<td>0.79</td>
</tr>
<tr>
<td>BoW-HOG</td>
<td>0.93</td>
<td>0.89</td>
<td>0.91</td>
<td>0.85</td>
<td>0.83</td>
<td>0.84</td>
<td>0.80</td>
<td>0.69</td>
</tr>
<tr>
<td>BoW-SURF</td>
<td>0.91</td>
<td>0.89</td>
<td>0.90</td>
<td>0.84</td>
<td>0.82</td>
<td>0.83</td>
<td>0.80</td>
<td>0.78</td>
</tr>
</tbody>
</table>

3.3. Dataset 3: Street View Images

Street view close-range images of damaged buildings captured by hand-held cameras after earthquakes in different geographic locations were used. These images were collected from two sources: (1) Governmental organization: the German Federal Agency for Technical Relief, THW; (http://www.thw.de); and (2) the Internet (various websites). The collected images vary in scale; however, the actual scale is unknown. Therefore, the 100 × 100 pixel patches generated from those images may cover small areas (e.g., an element of the building) or large areas (e.g., entire or major portion of the building). The collected images contain buildings with various kinds of damages, such
as complete collapse, partial collapse, inter-story collapse, and heavy spalling. In total 887 samples (384 damaged, 503 non-damaged) were considered to construct and evaluate the supervised model for the damage classification. Sample image patches used for the training and testing of the supervised model are depicted in Figure 6. The results are reported in Table 5.

Table 5. Performance of feature descriptors when associated with different learning algorithms for dataset 3, comprising patches from street view images (training samples = 620, testing samples = 267).

<table>
<thead>
<tr>
<th>Dataset 3</th>
<th>SVM</th>
<th>RF</th>
<th>Adaboost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision</td>
<td>Recall</td>
<td>Accuracy</td>
</tr>
<tr>
<td>Gabor</td>
<td>0.89</td>
<td>0.85</td>
<td>0.87</td>
</tr>
<tr>
<td>HoG</td>
<td>0.95</td>
<td>0.74</td>
<td>0.86</td>
</tr>
<tr>
<td>BoW-Gabor</td>
<td>0.99</td>
<td>0.91</td>
<td>0.95</td>
</tr>
<tr>
<td>BoW-HOG</td>
<td>1.0</td>
<td>0.93</td>
<td>0.96</td>
</tr>
<tr>
<td>BoW-SURF</td>
<td>0.99</td>
<td>0.89</td>
<td>0.94</td>
</tr>
</tbody>
</table>

3.4. Dataset 4: Datasets 1, 2, and 3 Are Combined

The samples from datasets 1, 2, and 3, which vary in scale, camera view, and platform and scene complexity, were combined into a single dataset in order to assess the generalization capability of the damage classification methods. In total 3109 samples (1564 damaged, 1545 non-damaged; subsequently termed COM3109) were used to develop and test the supervised models for damage classification. The results are reported in Table 6. For visual analysis, an UAV image of dataset 1 and a Pictometry image of dataset 2 were classified for the damage detection using the best performing model (BoW-Gabor with SVM). The classified images are depicted in Figure 7. The classification is quite accurate, showing only very few false positives and false negatives, which are also highlighted in the classified images (cf. Figure 7). The false positives and negatives are the examples where our assumption fails: *i.e.*, a surface with unusual radiometric characteristics is assumed to be damaged, while manmade objects are assumed to have a regular shape and uniform radiometric characteristics. For example, in Figure 7b.
the leaf-off tree was misdetected as damage, since it shows a strong irregular texture pattern. Similarly, the damaged regions are not detected as they show smooth texture.

Figure 7. Damage classification of images based on best performing supervised model: (a) UAV image of dataset 1 (left); detected damaged regions are highlighted in red, and the false positives are highlighted using yellow circles (right); (b) Subset of Pictometry image of dataset 2 (left); detected damaged regions are highlighted in red, and the false positives and false negatives are highlighted using yellow and green circles, respectively (right). Images © Pictometry; (c) Street view image of dataset 3 (left); detected damaged regions are highlighted in red, and the false positives and false negatives are highlighted using yellow and green circles, respectively (right).
Table 6. Performance of feature descriptors when associated with different learning algorithms for dataset 4 (COM3109) comprising patches from UAV, Pictometry, and street-view images (training samples = 2176, testing samples = 933).

<table>
<thead>
<tr>
<th>Dataset 4</th>
<th>SVM</th>
<th>RF</th>
<th>Adaboost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision</td>
<td>Recall</td>
<td>Precision</td>
</tr>
<tr>
<td>Gabor</td>
<td>0.79</td>
<td>0.75</td>
<td>0.77</td>
</tr>
<tr>
<td>HoG</td>
<td>0.81</td>
<td>0.62</td>
<td>0.73</td>
</tr>
<tr>
<td>BoW-Gabor</td>
<td>0.95</td>
<td>0.88</td>
<td>0.91</td>
</tr>
<tr>
<td>BoW-HOG</td>
<td>0.89</td>
<td>0.87</td>
<td>0.88</td>
</tr>
<tr>
<td>BoW-SURF</td>
<td>0.90</td>
<td>0.84</td>
<td>0.87</td>
</tr>
</tbody>
</table>

4. Observations and Analysis

For convenient analysis of the results, datasets 1 and 3, which were not cluttered and less affected by shadows and trees, are referred to as non-complex datasets, while datasets 2 and 4, where the image patches were mostly cluttered and severely affected by shadows and trees, are referred to as complex datasets. Also for convenience, the datasets are named based on the image characteristics and number of samples as described in Table 7.

Table 7. Naming of datasets based on the image characteristics and number of samples.

<table>
<thead>
<tr>
<th>Dataset 1</th>
<th>Name</th>
<th>Description</th>
<th>Scene Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset 2</td>
<td>UAV966</td>
<td>966 image patches generated from UAV images.</td>
<td>Non-complex</td>
</tr>
<tr>
<td>Dataset 3</td>
<td>PIC1256</td>
<td>1256 image patches generated from Pictometry images.</td>
<td>Complex</td>
</tr>
<tr>
<td>Dataset 4</td>
<td>SVI887</td>
<td>887 image patches generated from street view images.</td>
<td>Non-complex</td>
</tr>
<tr>
<td>Dataset 4</td>
<td>COM3109</td>
<td>Comprehensive dataset, where datasets 1, 2 &amp; 3 are combined, containing 3109 image patches.</td>
<td>Complex</td>
</tr>
</tbody>
</table>

4.1. Global Representation of HoG and Gabor Wavelet for Damage Classification

The results show that the global representations of HoG and Gabor wavelet feature descriptors have great potential to identify the damaged regions in the image, if the defined image patches are non-complex. For example, the supervised models constructed for UAV966 (non-complex) based on the global representation of Gabor wavelet and HoG features resulted in accuracies of 95% and 93%, respectively (Table 3). However, the same feature descriptors Gabor and HoG produced accuracies of 82% and 72%, respectively for PIC1256 (Table 4), where the defined image patches were mostly complex. Moreover, the same features Gabor and HoG produced highly inferior results for COM3109, which was more complex than the other datasets (Table 6). This clearly indicates that the robustness of the global representation of HoG and Gabor features declines with an increase in image patch complexity. This is because in the global representation the radiometric characteristics of the complex region (e.g., clutter, shadows, and trees) resemble the radiometric characteristics of damaged regions, which are generally more non-uniform than radiometric patterns of non-damaged regions (cf. Figure 8).

Consider an image patch that contains different objects with different dominant orientations. The global description of this image patch based on gradient orientation is the aggregation of all gradient orientation information within this patch. In such a case the image patch would seem to possess gradient orientations in many directions, which resembles the radiometric characteristics of damaged regions. For example, consider Figure 8a as an image patch that contains four different objects (annotated as A, B, C, and D) with different gradient orientation patterns. The gradient pattern derived locally for each object was overlaid on the corresponding object with a black background. These local patterns show that each object possesses dominant orientations that were more uniform. However, the global gradient pattern derived for the whole image patch was non-uniform and resembles the characteristics of damaged regions (cf. Figure 8a). Thus, it is difficult to classify an image patch based on global features in case it...
is cluttered. Also, trees and shadows possess irregular shapes and non-uniform gradient orientations, which also resemble the radiometric characteristics of the damaged regions. Hence, global feature descriptors-based damage classification did not efficiently classify the image patches that were strongly affected by trees and shadow.

Figure 8. (a) Local and global gradient pattern of an image patch that contains four objects with different dominant orientations; (b) gradient pattern of damaged regions.

4.2. BoW-Based Feature Representation for Damage Classification

The Gabor and HoG features produced superior results for all datasets when represented in a BoW framework than represented in conventional global scale for damage classification. Although the BoW approach produced superior results to the conventional approach, there was no significant difference in the performance between them when the considered image regions were not complex, e.g., UAV966 (cf. Table 3). However, in complex image regions there was a significant performance difference between the BoW and conventional feature representation. This is evident from the results for PIC1256, where the BoW-based Gabor and HoG produced maximum accuracies of 88% and 91%, respectively, which are 9% and 19% higher than the accuracies obtained by those features when represented at a global scale (cf. Table 4). This shows that BoW-based Gabor and HoG features are more robust to clutter, trees and shadows than when they are represented at a global scale. The following characteristics made the BoW approach more robust compared to the global representation:

(1) Unlike the global representation, the BoW approach does not aggregate the radiometric patterns within the image patch. Instead, it describes the image patch based on the number of salient points, where each point is described by the local radiometric pattern derived from its neighborhood. Therefore, in case of no damage, the image patch will be represented by points with a uniform radiometric pattern (gradient orientation), even if the image patch contains objects with different dominant orientations. On the other hand, if the image patch contains damage it will be represented by the points with non-uniform gradient orientations. The final damage classification is performed by analyzing the pattern of the occurrences of local radiometric patterns within the image patch. This
eliminates the ambiguity caused by mixed radiometric pattern typical for the global representation, making the BoW comparatively more robust.

(2) The BoW approach considers only the salient points as representatives to describe the image patch. The salient point selection method based on SURF mostly did not consider the pixels of shadows and trees as salient points. Figure 9a,b show the strongest 300 SURF points in the image, where most of the detected points are not corresponding to trees and shadows. Thus, the BoW approach largely eliminates the shadows and trees in the damage classification process, which was one of the major problems in the global descriptors-based damage classification. Moreover, the pixels corresponding to the damaged regions were often detected as salient points, as they show a stronger gradient than other objects (cf. Figure 9a). This ensures that the number of points corresponding to the damaged portion will always be proportional to the number of points corresponding to the non-damaged objects, even if only a small portion of the image patch is damaged (cf. Figure 9a,b). This specific characteristic of the SURF points made the BoW-based damage classification approach invariant to scale, clutter, and scene complexity.

![Figure 9](image_url)

**Figure 9.** Detected SURF points are plotted on the image. (a) Strongest 300 SURF points among 4032 × 3024 pixels; (b) strongest 300 SURF points among 977 × 835 pixels. Images © Pictometry.

### 4.3. Impact of Choice of Learning Algorithm

The results show that the choice of learning algorithm has a significant impact on damage classification performance, since the feature descriptors performed differently for different datasets when associated with different learning algorithms. The accuracies produced by SVM, RF, and Adaboost for all datasets when they were associated with different feature descriptors are depicted in Figure 10. The plot shows that (1) SVM and RF mostly outperformed Adaboost; (2) using the global feature descriptors the performances of RF and SVM varied with the datasets: RF produced superior results compared to SVM for UAV966 and SVI887, whereas it produced inferior results than SVM for PIC1256 and COM3109. This shows that the performance of the learning algorithm is highly dependent on the characteristics of the datasets, with SVM performing well for complex datasets and RF performing well for non-complex dataset. However, using the BoW approach the SVM mostly outperformed RF in the classification for all datasets, irrespective of the feature descriptors (cf. Figure 10). One overall conclusion from this is that the SVM-based, supervised models were more reliable and mostly showed better generalization performance than RF, particularly for the complex datasets.
5. Discussion

The primary objective of this paper was to develop a damage classification method that classifies a given image patch as damaged or non-damaged, irrespective of scale, image characteristics, and scene complexity. The damage classification method was developed by considering various feature descriptors (HoG, Gabor, and SURF), different feature descriptor representations (Global and BoW), different learning algorithms (SVM, RF, and Adaboost) and image datasets with different levels of scale and scene complexity. It was shown that the feature representation has a significant impact on the performance of the damage classification compared to other components such as feature descriptors and learning algorithms. For all datasets, the BoW-based damage classification models performed well for all combinations of feature descriptors and learning algorithms, compared to the models developed based on global representation. Particularly, concerning COM3109 (the comprehensive dataset), the accuracy obtained with the best-performing feature descriptor (Gabor) and learning algorithm (SVM) with the global representation improved by 14% (from 77% to 91%), when tested with the BoW representation (cf. Table 6). The choice of learning algorithm was found to be the second significant factor in the performance of the damage classification model: the SVM produced significantly better results than RF and Adaboost for all feature descriptors in the BoW representation (cf. Table 6). The considered feature descriptors performed equally well and, hence, the choice of feature descriptor was found to have least impact on the performance of the model. The Gabor features led to a 3% and 4% improvement in accuracy compared to HoG and SURF, respectively, when the image patches were classified with SVM in the BoW framework. This small improvement also may be
due to the additional information that Gabor features possess compared to HoG. For example, in Gabor, the gradient orientations information is extracted based on five different frequency scales, whereas in HoG the gradient orientation information is extracted at only one frequency scale (cf. Section 2.1.1). However, these improvements are significantly more modest when compared to the 14% improvement in accuracy between BoW and global representation (cf. Table 6). This highlights the importance of feature representation, regardless of the potential of features. Overall, SVM associated with Gabor feature descriptors in the BoW framework was found to produce the most robust and generalized damage classification model. Even visually, the damage classification was found to be more accurate when the images of different scales, camera views, and capturing platforms, and different levels of scene complexity, were classified by the best performing model (cf. Figure 7). Shadowed areas continue to pose a major problem in damage classification. Since the damaged regions covered by shadows show low contrast, they were not detected by our BoW-based approach (no SURF points in those areas). However, it is important to identify the damages in low-contrast regions as well; therefore, further tuning of the methods or identifying the optimal strategy that can make our approach work even in low contrast regions is required to increase the robustness of the model.

The BoW framework consists of various components such as feature descriptors, learning algorithms, and the visual word dictionary construction. The algorithms used for each component are associated with a number of parameters (cf. Table 1). The performance of the BoW-based damage classification model might be further improved by tuning the parameters of the algorithm or modifying/replacing the algorithm of the specific component. For example, the iterative $k$-means clustering was used to construct the visual word dictionary, whereas other feature encoding methods such as auto-encoders (e.g., Vincent et al. [51]), which encode the features differently compared to $k$-means, may produce a better visual word dictionary and thereby can potentially improve the performance of the model as well. Concerning the feature descriptors, all three-feature descriptors were used independently to construct the damage classification models, whereas the combined use of feature descriptors may also improve the performance of our model. Similarly, concerning the learning algorithm we used a single kernel-based SVM for constructing the damage classification model, whereas the multiple-kernel (e.g., Bucak et al. [52]) based learning may improve the performance of the model as well. We did not attempt to fine-tune the model by exploring all those possible approaches, because the principal focus of this paper was to analyze the potential of the BoW framework in damage classification.

The developed method can identify the damages related to debris/rubble piles that are strong indicators of building collapse or severe structural damage, which would be very useful for first responders involved in disaster response, but also other stakeholders such as governmental agencies assessing post-disaster construction needs, or insurance companies. However, for detailed building level damage assessment, this evidence alone is not sufficient to infer the complete damage state of the building, nor the total damage cost, as the latter also depends on internal (invisible) damage, and on building functions being affected, which is not always visible. However, along with other damage evidences such as cracks, inclined elements, etc., this evidence is also important in the damage classification process. From a practical point of view, especially, the observations we made using the combined dataset 4 (COM3109) are very interesting. Although the used patches vary significantly in terms of scale and complexity, an overall accuracy of around 90% was reached (cf. Table 6). Transferred to an actual disaster scenario, where quick interpretation of image data is needed, this would mean that an already existing database can be used to train a model and new images can be readily classified, and a similar overall accuracy might be expected. Hence, at least for a first damage assessment, the tedious manual referencing might not be necessary.

6. Conclusions and Outlook

A damage classification based on BoW was developed to classify a given image patch as damaged or non-damaged, irrespective of scale, image characteristics, and scene complexity. Various
combinations of image features (Gabor wavelets, HoG, and SURF) and supervised classifiers (SVM, RF, and Adaboost) are tested in both the BoW framework and conventional global feature representation approach using four different datasets. The BoW framework outperformed conventional global feature representation approaches in all scenarios (i.e., for all combinations of feature descriptors, classifiers, and datasets) and produced an average accuracy of approximately 90%. Although the developed model can identify the damaged regions in the images well, it cannot classify the detected damaged regions into specific types, such as debris, rubble piles, spalling, and inter-story collapse. We need contextual information and 3D geometric features such as shape, location, characteristics of the neighboring elements, and local height variation of the damaged region to identify the actual category of damage. For example, the damage patterns on large intact planar elements could be classified as spalling, whereas the damage pattern on the ground with large local height variations and no large 3D segments could be classified as debris. Therefore, the potential extension of this work will be the development of methods for classification of the detected damaged regions into actual damage categories.

As stated earlier, the feature descriptor component in the BoW framework has a significant impact on the performance of the model. Here, the texture features are chosen to examine our BoW framework as their potential in damage detection has been demonstrated well by previous studies, as highlighted in the Introduction. However, recent studies report that supervised feature learning methods such as convolutional neural networks (CNN) could learn the feature and its representation directly from the image pixel values chosen for a specific application [53]. Hence, these features are referred to as data-adaptive features and are found to be superior to well-proven handcrafted features such as Gabor and HoG for many computer vision applications including image classification [54,55]. Therefore, we intend to explore the potential of CNNs for damage classification in the future.

Acknowledgments: This work was funded by the EU-FP7 project RECONASS (Reconstruction and Recovery Planning: Rapid and Continuously Updated Construction Damage and Related Needs Assessment; grant no 312718). We thank Pictometry, Inc. for providing the imagery used in this study.

Author Contributions: Anand Vetrivel and Markus Gerke conceived the design and implementation of the method and wrote the manuscript. Norman Kerle and George Vosselman helped to improve the method and experiments, and also reviewed and improved the manuscript structure and writing.

Conflicts of Interest: The authors declare no conflict of interest.

References


34. Conde, C.; Moctezuma, D.; Martín De Diego, I.; Cabello, E. HoGG: Gabor and HoG-based human detection for surveillance in non-controlled environments. Neurocomputing 2013, 100, 19–30. [CrossRef]


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
An Automatic Procedure for Early Disaster Change Mapping Based on Optical Remote Sensing

Yong Ma1,2, Fu Chen1,*, Jianbo Liu1, Yang He1,2, Jianbo Duan1 and Xinpeng Li1

1 Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China; mayong@radi.ac.cn (Y.M.); liujb@radi.ac.cn (J.L.); heyang@radi.ac.cn (Y.H.); duanjb@radi.ac.cn (J.D.); lixp@radi.ac.cn (X.L.)
2 University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: chenfu@radi.ac.cn; Tel.: +86-010-8217-8158

Academic Editors: Zhenhong Li, Roberto Tomas, Norman Kerle and Prasad S. Thenkabail
Received: 27 October 2015; Accepted: 11 March 2016; Published: 26 March 2016

Abstract: Disaster change mapping, which can provide accurate and timely changed information (e.g., damaged buildings, accessibility of road and the shelter sites) for decision makers to guide and support a plan for coordinating emergency rescue, is critical for early disaster rescue. In this paper, we focus on optical remote sensing data to propose an automatic procedure to reduce the impacts of optical data limitations and provide the emergency information in the early phases of a disaster. The procedure utilizes a series of new methods, such as an Optimizable Variational Model (OptVM) for image fusion and a scale-invariant feature transform (SIFT) constraint optical flow method (SIFT-OFM) for image registration, to produce product maps including cloudless backdrop maps and change-detection maps for catastrophic event regions, helping people to be aware of the whole scope of the disaster and assess the distribution and magnitude of damage. These product maps have a rather high accuracy as they are based on high precision preprocessing results in spectral consistency and geometric, which compared with traditional fused and registration methods by visual qualitative or quantitative analysis. The procedure is fully automated without any manual intervention to save response time. It also can be applied to many situations.

Keywords: disaster emergency mapping; optical image; automatic procedure; OptVM; SIFT-OFM

1. Introduction

Natural disasters, such as earthquakes, landslides, avalanches and debris flows, occur unexpectedly and suddenly, claiming huge losses of life and property and causing significant damage to the surrounding environment. A large number of cases show that rapid emergency response is effective in reducing the casualties and the loss caused by disasters. However, effectively making proper contingency plans is a troublesome problem faced by governments and experts [1].

Remote sensing technologies have the unique ability to help emergency managers streamline response and recovery by providing a backdrop of situational awareness which can be invaluable for assessing the impacts of the damage and to guide the disaster rescue [2]. Remote sensing plays an important role for the disaster rescue from the very early phase of a disaster, right through to long-term recovery [3]. Especially during the early stages of disaster response, a coordinated and reasonable plan for search and rescue (S&R) activities, logistics planning, and monitoring can save lives and property. However, the information (e.g., the geographical scope of the disaster areas, the magnitude and the spatial distribution of damage, and transportation infrastructures conditions) to support rapid response is limited. Remote sensing is a critical option for obtaining a comprehensive information about the disaster and for assessing the scope of damage, in particular to remote areas, where other means of assessment or mapping either fail or are of insufficient quality [4].
Various type of remote sensing sensors, platforms and techniques can be considered for emergency mapping [5]. The choice is mainly based on the type of disaster, the approximate extent of the affected areas and the requirement for monitoring the event. Generally, the main source of data for response activities is satellite remote sensing, as this gives the ability to monitor a wide footprint on the ground with limited or no access. In addition, modern satellite platforms can be triggered to change the acquisition angle to cover the affected areas in a short time for increasing the observation frequency of the regions of interest. The satellite imagery of disaster-hit areas can be obtained every day or even every few hours. Regarding the sensor type, the synthetic aperture radar (SAR) data and optical high-resolution (HR) or very high-resolution images (VHR) are generally obtained when disaster occur. SAR or InSAR systems are of great value for disaster response as it is a utility further enhanced by its all-weather capabilities, especially when the persistent of cloud cover over the affected areas make optical data unusable [6]. It has been achieved success in detecting surface displacement and offsets and height change of disaster areas by using the intensity, coherence or phase information of post-event SAR or InSAR data. However, for structural damage or some temporary change (e.g., damage buildings and the shelter sites), SAR and InSAR data are insufficiently sensitive and the final mapping results have been less conclusive, and marked by uncertainties. And most SAR-based change detection approaches suffer from a lack of archive data with the same acquisition parameters as the post-crisis imagery [7]. In addition, SAR imagery is less intuitive and is difficult to be interpreted by nonexperts, who must rely on sophisticated software to analyze the interferometry or amplitude coherence that require a longer processing time. Almost from the very onset of the disaster, optical satellite imagery is always available and provides the first glimpse of the disaster devastation. Optical high resolution imagery is the preferred choice to delineate the feature of interest (e.g., building, tents) and their current status (e.g., destroyed, burned, or moved) [5]. Generally, optical data can provide useful information to discriminate between damaged or non-damaged areas. Even though no further classification with respect to damage levels can be retrieved, the information derived from optical imagery can be effective [8]. The advantage of optical data is that their interpretation is intuitive for nonexperts, it can provide an overview of affected regions and sufficiently detailed information for decision-makers or aid workers to make S&R plan. However, the presence of cloud and shadows and variations in observation angles and geometric distortions limit the application of optical imagery for rapid emergency mapping [9].

This paper focuses on optical remote sensing image processing for the disaster mapping. One widely used approach for producing damage mapping is visual interpretation [10,11], which is tedious and labor intensive. Several automatic methodologies have been presented for disaster mapping, however, geometric distortions and improper co-registration between pre- and post-disaster imagery can result in a high false alarm rate in these automatic change detection approaches [7]. Therefore, a balance must be found between the product accuracy and the labor-cost, and a compromise must be found between the response time, the analysis depth, and the mapping accuracy. In this respect, we propose an automatic Optimizable Variational Model (OptVM) and scale-invariant feature transform (SIFT) constraint optical flow method (SIFT-OFM) to improve the accuracy of change mapping to produce the preliminary damage map, which usually will have a rather coarse character. In the early period of disaster, these initial maps could quickly provide some crucial information to relief teams and support the planning and coordination of rescue operations for emergency response. At the same time, these maps would give professional analysts or the public a guide to further checking whether a detected change is correct to improve the accuracy of the damage map by visual validation. With the availability of further earth observation data and more in-depth image analysis, the maps can be updated to incorporate this new information and provide a refined damage assessment [7].

Therefore, we propose an automatic optical remote sensing process to produce a cloud free product as a situation awareness map for identifying the overall damage and potential cascading effects, a change detection map for estimating the impacts of damage, and a temporal change map of key areas (such as the heavily damaged areas and temporary resettlement of victims) quickly in
order to monitor the recovery process and guide the disaster recovery. This process can rely on the newest available optical data to rapidly produce the up-to-date maps, cover large parts of the affected area, and enable disaster managers to obtain an overview of the situation, assess the damage, and supply local logistic teams with reliable information on very short notice. The procedure and methods are introduced in Section 2. The experimental results and analyses, using the Nepal magnitude 8.1 earthquake on 25 April 2015, as an example, are given in Section 3. A discussion is presented in Section 4. The last section concludes the paper with the significance of this research and proposed future work.

2. The Procedure and Methods

The timeliness of receiving data and the length of time needed for data processing are key factors in rapid emergency mapping for disaster events. When a large disaster occurs (e.g., the Haiti earthquake, the forest fires in Australia, and the Nepal earthquake), many international organizations or commercial companies decrease the satellite revisit time to obtain the newest data for the affected areas by adjusting the acquisition angle [12,13]. This timely triggering provides enough data for disaster mapping. However, these agencies only offer the original satellite imagery, rather than processing or analysis results. The orderly organization and management for effective processing is scanty; and these multi-source and multi-temporal data are not sufficient for developing information applicable in humanitarian and natural crisis situations.

In this paper, a procedure is proposed to deal automatically with multi-source and multi-temporal remote sensing data, in order to produce a product that can serve in the early stages of disasters. The procedure is shown in Figure 1, which includes image fusion, image registration, and the production of maps. The procedure ultimately offers a cloudless remote sensing map, a change detection map, and a local temporal variation map for some typical areas during a short period of time. The proposed procedure is based mainly on the general rapid mapping framework proposed by Ajmar and Boccardo [14], but our proposed procedure is a specific implementation of a general framework except for map dissemination. In our procedure that uses some new techniques and methods, the accuracy of the products in each step has been greatly improved. Additionally, the whole procedure is fully automated without any artificial participation. The methods applied in each step of the procedure are described below.

![Figure 1. Procedure for dealing with multispectral and multi-temporal remote sensing images.](image-url)
2.1. Image Fusion

Recently high spatial resolution images, such as QuickBird, WorldView, GeoEye and GF-1, have been favored in many remote sensing applications, especially for the assessment of disasters [15]. Usually, high spatial resolution satellites provide image data with a low-resolution multispectral (LRM) image and a high-resolution panchromatic (HRP) image [16]. In order to obtain the high-resolution multispectral (HRM) images needed for detailed damage assessment, the image fusion process must be implemented. This is the first key step in our procedure.

There has been much research done on image fusion. The methods can be categorized into five types, including variable substitution (VS) methods, such as the IHS and PCA algorithms [17,18]; modulation fusion methods (MF), such as the Brovery algorithm [19]; multi-scale analysis (MSA) such as the wavelet fusion algorithm [20]; restoration based methods (RB), such as the adjustable model-based fusion algorithm [21]; and compressed sensing methods (CS) [16,22]. The first three types of methods can be viewed as traditional methods, which tend to cause color distortions. Additionally, spatial information weakens due to the wavelength extension of the new satellite HRP images. Inspired by the rapid development of super-resolution techniques, RB methods have been applied routinely to image restoration and fusion. These methods view the LRM image and the HRP image as the observations of the HRM image via an image degradation model. The HRM image is obtained by setting up a model and solving the model. The CS method is based on sparse representation, which utilizes the strong correlation between the LRM and HRP images to obtain a sparse coefficient, construct two coupled dictionaries, and reconstruct the fused image by the fused measurement [23]. This method usually needs a large collection of training images to obtain trained dictionary, requiring significant computing time.

In this paper, an improved fusion model, OptVM, is applied to produce high-quality HRM images automatically. The flowchart of the OptVM is shown in Figure 2. First, a PAN image is simulated by MS image, and the simulated PAN and original HRP image are matched. Next, the restored MS image is obtained by the relationship of spectrum and grayscale between the simulated PAN and LRM images; and finally, the model (Equation (1)) is optimized to get the HRM image. The model defines Pan-sharpening as an optimization of a linear over-determined system, which is based on gray-value and spectral sensitivity relationships between the original multispectral (MS), panchromatic (PAN), and fused images.

![Figure 2. Flowchart of image fusion method.](image-url)
The OptVM constructs a cost function to generate the HRM image based on three constrained hypotheses: (1) the radiation of an object in an LRM image should be equal to the integral average radiation of the same object in the HRM image; (2) spatial details of the HRP image and the HRM image (the target fused image) should be similar; (3) the values in the panchromatic band should be the summation of the respective multispectral band, if the spectral response range of the LRM image is nearly the same as the HRP image. According to the three assumption, the cost function can be defined as follows in Equation (1):  

$$E(P_H, X_{SH}, X_{SL}) = \lambda_1 \sum_{n=1}^{4} \int (X_{SL}^n(x) - k \times X_{SH}^n(x))^2 \, dx + \lambda_2 \sum_{n=1}^{4} \left( X_{SH}^n(x) - \sum_{s \in N(x)} W_{sx} X_{SH}^n(s) \right)^2 \, dx + \lambda_3 \frac{1}{2} \left( \sum_{n=1}^{4} W_{sn} X_{SL}^n(x) - P_H(x) \right)^2 \, dx$$

(1)

where $X_{SL}$, $X_{SH}$ and $P_H$ are the radiation of pixels in the LRM image, the HRM image and the HRP image respectively. The variables, $\lambda_1$, $\lambda_2$ and $\lambda_3$ represent the weight of the three models in the total energy function, which can be adjusted according the fusion result. Here, in next experiment, $\lambda_1$, $\lambda_2$ and $\lambda_3$ are set 1. The variables, $n$ is number of fusion band. The parameter of $k$ is based on the fact that the low-resolution pixels are formed from the high-resolution ones by low pass filtering followed by subsampling [24], and $k$ represents the low pass filtering followed by subsampling. The express of $k$ is shown as follows in Equation (2):

$$X_{SL}(x) = k \times \sum_{j \in M(x)} P_H(j)$$

(2)

In the Equation (2), the pixel $j$ is one of the pixels set $M$ in HRP image which corresponds to the pixel $x$ in LRM image. For example, if the resolution of LRM image is 8 m and HRP image is 2 m, each pixel in LRM image corresponds to $4 \times 4$ pixels set in HRP image in the same object. The variables in Equation (1), $W_{sx}$ represents the weighted value of the pixel $s$ which in the neighborhood window of pixel $x$ in HRP image, the express is as follows Equation (2), including $\delta_x$ is the standard deviation.

$$W_{sx} = e^{-(P_H(x)-(P_H(s))^2)/2\delta_x^2}$$

(3)

In the Equation (3), the size of neighborhood window is set $9 \times 9$ in the next experiment. The variables in Equation (1), $W_n$ is the band weights of multispectral image in spectral response relationship. The express is shown as follows in Equation (4):

$$P_H(\lambda) = \sum_{n=1}^{4} W_n X_{SL}(\lambda)$$

(4)

In the Equation (4), $\lambda$ represent the Spectral value. The process of image fusion is the global minimum of $E(P_H, X_{SH}, X_{SL})$, which is the linear constraint of the three basic models. The least square technique—LSQR [25] is applied to optimize the model. An important step of the processing is the image registration of the HRP image and the simulated panchromatic image by the LRM image to avoid fusion image distortion and spectral offset based on the optical flow and feature constraint method (introduced in Section 2.2).

2.2. Image Registration

Image registration is an important procedure in remote sensing image applications [26] and a crucial step in our chain. Previous studies of image registration methods can be divided into three
categories, including area-based method, such as Cross-Correlation (CC) method [27], the mutual information (MI) [28] and the Sequential Similarity Detection Algorithms (SSDAs) [29]; feature-based method such as SIFT [30–32] and physically-based methods, such as multi-resolution elastic model [33] and the optical flow model [34,35]. The first two methods, which are based on the gray information and the spatial features of the images to register the images, cannot meet the real time requirements, as they require large computing capacity and highly complex algorithms, They are also not very robust. The physically based methods utilize physical models to describe the process of image deformation and solve the image physical model to achieve image registration. These methods have great advantages for speed and accuracy. In this paper, a physically-based method, the SIFT feature constraint optical flow method (SIFT-OFM), is proposed to register images. Figure 3 shows the flowchart of the algorithm. First, homologous points are obtained for the reference image and the sensed image using the SIFT feature matching method [36]. Then the affine transform coefficients are calculated to be a constraint for an optical flow model. Finally, the parameters of optical flow model are optimized through an iterative process to obtain high accuracy registration images.

![Flowchart of image registration method.](image)

The optical flow model method is based mainly on the assumption of the local gradient model and the overall smoothness model of the pixel gray values, which regard the change of image gray scale as the change of optical flow [35]. The algorithm is able to get the high-precision matched image by considering the optical flow change as the displacement of the reference image and the sensed image to the register.

The optical flow model is based on the assumption of the constancy of brightness, which means that the gray value of the same object is not changing with time and the radiance consistent is in...
motion [37]. Thus, the spatial variation of the optical flow model can be expressed as follows in Equation (5):

\[
E_{\text{gray}} = \int \int \varphi \left( (I(x + u, y + v) - I(x, y))^2 \right) dx dy
\]  

(5)

where \( I \) is the image intensity function of interval \( \Omega \) \((x, y) \in \Omega \); \( u \) and \( v \) are the displacements in the \( x \) and \( y \) directions, respectively, which is the offset between the reference image and the sensed image. The symbol, \( \Psi \), represents a differentiable function on the interval \( \Omega \).

To reduce the impact of light changes, the gradient energy equation is added into the total energy formula. The gradient energy equation is defined as follows in Equation (6):

\[
E_{\text{gradient}} = \int \int \varphi \left( |\nabla I(x + u, y + v) - \nabla I(x, y)|^2 \right) dx dy
\]  

(6)

where \( \nabla I = [I_x, I_y]^T \), and \( \nabla I \) is the gradient vector of the image. For every pixel, there is only one equation (the sum of Equations (2) and (3)); but there are two variables, \( u \) and \( v \), which can cause uncertain solutions and aperture problems. Thus, there is a need to add a regularization constraint. The regularization constraint plays an important role in the estimation intensity, the edge smoothness, and the edge retention of the optical flow field. In this algorithm, a global smoothness constraint proposed by Horn [37] is used to minimize the constraint term. The global smoothness constraint is described as follows in Equation (7):

\[
E_{\text{smooth}} = \int \int \varphi \left( |\nabla u|^2 + |\nabla v|^2 \right) dx dy
\]  

(7)

where \( \nabla := (\partial x, \partial y) \) is a two-dimensional differential operator.

Optical flow model is not suitable for matching large displacement images since the optical flow model can be expressed by expanding the Taylor series and omitting the high-order terms. Therefore, the SIFT-OFM adds the SIFT feature as a constraint into the optical flow model. The generation of the SIFT feature points includes four steps: extreme space detection, precise positioning of the SIFT feature point, determination of the main direction of the feature point, and the description of the building features. Based on these SIFT feature points, the affine coefficients are calculated as the initial value of the optical flow field to match the image. The total energy equation of SIFT-OFM is defined as follows in Equation (8):

\[
E = E_{\text{gray}} + \lambda E_{\text{gradient}} + \alpha E_{\text{smooth}} + \tau E_{\text{SIFT}}
\]

\[
= \int \int \varphi \left( (I(x + u, y + v) - I(x, y))^2 \right) dx dy \]

\[
+ \lambda \int \int \varphi \left( |\nabla I(x + u, y + v) - \nabla I(x, y)|^2 \right) dx dy \]

\[
+ \alpha \int \int \varphi \left( |\nabla u|^2 + |\nabla v|^2 \right) dx dy \]

\[
+ \tau \int \int \varphi \left( (u - u_{\text{SIFT}})^2 + (v - v_{\text{SIFT}})^2 \right) dx dy
\]

(8)

where \( \lambda, \alpha \), and \( \tau \) are the coefficient of the last three component. Based on previous experience, these are set to \( \lambda = 1, \alpha = 50 \) and \( \tau = 1 \) in next experiment.

The minimum of \( \vec{u} \) and \( \vec{v} \), are solved based on the energy function \( E \), using Successive Over Relaxation (SOR) iterative method [38]. This is the displacement of the reference image and the sensed image. Then calculate the displacement based \( \vec{u} \) and \( \vec{v} \) and resample to get high-precision matched image.

2.3. Map Production

The procedure mainly provides two kinds of products. These are the cloudless maps for the disaster regions and the change detection maps for the early stage of the disaster. The flowchart for map production is shown in Figure 4. After obtaining the high precision matched disaster image, the clouds on the images are detected and removed to get cloudless remote sensing maps for determining
the overall scope of the disaster and for automatically detecting the change between the pre- and post-disaster cloudless images. For the automated change map, artificial checking is needed to exclude the error detection, which is caused by the different types of sensors and the observed conditions between the two images. In cases of emergency, the automated coarse change map can be used directly to offer approximate information for managing crisis issues.

Figure 4. Flowchart of products processing.

Much research has been focused on the task of cloud removal to relieve the influences of clouds. In fact, the process of cloud removal is an information reconstruction process for missing information. The proposed approaches can be grouped into three categories [39,40]: (1) spatial-based methods without any other auxiliary information source; (2) spectral-based methods, which extract the complementary information from other spectra; and (3) temporal-based methods, which extract the complementary information from other data acquired at the same position and at different time periods. Spatial-based methods are often used for some small missing regions by interpolation, such as for the scan line (SLC) missing from Landsat-7 [41]. It is not appropriate for large-scale cloud removal. Spectral-based methods are based on the relationship of spectra in clear regions and restore the contaminated band of data by modeling a relationship between the contaminated band and the auxiliary band by methods such as the HOT method [42], and the STARFM method [43]. It is usually constrained by the spectral compatibility and has difficulty with thick clouds. Temporal-based methods utilize multi-temporal clear data for the same region to simulate the cloud cover area data for acquiring cloud-free images, such as local linear histogram matching (LLHM) approach [44], spatio-temporal MRF model [45] and etc. In this paper, we produce the cloud free maps by the temporal phases simulation method proposed by Gui [46], which utilize spectrum relationship, vegetation index and HOT parameters [47] to detect clouds on the images and calculate the feature difference between two temporal images to make up the cloud regions. The simulation method uses the minimum value of the feature difference and a multi-grid optimization algorithm. It converts the features of one image to the features of the reference image and then fills the areas of the clouds by the converted image that has no
cloud in the same region. This generates a cloud free image in the disaster districts. In this study, time to acquire the difference between the images is very short, so the generated image can truly reflect the overall situation after the disaster by this method.

Several authors have presented semi- or fully-automatic methods to analyze images for damage change detection and assessment by optical data [3, 48–50]. But geometric distortion, improper co-registration of pre- and post-disaster images can result in a high false alarm rate of automatic change detection methods. In this procedure, the new methods, OptVM and SIFT-OFM, are utilized to improve the accuracy of preprocessing results. And based on the high-precision preprocessing images, the Iteratively Re-weighted modification to the Multivariate Alteration Detection (IR-MAD) method proposed by Nielsen [51, 52] is used for detecting change automatically in this procedure. IR-MAD is mainly based on the means of canonical correlation analysis (CCA). First, the maximum canonical correlation coefficient after linear transformation is determined for the pre- and post-disaster images [53]. The D-value and the variance of the linear transformed images are calculated. Finally, the pixel change is checked by the Chi-square distribution of the D-value and the variance. In order to highlight the changing information of the pixels to improve the accuracy of the change detection map, a weighted iteration method is used during the process of transformation. The iterations continue until the absolute value is less than the error limit (for this paper, the error limit is set as 0.001). The absolute value is the difference of correlation coefficient maximum value from two iterations, namely, before and the current conversion.

However, the change detection map has some error detection points because the map, which is generated by IR-MAD, could be influenced by the different conditions during the acquisition of the two images. Artificial checking is essential for removing the errors in the changed regions. Changes are totally possible during the early stages of disasters, since many people and agencies from other regions or countries have a great enthusiasm and willingness to help the victims and local government in relief and reconstruction activities. There are many public community remote sensing platforms, such as TomNod [54] and GIS Corps [55] which could add and update the information (not only in vector format but also by means of geotagged pictures and textual information), validate the information conveyed, and actively participate in the map production phase by coordinating the efforts of ordinary people and experts [56]. It could gather every tiny force to provide significant assistance for the victims and decision making department in the disaster region [57]. The automatic production change map can be used in on a community remote sensing platform to reduce labor costs and to provide a rapid response to the disaster.

3. Experiments Results and Analysis

3.1. Study Area and Data Sets

A powerful earthquake shook Nepal near its capital, Kathmandu, at 11:56 NST on 25 April 2015, killing more than 8000 people and flattening section of the city’s historic center. Some of the oldest parts of the city, including seven UNESCO World Heritage sites, were severely damaged in the earthquake. This paper uses the Nepal earthquake as an example. The location of the study area, the center of Kathmandu, is shown in Figure 5. Four images from a GF-1 PMS sensor were selected, including a pre-earthquake image on April 11 and post-earthquake images on 27 April, 1 May, and 2 May, as shown in Figure 6. The specific observation parameters of the data set are shown in Table 1, and the parameters for the PMS GF-1 sensor are shown in Table 2. The proposed procedure was followed without any human intervention. The procedure will be discussed clearly, including the intermediate results of the image fusion and registration and the final products of the cloudless maps and the change map.
Figure 5. Spatial location of the study area (the resolution of left map is 2 m after fusion).

Table 1. The specific observations parameters of data set.

<table>
<thead>
<tr>
<th>Acquisition Date</th>
<th>Sensor</th>
<th>Viewing Angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 April 2015</td>
<td>PMS-1</td>
<td>0°</td>
</tr>
<tr>
<td>27 April 2015</td>
<td>PMS-2</td>
<td>−24°</td>
</tr>
<tr>
<td>1 May 2015</td>
<td>PMS-1</td>
<td>−23°</td>
</tr>
<tr>
<td>2 May 2015</td>
<td>PMS-1</td>
<td>23°</td>
</tr>
</tbody>
</table>

Table 2. Parameters of GF-1 PMS sensor.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>PMS-1/MS-2 Sensor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral range</td>
<td>PAN 0.45–0.52 μm</td>
</tr>
<tr>
<td></td>
<td>MS 0.52–0.59 μm</td>
</tr>
<tr>
<td>Spatial resolution</td>
<td>PAN 2 m</td>
</tr>
</tbody>
</table>
3.2. Image Fusion Results and Analysis

In this section, the image from 27 April is taken as an example to show the accuracy of the image fusion results. Figure 7 shows part of the whole image to exhibit the image details clearly. The left image is multi-spectral data, and the right is panchromatic data. The upper left corner image in each view is an enlarged view of the small red box. Figure 8 shows the result of PCA fusion, GS-pansharpening fusion, wavelet transform fusion, and the proposed OptVM method. Visually, the four methods have achieved spatial resolution enhancement and keep a certain spectral fidelity; but the results of PCA, GS, and wavelet transforms have virtual phenomena in the edge of objects, especially the wavelet transform result being the worst. In detail, from the enlarged image of the small red box, the resulting color recovery of OptVM is finer than with the other three algorithms. For example, observe the red roof in the enlarged image in Figure 8. The OptVM method restored the red

![Image of data set](image-url)
color of the roof regularly, while the color recovery results of the other three algorithms deviate from the red color, with some parts of roof being white.

![Image of multispectral and panchromatic data](image)

**Figure 7.** Multispectral (left) and panchromatic data (right) of the original image on 27 April (upper-left small image is the detail from the area of the red rectangle).

Six typical surface features are shown in Figure 9 which are sampled from the fused results (seen in Figure 8). These are sampled in relatively uniform areas to assess the spectral fidelity and the spatial location of the six points marked in Figure 9A. These pixels cover a large set of material in the scene and are uniform around their neighborhood. The spectra of the original MS image are used as ground truth for comparisons based on the reasonable assumption that the spectra around such uniform areas are most likely unchanged. The spectra results of the image fusion methods are shown in Figure 10; and the Euclidean distance difference and the spectral difference, expressed in spectral angle difference, are shown in Table 3. For uniformity of comparison, the values of the fusion results are normalized to 0–255. From Figure 10 and Table 3, it can be seen that the result of OptVM method can preserve the spectra of these pixels well, whereas the spectra from the PCA, the Gram–Schmidt (GS), and the wavelet fusion are very different from the truth. For example, the red roof of pixel #1 is shifted seriously in the result from wavelet fusion. For the ordinary roof (pixel #4) and bare land (pixel #6), the results of PCA and GS are lower than the truth, especially in red band, and the difference of the spectral angle is larger. In the vegetation area, the colors of the results based on the wavelet method (see Figure 9D) lack fidelity, while our algorithm performs very well in most regions (Figure 9F). In our method, a crucial step is that the simulated PAN and the original PAN need to be matched to reduce the distortion that is caused by the geometric deviation of the original MS and PAN images.
Figure 8. Result of various image fusion methods using (A) PCA; (B) wavelet; (C) GS; and (D) OptVM.
Figure 9. Comparison of the various algorithms on the urban area scene. (A) the PAN image; (B) the MS image after resampling using the nearest neighbor; (C) the PCA result; (D) the wavelet result; (E) the GS result; and (F) the OptVM result.

Figure 10. Spectral comparison between the various approaches in the urban area scene. The site of pixels marked in Figure 9A.
Table 3. Spectral angle difference and Euclidean distance difference of marked pixels.

<table>
<thead>
<tr>
<th>Category</th>
<th>Spectral Angle Difference</th>
<th>Euclidean Distance Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PCA</td>
<td>Wavelet</td>
</tr>
<tr>
<td>Red roof</td>
<td>29.75</td>
<td>25.71</td>
</tr>
<tr>
<td>Water</td>
<td>41.33</td>
<td>16.79</td>
</tr>
<tr>
<td>Grass</td>
<td>26.65</td>
<td>10.72</td>
</tr>
<tr>
<td>Roof</td>
<td>39.01</td>
<td>14.90</td>
</tr>
<tr>
<td>Road</td>
<td>26.44</td>
<td>12.37</td>
</tr>
<tr>
<td>Bare land</td>
<td>26.44</td>
<td>9.00</td>
</tr>
</tbody>
</table>

In order to comprehensively assess the quality of fused images, four objective evaluation metrics are selected, which include the spectral angle mapper (SAM), the error of relative dimensionless global in synthesis (ERGAS) metric, the Q4 index [58] and Quality with No Reference (QNR) metric [59]. The equations of four metrics are presented in the Appendix. SAM and ERGAS are measures of the global radiometric distortion of the fused images, and a smaller value indicates better fusion results. Q4 and QNR reflect the spectral and spatial quality of the fused images. Values closer to 1 denote the better result. The results of the four metrics are shown in Figure 11 and Table 4. In regards to the fidelity of the image spectrum and the image clarity, the fused images by GS and by our proposed algorithms appear better than those obtained by the PCA and the wavelet fusion methods. The QNR index obviously displays this difference, since replacing the original MS image information leads to the loss of information. Compared with the other fused methods, the SAM value of the GS fused image is very large, with the degree of sharpening being excessive in some parts of the results. The values for the four metrics for our method show that it is the best of the four methods. In fact, most of the existing image fusion algorithms are based on the band-to-band correlation, whereas our method synthetically utilize the relationship of spectrum, grayscale and spatial to produce an image with fewer distortions.

Figure 11. Line chart of the four metrics comparing fusion methods (ERGAS and SAM values are better when approaching 0, while Q4 and QNR values are better when approaching 1).

Table 4. Comparison of the value of above four index.

<table>
<thead>
<tr>
<th>Index</th>
<th>PCA</th>
<th>Wavelet</th>
<th>GS</th>
<th>OptVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERGAS</td>
<td>4.118</td>
<td>6.551</td>
<td>4.214</td>
<td>3.485</td>
</tr>
<tr>
<td>SAM</td>
<td>3.566</td>
<td>2.957</td>
<td>3.656</td>
<td>2.492</td>
</tr>
<tr>
<td>Q4</td>
<td>0.533</td>
<td>0.527</td>
<td>0.846</td>
<td>0.894</td>
</tr>
<tr>
<td>QNR</td>
<td>0.878</td>
<td>0.886</td>
<td>0.86</td>
<td>0.931</td>
</tr>
</tbody>
</table>

3.3. Image Registration Results and Analysis

Image registration is based on the results of image fusion and select two images as the example (pre-earthquake on 11 April 2015, and post-earthquake on 27 April 2015). The pre-earthquake image is the reference image to match with the post-seismic image of 27 April. In Figure 12, the middle image
is a combined image, which takes a $100 \times 100$ block from the reference image and also from the sensed image, alternately. Figure 12A–D are the detailed enlarged images, which show the position offset of the reference image and of the sensed image. In Figures 13–16 the middle combined image is obtained in the same way as Figure 12. These figures are based on the reference image and the matched image using the optical flow model algorithm (Figure 13), the SIFT feature algorithm (Figure 14), the ENVI software registration workflow (Figure 15), and our proposed SIFT-OFM method (Figure 16).

**Figure 12.** The combined image of the reference image and the sensed image (A–D show the detail of each red rectangle on the combined image).

**Figure 13.** The combined image of the reference image and the matched image only using the optical flow model (A–D show the detail of each red rectangle on the combined image).
Figure 14. The combined image of the reference image and the matched image only using the SIFT algorithm (A–D show the detail of each red rectangle on the combined image).

Figure 15. The combined image of the reference image and the matched image by using GCP in ENVI software (A–D show the detail of each red rectangle on the combined image).

Figure 16. The combined image of the reference image and the matched image using SIFT-OFM (A–D show the detail of each red rectangle on the combined image).
From the detailed images in Figure 12, the dislocation between the reference image and the sensed image is very obvious, so it is necessary to register the two images for producing the products. From Figure 14 to Figure 16, the results are good. Figure 13 shows that the result that is based only on the optical flow model algorithm is the worst. It is because optical flow algorithm model is not suitable for the large displacement of two images, generally not more than two pixels; but in this study, the difference of the two images significantly exceeds at least 10 pixels. Furthermore, comparing the detailed image from Figure 14 to that in Figure 16, the SIFT algorithm performed well in Figure 14A,D; but in Figure 14B, the image of the dislocation of the bridges is very clear, while the road in the Figure 14C image is broken. The linear correction model of the affine transformation is utilized in SIFT feature algorithm and can cause some nonlinear distortions to be uncorrected. The result of Figure 15 is based on ENVI registration process, in which 34 ground control points (GCP) are marked by artificial selection and the overall error of these GCP is not more than 0.43 pixels. However, in both Figure 15A,B the discrepancy of the bridges is obvious and serious. This is related to the maldistribution of the GCP and the linear calibration model in the registration process. There is no obvious displacement or deviation in any of the images of Figure 16 produced by the SIFT-OFM method, which restricts the larger deviation by the SIFT feature and corrects the local linear or nonlinear distortion by the optical flow model.

The structural similarity index measurement (SSIM) [60] is the quantitative evaluation metric used to evaluate the precision of a registration image. The SSIM maps in Figure 17 are obtained by calculating the SSIM between the reference image and the matched image in a sliding window within $9 \times 9$. In the SSIM map, red indicates SSIM values close to 1; that is, the areas with high precision registration. It can be seen that the precision of the proposed method is higher for the SIFT feature and the ENVI automatic registration process. The existence of clouds and shadows in sensed images and the lack of clouds in the reference image leads to the low SSIM value in the clouds and shadows regions. Therefore, in the Table 5, the overall SSIM value in the whole image shows relatively low.

<table>
<thead>
<tr>
<th>Index</th>
<th>SIFT</th>
<th>ENVI</th>
<th>SIFT-OFM</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSIM</td>
<td>0.6266</td>
<td>0.6415</td>
<td>0.7442</td>
</tr>
</tbody>
</table>

Figure 17. SSIM maps of three methods (the red area indicates SSIM values close to 1; that is, the more high precision registration regions).

3.4. Cloudless Product and Analysis

The generation of cloudless maps includes cloud detection and temporal simulation based on the result of image fusion and image registration. The four images after image fusion and image registration are shown in Figure 18. From Figure 18, all the post-event images (images of 27 April, 1 May, and 2 May) included clouds and shadows, which affect the ability of the decision makers and analysts to estimate and grasp the overall situation. Therefore, cloud and shadow removal and temporal simulation for producing a cloud free image is indispensable.
The data set of four images after image fusion and registration (compared with Figure 6, the resolution and the registration accuracy have been improved greatly).

The cloud detection results by our method introduced in Section 2.3 are shown in Figure 19, in which red areas represent cloud regions and green areas represent shadow regions. By visual comparison, most of the cloud and shadow have been detected, while only a small portion of the thin clouds have not been detected. The cloud and shadow based on the result in Figure 19 is used to produce the cloud free image by temporal simulation based on the adjacent time image after the earthquake. The cloudless image is shown in Figure 20, in which Figure 20A is the image of 1 May and Figure 20B is the image of 2 May. Figure 20C is the cloudless map. The image, Figure 20C, is based mainly on the Figure 20B image, with the cloud and shadow regions made up from the Figure 20A image. Figure 21 show the detailed enlarged images of the rectangle in Figure 20. In Figure 21, the scenes in the cloud and shadow regions are recovered by the temporal simulation method and keep consistency in color and texture using neighbors. It provides a high quality cloudless image and identifies the overall damage and potential cascading effects accurately. Figure 22 shows the cloudless maps of the post-event on 27 April, 1 May, and 2 May. Comparing with the original image of Figure 18, the cloudless maps eliminate the influence of clouds and shadows and can be used easily as backdrop maps or situation maps. The cloud and shadow regions of the image on 27 April (seen Figure 18
right-up image and Figure 19 right-up image) is replaced by the pre-event cloudless image on 11 April (seen Figure 18 left-up image) to obtain a cloudless map (seen Figure 22 left image), as there is no post-event image before 27 April to make up the cloud and shadow regions. The left cloudless image of Figure 22 contains some pre-event information, which may make a confusion by users if they only use the single cloudless map, so some red vectors of cloud and shadow regions are added to distinguish the pre and post event data. In the internal regions of the vectors are the pre-event information, while the external region are the post-event information. This kind maps is only used in the short time after the event when there are few post-event images and the images still have overlapping cloud-covered areas.

![Figure 19](image1.png)

**Figure 19.** The cloud detection products of the four image in Figure 18 (red areas represent cloud regions and green areas represent shadow regions).

![Figure 20](image2.png)

**Figure 20.** The cloudless map of 2 May. (I and II are the original images of 1 May and 2 May, III is the cloudless map after removing the cloud and shadow). (A) and (B) are the shadow and cloud regions in the original image; (C) and (D) are the same areas after removing the influence of cloud and shadow corresponding to (A) and (B) regions. The detail of A–D are shown in Figure 21).
Figure 21. The detailed enlarged images of the rectangle in Figure 20. (A and B are the original images from 2 May, C and D are the simulated cloudless images).

Figure 22. The cloudless maps for the post-event images from 27 April, 1 May, and 2 May.

3.5. Change-Detection Product and Analysis

The change-detection map produced by IR-MAD in this paper is a coarse detection map and needs further manual checks to improve the accuracy of the product. In this section, the four images of Figure 18 are tested. The image in Figure 23I is the change-detection map between 11 April and 27 April,
in which clouds and shadows are considered as the change, and Figure 23II is the change-detection map after removing the cloud and shadow regions based on the Figure 23I image. Figure 23A–D give the detailed enlarged images of the red rectangle in the large maps. Figure 23A illustrates the change of open grass land before and after earthquake, where some tents have been put up as temporary refuge. Figure 23B indicates the landmark building of the Kathmandu, Bhimsen Tower, completely collapsed after the earthquake. Figure 23C represents an error change detection result, as the different imaging angle of two images result in the different positions of tall buildings and their shadows in the images. Though the images are matched by the method proposed in Section 2.2, the differences caused by the viewing angle is very difficult to be corrected. Figure 23D also shows an error change-detection result, because part of the shadow area is missing from detection. From the above analysis, the change-detection map produced by our proposed procedure is a coarse result without distinguishing the type and grade of change and many errors are detected, which require manual checking to refine the work.

The image in Figure 24I is the same image shown in Figure 23II, which is the change map between 11 April and 27 April. The image in Figure 24II is the change between 1 May and 11 April and the image in Figure 24III is the change between 27 April and 1 May. All of these eliminate the impact of cloud and shadow. Figure 25 shows the detailed enlarged images of the red rectangles of Figure 24. In Figure 25A, the Durbar Square of Kathmandu, which damaged severely in the earthquake, is shown, but this change (the damage) is not discovered because the clouds and shadows exist in the image of 27 April (seen in Figure 25A). In Figure 24III, the change is observed, since the image of 1 May is cloudless in this region. For B sense in the Figure 24, the reason of discovering the change in II, not in I and III map is the same with the A sense as the influence of shadow (seen the detail in Figure 25B). Figure 25C shows the Kal Mochan Temple of Kathmandu, which also collapsed after the earthquake. In Figure 24I,II, the change area is the same large, but small in Figure 24III map. This is because the feature on the ground has no major changes in the post-earthquake images (27 April and 1 May), while the feature changed obviously between pre- and post-earthquake images (Figure 24I is the result of 11 April and 27 April, Figure 24II is the result of 11 April and 1 May). There are just some rocks from the collapsed temple that have been consolidated during the period from 27 April to 1 May to clean
the area of the ruins (shown in Figure 25C). By these comparisons, the progress of the rescue will be learned. The same reason applies to the Figure 25D,E scenes. From these images, we see that many tents were put up in some open space nearby the houses of victims as temporary shelters because of the constant large aftershocks. The temporary shelters were migrated or revoked on 1 May as the rescue proceeded. So some temporary facilities would disappear on the image of 1 May (seen in Figure 25D,E); and no or few change areas would be detected in image III of Figure 24. The results of this paper mainly provide a quick change detection map to determine the urgent requirements, as well as offer guidance to check the change regions precisely through a community remote sensing platform for experts or the public, which is possible to reduce labor costs and improve efficiency in the maximum extent. Figure 26 show the detected changes vector superposition on cloudless maps of 27 April. The vector of the left hand map is the final result, which was extracted from the image I of Figure 26; and the vector of the right hand map removes the error regions of the final result by visual checking. The right vector map is the combination of visual interpretation by four independent people. This vector map can be considered the relative correct result for comparing. The simple statistic is shown in Table 6. The number of change regions in the original vector map (shown in Figure 26 left hand map) is 758, while the number in the checked vector map (shown in Figure 26 right hand map) is 556, and the error detected change regions are 202. The percentage of truth change regions and original change detected regions is 73.4%. This can meet the application requirements to a certain degree. In the true change regions, most of them are tents or shelters (shown Figure 23A, Figure 25D,E, and Figure 27A), some are cars on the street or other temporary features, and a few are the collapsed or damaged buildings (shown in Figures 25C and 27B). In the error change regions, as analyzed above, most of the error is caused by the changes in different view angles between pre-event maps and post-event maps, and some are caused by undetected clouds, thin clouds, shadows, or other unknown errors.

Figure 24. The change-detection map of image between 11 April, 27 April and 1 May after removing the cloud and shadow (I is the change map of 11 April and 27 April, II is the change map of 11 April and 1 May, III is the change map of 27 April and 1 May. A–E region are the different sense and the detail show in Figure 25).
Figure 25. The detailed enlarged images of the red rectangle in Figure 24 (A and B are influenced by clouds and shadows; C is a collapsed temple; D and E are the open spaces where many tents are being set up).

Figure 26. The detected changes vector superposition on the cloudless map of 27 April. (the vector of the left hand map is extracted from the I image of Figure 24, and the vector of the right hand map has removed the error regions of the left hand vector map by visual checking).
Table 6. The simple statistic of detected changes vector map and checked changes vector map.

<table>
<thead>
<tr>
<th>Method</th>
<th>Number of Regions</th>
<th>Change Category</th>
<th>Number of Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total detected change</td>
<td>758</td>
<td>Damage Building (collapsed)</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tent/shelter</td>
<td>396</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cars or other temporary feature</td>
<td>139</td>
</tr>
<tr>
<td>Visual interpretation checking real change</td>
<td>556</td>
<td>changes as different view angles</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td></td>
<td>change as undetected thin cloud or shadow</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>other error detected changes</td>
<td>56</td>
</tr>
</tbody>
</table>

Figure 27. The temporal change detail map of two scenes (A shows a series of changes in Center Park from 27 April to 2 May and B shows the change of a collapsed tower).

Through the change detection map, the rescue progress can be observed. An example is shown in Figure 27. The image in Figure 27A shows the difference over four days in the open space of a park in the center of Kathmandu. The image from 27 April, the day after earthquake, indicates that the situation in this open space is in chaos, with a variety of temporary residences set up. However, by the sixth day after the earthquake, on the 1 May, the temporary settlements are orderly. This phenomenon is more obvious and more and unified residences and relief tents are set up by 2 May. The image in Figure 27B shows the series of changes of Bhimsen Tower. From 27 April to 2 May, the bricks and masonry from the ruins are cleaned up effectively and the situation changed from one of chaos to order. These temporal change maps indicate the progress of the disaster rescue, allowing for more objective guidance to formulate the next plan of rescue and allocate rescue resources reasonably.

4. Discussion

The goal of this work was to solve the problems related to the use of optical remote sensing data in disaster emergency mapping. The main problems are the impact of clouds in obtaining a clear situation backdrop map after the disaster and the contradiction between obtaining the poor precision change detection maps by automated methods and the relatively labor intensive process by visual interpretations. Through a series of new methods proposed in this paper, the accuracy of optical data
pre-processing results have been improved greatly and the produced change detection maps can fulfill the fundamental requirement of disaster S&R activities, which is the information regarding the disaster in terms of the number, location, distribution, arrangement, extent, and impact of the devastation. These rather coarse change maps also can be used as a guideline and the basis for visual observation to check the correctness of the detected change in objects. This can then generate more refined damage assessment maps and up-to-date maps for evaluation.

4.1. Map Accuracy and Timeliness

Reliability of the results and timeliness of the process are the key factors in emergency mapping. In previous studies, the biggest problem of optical data used in damage mapping by automatic change detection method was the calibration stage, which is a time consuming procedure with poor accuracy of results. In this paper, OptVM for image fusion and SIFT-OFM for image registration are proposed for producing a high accuracy fused image and matched image to obtain a rather high precision change detection map. Comparing the results based on traditional image fusion and registration methods, the results based on our proposed methods are the best overall (see the analysis in Sections 3.2 and 3.3). The false alarm rate of the final change detection scenes, which are caused by improper co-registration, can be reduced by relying on the rather high accuracy of the fused image and the matched image. Timeliness is also an issue. The time for processing the images of 11 April, 27 April, and 1 May from the Kathmandu example was only about 46 min to obtain the cloudless map shown in Figure 20 and the change map shown in Figure 24 using a PC with 8GB RAM and a 64-bit operating system. This is well worth the effort to obtain a cloudless situation backdrop map and a rather coarse change map. The time can be shortened if cluster or GPU parallel computing is used.

4.2. Limitations

Although a series of methods or strategies are used in this study to reduce the effects of optical imagery limitations, some limitations are still presented in the maps produced by this study. First, cloudless situation maps are available only when cloud cover is less than 50% of the image taken on the same day; otherwise, the cloudless maps contain a lot of information from the previous images. When the cloud coverage in the image is more than 80%, the image is completely unavailable, and only the SAR data is available to determine the situation of the cloud and shadow areas. The other limitation is that the generation of cloudless maps requires at least two post-event images with non-overlapping cloud covered areas. Post-event data are a bit more challenging due to cloud cover during the time of satellite passes. In some cases, getting a post-event cloudless map of the entire damaged area requires a great deal of time. However, multiple satellites passes and multiple data increase the possibility of obtaining cloudless images. Generally, a single post-event imagery is available in the first hours after the event. If it has cloud covered areas, there is no method other than to just replace the cloud areas with the nearest pre-event clear data when we want to get cloudless maps. Our proposed method can produce a cloudless map in this case. It is also useful for rescue teams to learn the situation of the affected areas using cloudless maps before they start search and rescue (S&R) activities within a few hours after the disaster. In parts of the country without internet access, the professional rescue teams need to take a cloudless printed map. It is important to provide cloudless maps even though cloud cover areas are replaced by pre-event clear data. In this kind cloudless map (seen Figure 22 left image), the vectors of detected cloud and shadow regions are provided to distinguish where is the pre-event or post-event data to avoid confusion for users.

Next, undetected clouds and shadows, variations in solar illumination, different off-nadir angles, and some temporary objects (e.g., car traffic), which could influence the results of change detection, still exist and reduce the accuracy of change maps in this example (as shown in Figure 23C,D). These errors are difficult to exclude from automatic change detected maps. It is only by manual visual checking that errors can be removed to obtain more precise change maps. That is why we emphasize the important of manual participation for further checking the change maps in this study. These change maps are just
the coarse result, which provides some basic main information for the affected regions. In addition, the detected changes in the product maps are not divided into categories and degrees, which also rely on manual interpretation to measure.

Then, a main limitation is the data resolution. As the policy of data acquisition and fund problems, the highest data resolution obtained was only 2 m with GF-1. This is not sufficient for detecting and assessing the damaged buildings, except those that are completely collapsed (shown in Figure 27B) or partially collapsed (shown in Figure 25C). The main purpose of the change detection maps in this study is to detect the scene after the earthquake to determine whether it was changed or not. The change detection maps could provide guidance of the changed regions for further checking as mentioned above.

Lastly, though image-based mapping has played an important role in rapid post-disaster damage assessment, the detail and accuracy of assessment has not reached the level of ground-based surveys [9]. The limitation of image-based damage assessment is not only related to the spatial resolution of the sensors, but also the reflected information reflected by image itself. Generally, most of the satellite images for disaster mapping are acquire by vertical angle (or small angle offset vertical), and the provided information of images, such as for building, is the situation of roofs. The roof information is well suited for the identification of extreme damage states, i.e., completely destroyed structures [61]. But for some not collapsed buildings with cracks or inclined walls, the results of image-based mapping are largely missed and lead to assessment deficiency of this type damage in the final results. Recently oblique or multi-perspective data, which provide information of both roofs and facades, has been identified as a potential data source for a more comprehensive building damage assessment [62,63].

5. Conclusions and Future Work

Using remote sensing for emergency mapping is now a common operational approach for disaster risk management. Both optical and radar data are the main input data when disasters occur. However, for damage assessment or analysis purposes, a number of results have been presented after every event mostly based on optical data. These can be easy to read and analyze to generate fast maps showing the location, scale, or the extent of disaster or crisis situation. In this study, a rapid procedure for transforming optical remote sensing data is proposed to produce cloudless situation backdrop maps and change detection maps for the emergency needs in the early phases of disaster recovery. On the one hand, the precision and reliability of the products has improved significantly by using the new methods, which include OptVM and SIFT-OFM. On the other hand, the automatic processing time is shortened by proposed procedure. It is useful to provide some rapid and actionable information for the governmental agencies, multilateral organizations, and NGOs in a timely fashion. While only the GF-1 temporal data were tested in this study as the limitation of data acquisition, the proposed procedure is also applicable to the other remote sensing data. It can even provide up-to-date situation awareness, loss estimation, and recovery monitoring maps to support decision making by processing multi-source and multi-temporal optical remote sensing data together.

The accuracy of the results from the automatic methods is obviously not as good as the manual visual interpretation by good cooperation of professional image analysts or skilled operators [63]. Manual visual interpretation, is always a rather labor intensive task. An ideal solution is to minimize the impact of restrictions to improve the accuracy of automatic methods results continuously. These automatic methods results can save the labor-cost to the maximum extent and provide some guideline for further refine the results by manual visual checking, though the results is rather coarse. This is also our main purposes of this study. International cooperation and volunteer contributions play an important role in mapping for disaster impact assessment at the urban level. International agencies and volunteers should be coordinated properly to ensure that common mapping guidelines and proper communication mechanisms are in place, for more effective cooperation among the involved actors especially during major emergencies.
Remote sensing is an effective tool at the disposal of a disaster manager. It can supply authentic and objective information for various rescue resources and the general public, especially in remote or inaccessible areas during the early period of the disaster. It would be strongest when allied with other tools, such as catastrophe risk models or GIS tools \[64\], to provide a powerful guidance to streamline evacuation, issue disaster declarations, coordinate S&R, coordinate logistics, conduct damage assessment, distribute relief supplies, and aid in long-term recovery. Some emerging technologies and initiatives will provide updated change and geospatial information as well as the resources distribution and demand data to analyze it in a time frame suitable for emergency response purposes. In the future, integrating different remote sensing data based on emergency mapping and other GIS information or analysis tools will be the popular trend to reduce economic losses, ensure public safety, and meet the basic subsistence needs of the people affected in an emergency response context.

Acknowledgments: This work was supported by the civil aerospace pre-research project during the “12th Five-Year Plan”, the satellite special project of National Development and Reform Commission: the system of remote sensing satellite data real-time active service and the 135 Strategy Planning of Institute of Remote Sensing and Digital Earth, CAS.

Author Contributions: Yong Ma developed the methods, carried out the experiments and analyzed the results. Fu Chen and Jianbo Liu supervised the research. Yang He, Jianbo Duan and Xinpeng Li prepared the data. Yong Ma and Fu Chen wrote the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix

The four indices used in this paper are defined as follows

1) **The Spectral Angle Mapper (SAM):** The SAM measures the absolute angle between the spectral vectors between the reference and fused images, which reflects the spectral distortion. The best value is 0, and the equation is shown as follow

\[
SAM(A(i,j), B(i,j)) = \arccos \left( \frac{\langle A(i,j), B(i,j) \rangle}{\| A(i,j) \|_2 \cdot \| B(i,j) \|_2} \right)
\]

where \( A(i,j) \) and \( B(i,j) \) are the value of pixel \((i, j)\) on the reference and fused images.

2) **The error of relative dimensionless global in synthesis (ERGAS):** ERGAS is a comprehensive index, which gives a global depiction of the radiometric difference between the reference and fused images. The best value is 0 and it is defined as

\[
ERGAS(A_k, B_k) = \frac{100}{\text{ratio}} \sqrt{\frac{1}{N} \sum_{k=1}^{N} \left( \frac{\text{RMSE}(A_k, B_k)^2}{\text{Mean}(B_k)} \right)}
\]

where \( A \) is the fused image and \( B \) is the reference image, ratio is the ratio of the spatial resolutions of the PAN image and the MS image, RMES is the root mean square of image \( A \) and \( B \), Mean represents the average value of the image.

3) **The Q4 Index:** the Q4 index \[34\] is an extension of the quality index, which is a product of the correlation coefficient, mean bias, and contrast variation. The best value is 1 and it is described as follow:

\[
Q4D \times D = \frac{\delta_{Z_1} \delta_{Z_2}}{\delta_{Z_1} \delta_{Z_2}} \cdot \frac{2 \delta_{Z_1} \cdot \delta_{Z_2}}{\delta_{Z_1} \cdot \delta_{Z_2} + \delta_{Z_1} \cdot \delta_{Z_2}} \cdot \frac{2 |Z_1| \cdot |Z_2|}{|Z_1|^2 + |Z_2|^2}
\]

\[
Z_1 = X_1 + i \times X_2 + j \times X_3 + k \times X_4
\]

\[
Z_2 = Y_1 + i \times Y_2 + j \times Y_3 + k \times Y_4
\]

\[
Z_1^* = X_1 - i \times X_2 - j \times X_3 - k \times X_4
\]
\[ |Z_1| = \sqrt{Z_1 \times Z_1^*} = \sqrt{(X_1^1)^2 + (X_1^2)^2 + (X_1^3)^2 + (X_1^4)^2} \]

where \( |Z_2| \) is the same way with \( |Z_1| \), \( X_1 \) and \( Y_1 \) are the first band of image \( A \) and \( B \) respectively, and so are the other bands. \( \sigma_{Z_1Z_2} \) is the covariance of \( Z_1 \) and \( Z_2 \), and \( \sigma_{Z_1} \) and \( \sigma_{Z_2} \) is the variance of \( Z_1 \) and \( Z_2 \). The first term of the equation is the modulus of the hyper-complex CC between \( Z_1 \) and \( Z_2 \), and the second and third terms measure contrast changes and the mean bias on all bands, respectively. \( D \) is block size of image and in this paper the size is 32*32.

4) Quality with No Reference (QNR): QNR measure the quality of fused image and reference image from the spectral and spatial distortion. The best value is 1 and it is defined as follow:

\[
D_\lambda = \frac{1}{L(L-1)} \sum_{l=1}^{L} \sum_{r=1}^{L} \left| Q \left( \hat{G}_l, \hat{G}_r \right) - Q \left( \tilde{G}_l, \tilde{G}_r \right) \right|^{p}
\]

\[
D_s = \frac{1}{L(L-1)} \sum_{l=1}^{L} \left| Q \left( \hat{G}_l, P \right) - Q \left( \tilde{G}_l, \tilde{P} \right) \right|^{q}
\]

where \( L \) is the number of bands, \( \hat{G} \) is fused HRM image, and \( \tilde{G} \) is LRM image. \( p \) is integer index of amplifying the spectrum distortion, generally taking \( p = 1 \). \( P \) is HRP image and \( \tilde{P} \) is panchromatic image after processing by low-pass filter. \( q \) is integer index of amplifying of spatial distortion, generally taking \( q = 1 \), and \( \alpha = \beta = 1 \).

References


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
Crowdsourcing Rapid Assessment of Collapsed Buildings Early after the Earthquake Based on Aerial Remote Sensing Image: A Case Study of Yushu Earthquake

Shuai Xie1,2, Jianbo Duan1,*, Shibin Liu1, Qin Dai1, Wei Liu1, Yong Ma1, Rui Guo1 and Caihong Ma1

1 Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China; xieshuai@radi.ac.cn (S.X.); liusb@radi.ac.cn (S.L.); daqing@radi.ac.cn (Q.D.); liuwei@radi.ac.cn (W.L.); mayong@radi.ac.cn (Y.M.); guorui@radi.ac.cn (R.G.); mach@radi.ac.cn (C.M.)

2 University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: duanjb@radi.ac.cn; Tel.: +86-10-8217-8151

Academic Editors: Roberto Tomás, Zhenhong Li, Gonzalo Pajares and Prasad S. Thenkabail

Received: 14 July 2016; Accepted: 9 September 2016; Published: 14 September 2016

Abstract: Remote sensing (RS) images play a significant role in disaster emergency response. Web2.0 changes the way data are created, making it possible for the public to participate in scientific issues. In this paper, an experiment is designed to evaluate the reliability of crowdsourcing buildings collapse assessment in the early time after an earthquake based on aerial remote sensing image. The procedure of RS data pre-processing and crowdsourcing data collection is presented. A probabilistic model including maximum likelihood estimation (MLE), Bayes’ theorem and expectation-maximization (EM) algorithm are applied to quantitatively estimate the individual error-rate and “ground truth” according to multiple participants’ assessment results. An experimental area of Yushu earthquake is provided to present the results contributed by participants. Following the results, some discussion is provided regarding accuracy and variation among participants. The features of buildings labeled as the same damage type are found highly consistent. This suggests that the building damage assessment contributed by crowdsourcing can be treated as reliable samples. This study shows potential for a rapid building collapse assessment through crowdsourcing and quantitatively inferring “ground truth” according to crowdsourcing data in the early time after the earthquake based on aerial remote sensing image.

Keywords: crowdsourcing; building collapse assessment; earthquake; aerial image; EM algorithm

1. Introduction

Building collapse is one of the most serious types of earthquake damage. Most casualties from earthquakes are associated with collapsing buildings [1]. The extent of buildings damage reflects seismic intensity, which is important information to assess the losses of life and property in an earthquake-hit area [2]. Rapid assessment of collapsed buildings early after the earthquake can be instrumental in search and rescue during an emergency. It is hard to obtain the whole in-situ information of building damage in a short time after the earthquake, because the earthquake-damaged zones are not accessible in most cases. However, the aerial or satellite remote sensing can provide the image of the whole disaster area, making it possible to estimate the building damage of large disaster-affected regions in the early time. Many methods to visually interpret or automatically extract the building damage after the earthquake were proposed based on high-resolution aerial or satellite remote sensing image over the past ten years, which made a great contribution in estimating damage.
Remote Sens. 2016, 8, 759

extent of buildings caused by earthquake using remote sensing data. Chiroiu uses post-earthquake Ikonos imagery to assess the collapsed buildings by visual interpretation [3]. Saito et al. demonstrated that more collapsed buildings were recognized using pre- and post-earthquake QuickBird imagery, because the pre-earthquake imagery is a good reference of the building outlines [4]. Vu et al. uses the region-independent edge detection algorithm to detect the collapsed buildings based on Ikonos imagery. The results of “very heavy damage” and “destroyed” are consistent with visual interpretation and site survey [5]. Huyck et al. uses texture change detection algorithm based on pre- and post-earthquake imagery of mono-sensor and multi-sensor, respectively, finding that the results are quite different and only “hardest hit area” is recognized consistently [6]. Hutchinson et al. firstly extracts building boundary based on pre- and post-earthquake satellite imagery, then calculated the boundary compactness index defined as the ratio of the number of boundary pixels in the post- and pre-earthquake house, finally identifies the damage buildings through a threshold [7]. Chen and Hutchinson proposed a probabilistic classification framework by means of a multiclass classifier based on bitemporal satellite images to address the major limitation in past attempts which is the use of deterministic approaches to classify damage levels [8]. Geiß et al. quantitatively evaluates the suitability of multi-sensor remote sensing to assess the seismic vulnerability of buildings showing potential for a rapid screening assessment of large areas [1].

Due to the complex image characteristics of post-disaster ground objects and the limitation of resources, automated damage detection techniques with remote sensing are still in the preliminary stage [9]. The last decade has seen a proliferation of sophisticated sensors and technology capable of capturing, transmitting and storing immense amounts of data, like remote sensing image, increasing the importance and demand for fast and reliable methods of analysis [9]. Web2.0 technology has resulted in changes in the way that data are created. Individuals, who have the characteristics of large quantity, flexible time and uncertain location, now provide vast amounts of information to websites and online databases, much of which is spatially referenced [10]. This phenomenon called “crowdsourcing” is the product of the network society, which is an online and distributed pattern of problem-solving and producing. Therefore, remote sensing combined with crowdsourcing was used to quickly and accurately analyze large data sets by creating and leveraging a distributed network of human analysts. Crowdsourcing geographic information for disaster response has become a research frontier [11]. The Virtual Disaster Viewer, which was a pilot project following the May 2008 Wenchuan, China earthquake, provided damage assessments through crowdsourcing by having experts interpret pre- and post-event satellite imagery [9]. After the 2010 earthquake in Haiti, the GEO-CAN initiative utilized crowdsourcing through the recruitment of experts to make critical damage assessments based on high-resolution post-event satellite and aerial imagery [12]. Building upon the GEO-CAN effort in Haiti, damage assessment after the Christchurch 2011 was improved by asking participants, including non-experts, to delineate damaged buildings use a polygonal tool, in order to making crowdsourcing damage assessments of disaster areas faster and more accurate [13]. The above studies have clearly demonstrated the power of crowdsourcing for damage assessment to improve disaster response. As such, it offers substantial advantages, but suffers from a general lack of quality assurance [14]. The participants with different professional background and knowledge level have different understanding of remote sensing image. The interpretation results of the same image may be different. It is necessary to quantitatively evaluate the quality of crowdsourcing data to ensure the accuracy of damage assessment.

In this paper, Yushu earthquake is chosen as a case study. A web-based platform is built to collect the post-earthquake building damage assessment results contributed by public participants. High-resolution aerial remote sensing images are used due to the advantage of being captured and processed faster and higher spatial resolution than satellite imagery [15]. The problem we focus on is damage “extent” identification for buildings which is relatively straightforward and fast by means of RS images, instead of damage “level” assessment which is based on ground evaluations requiring a considerable amount of time and effort [16]. The next section describes the details of
study area. The data used for this research and overview of processing are introduced in the third section. The probabilistic model is applied to estimate individuals’ error-rates and infer ground truth. The experiment results are presented in the fourth section. Following the results, some discussion is provided regarding the accuracy of EM algorithm compared to the “majority” method and the variation distribution of assessment results among participants on each building. In addition, the features of each crowdsourcing-derived damage type are analyzed, which can be regarded as reliable samples to train machine learning to recognize objects of interest. Our main conclusions are presented in the final section and future work is proposed.

2. Materials and Methods

2.1. Study Area and Data

On 14 April 2010, a 7.1 magnitude earthquake occurred near Yushu, China, at 7:49 a.m. local time. The epicenter was located at 33.1 degrees north latitude, 96.6 degrees east longitude and focal depth was 14 km. The terrain is mainly mountainous, with an average elevation of 4493 m. The earthquake caused a large number of casualties and collapsed houses. The site survey data of Chinese scholars after the earthquake demonstrate that houses in the central area of Jiegu Town are mainly with brick and concrete structure, while other houses are mainly with brick and civil structure. Some of the brick and concrete structures in the town center area suffered serious damage, and almost all the brick and civil structures in the western and southern regions of the town were totally damaged [17]. Figure 1 below shows the location of Jiegu Town, where the earthquake happened.

![Maps about administrative areas and terrains](image-url)

*Figure 1. Maps about administrative areas and terrains: (a) the geographical location of Qinghai Province, China; (b) the geographical location of Yushu County in Qinghai Province; (c) the geographical location of Jiegu Town in Yushu County; and (d) the topographic map of Jiegu Town made by Landsat 8 OLI image (true color).*

The data used for this application are a 0.4 m resolution multispectral aerial image of the damaged zones of Jiegu Town captured on 14 April 2010, provided by Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences (Figure 2).
Figure 2. The aerial image of Jiegu Town captured on 14 April 2010, which was the area severely damaged by the earthquake.

2.2. Methods

2.2.1. Architecture of Processing

The architecture consists of four parts, which are, in order, basic imagery preparation, damage assessment collection, data quality evaluation and damage map export (Figure 3).

(1) Basic imagery preparation

In this part, we take the post-event high resolution aerial image (Section 2.1 mentioned) as basic imagery, which crowdsourcing participants use to visually interpret collapsed buildings. We obtain two images that cover the different part of the study area, respectively, but having the overlapping zone. To generate the entire region image, image registration, mosaic and clipping are applied to the two images we obtain.

(2) Damage assessment collection

We publish the pre-processed image online that is accessible to participants. Each visible building in the image is assigned one of the damage types according to the observed damage.

(3) Data quality evaluation

Because of the different professional background of participants, the quality of the data provided by the individual is uneven. The term crowdsourcing has three distinct meanings proposed by Goodchild [14]. The first meaning refers to the solution of a problem by referring it to a number of people, without respect to their qualifications. The second meaning refers to the ability of a group to validate and correct the errors that an individual might make. The third interpretation refers to the ability of the crowd to converge on the truth. We could not get the field data in a short time after the earthquake. However, we could infer ground truth from subjective labeling of the post-event high-resolution aerial image by participants. In the next section, we introduce the details of how we use maximum likelihood estimation, Bayes’ theorem and EM algorithm to estimate the ground truth and individual error rate.

(4) Damage map export

In the last part, the results integrated by the previous part are visualized, generating a damage map that shows the spatial distribution of damage types. The damage distribution map is generated with individual buildings rendered with colors representing the type of damage.
The process above enables assessing the damage of buildings rapidly and flexibly through crowdsourcing without considering the constraint conditions when using algorithm to process image. Each participant is regarded as a “classifier”, which classifies post-disaster buildings into three categories according to human’s understanding of the image, and then we evaluate the accuracy of each person and integrate the multiple people’s results into a final reliable result.

2.2.2. Probabilistic Model

Note that there are \( W \) participants to assess \( I \) buildings, which may be damaged using \( K \) damage types. It is assumed that all responses given by a single participant are independent and all the participants interpret independently. In addition, a participant may interpret the same building more than once. Note that \( a_{kl}^{(w)} (k = 1, \ldots, K; l = 1, \ldots, K; w = 1, \ldots, W) \) are the probability that a participant \( w \) will label \( l \) given \( k \) is the true type, which are called the individual error rate. \( n_{il}^{(w)} (i = 1, \ldots, I; l = 1, \ldots, K; w = 1, \ldots, W) \) are the number of times participant \( w \) label building \( i \) as \( l \), and \( p_k (k = 1, \ldots, K) \) are the probability that the true damage type of building is \( k \). Let \( G_{ik} (k = 1, \ldots, K) \) be a binary variable of building \( i \). If \( t \) is the true damage type of building \( i \), then \( G_{it} = 1 \) and \( G_{ik} = 0 \) \((k \neq t)\), namely, \( p(G_{ik} = 1) = p_k \). We follow a general model for subjective labeling originally proposed by Dawid and Skene [18] and apply it to the building damage labeling problem. The data from all participants are assumed to be independent and all the true damage types of buildings are assumed to be available. Generally, the likelihood function for the full data is

\[
\prod_{i=1}^{I} \prod_{k=1}^{K} p_k \prod_{w=1}^{W} \prod_{l=1}^{K} \left( a_{kl}^{(w)} \right)^{n_{il}^{(w)}} G_{ik} G_{il} G_{ikl} \quad (1)
\]

Using maximum likelihood estimation, and we obtain estimators

\[
\hat{k}_{kl}^{(w)} = \frac{\sum_i G_{ik} H_{il}^{(w)}}{\sum_l \sum_i G_{ik} n_{il}^{(w)}},
\]

\[
\hat{a}_{kl}^{(w)} = \frac{\sum_{i} G_{ik} n_{il}^{(w)}}{\sum_{i} G_{ik} H_{il}^{(w)}}.
\]
When $p_k (k = 1, \ldots, K)$ are unknown, these can be estimated:

$$\hat{p}_k = \frac{\sum G_{ik}}{I}.$$  \hspace{1cm} (3)

At this point, the true damage types of buildings are unknown. We use Bayesian theory to estimate the binary variable $G_{ik} (k = 1, \ldots, K)$,

$$p(G_{ik} = 1 | \text{data}) = \frac{p(\text{data} | G_{ik} = 1) p(G_{ik} = 1)}{\sum_{i=1}^{K} p(\text{data} | G_{ii} = 1) p(G_{ii} = 1)}.$$  \hspace{1cm} (4)

Therefore,

$$p(G_{ik} = 1 | \text{data}) = \frac{\prod_{w=1}^{W} \prod_{l=1}^{K} (a_{kl}^{(w)})^{n_{il}^{(w)}} p_k}{\sum_{i=1}^{K} \prod_{w=1}^{W} \prod_{l=1}^{K} (a_{il}^{(w)})^{n_{il}^{(w)}} p_l}.$$  \hspace{1cm} (5)

Then, we use EM algorithm for finding maximum likelihood estimates of parameters in the model above, due to the dependency of the hidden variables $G_{ik}$. EM algorithm is short for Expectation Maximization algorithm, which was described by Dempster et al. in 1977 [19]. It is an iterative optimization method for maximum likelihood estimation of parameters, which can estimate the parameters from incomplete data set.

In this problem, we treat $G_{ik}$ as missing data then the conditions of the EM algorithm are satisfied. The iterative procedure is as follows:

1. Give initial estimates of the $G$s.
2. Use Equations (2) and (3) to obtain estimates of the $p$s and $\alpha$s.
3. Use Equation (5) and the estimates of the $p$s and $\alpha$s to calculate new estimates of the $G$s.
4. Repeat Steps 2 and 3 until the results converge.

In Step 1, we use the equation below to calculate initial estimates of $G$s,

$$\hat{G}_{ik} = \frac{\sum_{w} n_{ik}^{(w)}}{\sum_{w} \sum_{l} n_{il}^{(w)}}.$$  \hspace{1cm} (6)

3. Experiment Results

Our project asked the participants to classify the post-earthquake damage buildings into one of three damage types: (1) basically intact; (2) partially collapsed; and (3) completely collapsed. These type numbers are used in subsequent tables. Here, the Yushu earthquake case study was selected to illustrate the results. The experiment area was a sub-region of Jiegu Town, which had visible various types of damage extent, shown in Figure 4, and contained 3456 data points labeled by 27 volunteers, describing the damage buildings at 127 locations. As can be seen in Figure 4, “basically intact”, “partially collapsed” and “completely collapsed” are represented by green, yellow and red points, respectively.

The system consists of a database for damage assessment accessed through a browser-based interface built using the ArcGIS API. Data from the participants are collected in the browser and transferred into the database through AJAX and PHP. One selects a damage type and then draws a point on the corresponding building. When a participant assesses a building’s damage, the record is stored along with the longitude and latitude of the point he draw, the damage type he select and his id. The data we collected were used later to estimate the error-rate of each participant in identifying the damage extent of each building and infer the ground truth.

Figure 5 gives the variation tendency of marginal probabilities of the three damage types with the iteration of EM algorithm. As shown in Figure 5, the results converge when the iteration is...
12 times. Table A1 (see Appendix A) gives the estimates of the individual error-rates ($\alpha$) of the 27 participants. The diagonal elements of each matrix are the estimate of the probability of a correct allocation by a participant. Table A2 (see Appendix A) gives the estimated probabilities for the Gs for each building. For most buildings, the posterior probability is 1 for one damage type, and the consensus appears obvious.

**Figure 4.** The experiment area and the location of the 3456 data points, which were contributed by 27 participants online. In the map, the green, yellow and red points indicate the damage type of “basically intact”, “partially collapsed” and “completely collapsed”, respectively.

**Figure 5.** The red, blue and green lines demonstrate the fluctuation of marginal probability of “basically intact”, “partially collapsed” and “completely collapsed”, respectively, with the increase of the number of iterations. When the number of iterations reaches 12, EM algorithm converges.
4. Discussion

Of the 3456 damage assessments received for the experiment region, we find that “basically intact” annotations made up 52.14%, “partially collapsed” made up 34.64% and “completely collapsed” made up 13.22%. There is no clear bias towards one or two damage types. However, if using the EMS-98 scale, the distribution of annotations reveals an overall bias to assess a building as “No Damage” or “Destroyed” [13]. In order to demonstrate the advantage of EM algorithm in terms of inferring ground true, we make a comparison with “majority” method. Figure 6a,b shows the assessment results of EM algorithm and “majority” method, respectively.

![Comparison of the results between the two different methods: (a) the result of EM algorithm; and (b) the result of “majority” method, which are generated with individual buildings rendered with colors representing the type of damage. The rectangles represent the buildings, and green, yellow and red indicate “basically intact”, “partially collapsed” and “completely collapsed”, respectively. The rectangles with the id number on them have the different results between the two methods.](image)

As shown in Figure 6, there are 11 buildings that have different results between the two methods: 21, 24, 29, 33, 61, 66, 67, 75, 78, 105 and 107. The first nine buildings are completely collapsed in EM results while are partially collapsed in “majority” method. The No. 105 building is basically intact in EM results while is partially collapsed in “majority” method. The situation of No. 107 building is opposite to No. 105. The “majority” method does not take into account the accuracy
of the participants, and EM may not choose the majority damage types as the final result of one building due to the low accuracy of participants who make the assessment. For example, the No. 33 building received the assessment results of 26 participants. Among them, there are 19 participants who gave the “partially collapsed” result and six participants who assessed the No. 33 building as “completely collapsed”. Consequently, the “majority” method regards the No. 33 building as “partially collapsed”. According to Table A1, we calculate the average accuracy of 19 participants who label as “partially collapsed” when the true is partially collapsed and the average accuracy of six participants who label as “completely collapsed” when the true is completely collapsed. The calculation results are shown in Table 1. We also calculated the corresponding incidence defined as the product of individual accuracy and marginal probability of damage type, as seen in Table 2. Obviously, the average accuracy of the latter is larger than the former and so is the average incidence, indicating that the participants who label the No. 33 building as “completely collapsed” have more “weight”.

Table 1. The accuracy of 19 participants who label No. 33 building as “partially collapsed” and 6 participants who label No. 33 building as “completely collapsed”. The averages are calculated in the last line of the table.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Accuracy of 2</th>
<th>Accuracy of 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5021</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.2699</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.3825</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.6200</td>
</tr>
<tr>
<td>5</td>
<td>0.6806</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.7419</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.5761</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>0.8492</td>
</tr>
<tr>
<td>10</td>
<td>0.6132</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.8465</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.7924</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>0.6200</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>0.5246</td>
</tr>
<tr>
<td>15</td>
<td>0.6887</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.5912</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.6015</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.5011</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.7608</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>0.7306</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td>0.7602</td>
</tr>
<tr>
<td>23</td>
<td>0.1905</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>0.3056</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>0.9523</td>
</tr>
<tr>
<td>26</td>
<td>0.2712</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>0.5399</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.5570</td>
<td>0.7211</td>
</tr>
</tbody>
</table>

To demonstrate the variation distribution of assessment results that participants give on each building, the percentage of each damage type on each building is presented in Figure 7, in which the building ids are sorted by the percentage of “basically intact” from small to large. Besides, the standard deviation of participants’ assessment results on each building is shown in Figure 8. No clear agreement between participants on each building apart from the No. 99 that all participants label as “basically intact”. This is because participants with different professional background have different cognition, or due to the limitation of image inherent characteristics such as spatial resolution and angle of view. Whatever limitations the professional faces naturally also apply to volunteers [20]. However, no obvious bias towards extreme value indicates that a majority of participants worked without malice.
Table 2. The incidence of 19 participants who label No. 33 building as “partially collapsed” and 6 participants who label No. 33 building as “completely collapsed”. The averages are calculated in the last line of the table.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Incidence of 2</th>
<th>Incidence of 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1030</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0553</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.0764</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.1024</td>
</tr>
<tr>
<td>5</td>
<td>0.1395</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.1521</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.1181</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>0.1402</td>
</tr>
<tr>
<td>10</td>
<td>0.1257</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.1736</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.1625</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>0.1024</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>0.0866</td>
</tr>
<tr>
<td>15</td>
<td>0.1412</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.1212</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.1233</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.1027</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.1560</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>0.1498</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td>0.1255</td>
</tr>
<tr>
<td>23</td>
<td>0.0391</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>0.0627</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>0.1572</td>
</tr>
<tr>
<td>26</td>
<td>0.0556</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>0.1099</td>
<td></td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.1142</strong></td>
<td><strong>0.1190</strong></td>
</tr>
</tbody>
</table>

In remote sensing applications, “ground truth” data are often used as the basis for training pattern recognition algorithms to detect objects of interest [21]. Many semiautomatic techniques have been designed to exploit Earth Observation (EO) data for earthquake damage assessment to the maximum possible extent, yet visual inspection still remains the best way to achieve meaningful results [16]. Combined with multiple participants, probability model-derived final results are used as the reliable samples to analyze the features of each damage type.

![Figure 7](image)

**Figure 7.** The percentage of each damage type on each building. The green, yellow and red bars represent the percentage of “basically intact”, “partially collapsed” and “completely collapsed”, respectively, on each building. The building ids depicted by X-axis are in ascending order according to the percentage of “basically intact” on each building.
Figure 8. The standard deviation of participants’ assessment results on each building. The building ids depicted by X-axis are also in ascending order according to the percentage of “basically intact” on each building.

We select some samples of each damage type from the results of EM algorithm, as shown in Table 3. The number 1, 2 and 3 represent the damage type of basically intact, partially collapsed and completely collapsed, respectively. The common features of damage type 1 have a clear outline and regular shape, and an intact shadow. Damage type 2 has fuzzy boundaries, irregular shape, offset orientation, and loss of shadow effects. Damage type 3 has no visible characteristics of man-made objects.

Some limitations of our experiment are discussed below.

Table 3. The crowdsourcing-derived samples of each damage types.

<table>
<thead>
<tr>
<th>Damage Types</th>
<th>Sample 1</th>
<th>Sample 2</th>
<th>Sample 3</th>
<th>Sample 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1" alt="Sample 1" /></td>
<td><img src="image2" alt="Sample 2" /></td>
<td><img src="image3" alt="Sample 3" /></td>
<td><img src="image4" alt="Sample 4" /></td>
</tr>
<tr>
<td>2</td>
<td><img src="image1" alt="Sample 1" /></td>
<td><img src="image2" alt="Sample 2" /></td>
<td><img src="image3" alt="Sample 3" /></td>
<td><img src="image4" alt="Sample 4" /></td>
</tr>
<tr>
<td>3</td>
<td><img src="image1" alt="Sample 1" /></td>
<td><img src="image2" alt="Sample 2" /></td>
<td><img src="image3" alt="Sample 3" /></td>
<td><img src="image4" alt="Sample 4" /></td>
</tr>
</tbody>
</table>

The study area in this paper is one part of the whole damage area, which we used as an experimental area to apply our methods on it and present the results clearly. In addition, the approach is also applicable in the wide area, which is the most important purpose of the crowdsourcing damage assessment. Because the model is proposed for the general circumstance without the limits of area, we will extend the application to wide interpretation in future work. The raw data we collected show that most participants interpret only once, although the method allows participants to interpret multiple times. We could not determine the minimum number of the participants, because the web-based interface is open to the public and any case could appear. We could calculate a result based on a span from the start to the time we choose, for example a week. The method would give a result based on any data collection phase.

Although “actual” ground truths of building damage are necessary in order to truly discuss the applicability of the proposed method, we could not find ground photographs or broadcasted videos on
TV of the damaged buildings in the target area, only text information or papers are available. We focus on using the wisdom of the public to find more damaged buildings in the early stage of the earthquake. The assessment results of the target area in this paper are highly consistent with results of the existing study published by Dou et al. [17] in the corresponding area.

5. Conclusions and Future Work

Building damage assessment in the early time after an earthquake is a very crucial problem. Knowing where the collapsed buildings are and to what extent buildings have been damaged are closely related to life-saving for emergency response. However, it is hard to survey the whole in-situ information in a short time after an earthquake. Satellite or aerial remote sensing technology has the capability of earth observation, and becomes a useful tool for damage estimation without being physically present in disaster area. Aerial remote sensing images, which are captured and processed faster and have higher spatial resolution, make it possible to rapidly assess collapsed buildings early after the earthquake. A web-based interface was built. Anyone who accessed our website was required to assign one of the three damage types for each buildings based on the aerial image. The 3456 data points from 27 participants on the experimental area, which is a sub-region of Yushu, were collected. MLE, Bayes’ theorem and EM algorithm were applied to estimate the individual error-rates and infer “ground truth” according to the 3456 data points of 27 participants. The results suggest that EM algorithm is better than “majority” method and there is no clear bias towards extreme value among damage types contributed by participants. This study shows that the variation of image understanding among participants exists, due to their different professional background. We demonstrated how to collect and store the data created by individuals online, how to make them contribute their results flexibly and easily, drawing a point instead of a polygon, and how to quantitatively estimate individuals’ accuracy and the “ground truth” of each building, using a probabilistic model. By means of a sequential procedure of RS image pre-processing, publishing RS image online, collecting crowdsourcing building damage assessment data, evaluating the quality of data contributed by crowdsourcing and damage mapping, the building collapse can be rapidly assessed with viable accuracies in the early time after the earthquake. We conclude that RS data combined with crowdsourcing have a high capability to support large-area assessments of building collapse, meeting the need of disaster emergency response. A new processing framework is proposed to establish the connection between remote sensing image and crowdsourcing, demonstrating potential for crowdsourcing rapid assessment of building collapse early after the earthquake based on aerial remote sensing image. Future efforts will focus on providing multi-source and multi-temporal remote sensing image. Multi-source RS data, such as oblique images, can provide more views of buildings on the ground. Multi-temporal RS data, such as pre-earthquake images, are considered to be very useful as a reference in identifying the damage in the post-event image [4]. Although these adjustments will refine the final results towards rapid damage assessment, a problem should be considered: how to balance the operational complexity of system and the improvement of results, because ordinary people are more inclined to use easy-to-operate systems without spending too much time. In summary, the aim is that the rapid assessment results through crowdsourcing could meet the needs of deploying rescue forces in the early time after the earthquake. More details about the damage level of buildings surveyed on the spot afterwards are beyond the scope of this paper.

Acknowledgments: The paper is funded by “Research on the model of remote sensing disaster monitoring and assessment based on crowdsourcing” project, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences. And we thank Airborne Remote Sensing Center, Chinese Academy of Sciences a lot for providing post-earthquake aerial remote sensing image of Yushu.

Author Contributions: Jianbo Duan conceived and designed the experiments; Jianbo Duan, Rui Guo and Caihong Ma built the experimental platform; Shuai Xie prepared and processed the aerial data, performed the experiments, analyzed the data and wrote the paper; Shibin Liu supervised the research; and Qin Dai, Wei Liu and Yong Ma gave comments and revised the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.
Appendix A

Table A1. The estimates of the individual error-rates.

<table>
<thead>
<tr>
<th>Participant 1</th>
<th>Participant 15</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td><strong>1</strong></td>
<td>0.64</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>0.38</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant 2</th>
<th>Participant 16</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td><strong>1</strong></td>
<td>0.83</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>0.5</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant 3</th>
<th>Participant 17</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td><strong>1</strong></td>
<td>0.9</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>0.62</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>0.24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant 4</th>
<th>Participant 18</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td><strong>1</strong></td>
<td>0.66</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>0.23</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>0.09</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant 5</th>
<th>Participant 19</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td><strong>1</strong></td>
<td>0.94</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>0.32</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>0.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant 6</th>
<th>Participant 20</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td><strong>1</strong></td>
<td>0.46</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>0.12</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant 7</th>
<th>Participant 21</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td><strong>1</strong></td>
<td>0.47</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>0.46</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>0.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant 8</th>
<th>Participant 22</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td><strong>1</strong></td>
<td>0.9</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>0.42</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>0.08</td>
</tr>
</tbody>
</table>
### Table A1. Cont.

<table>
<thead>
<tr>
<th>Participant 9</th>
<th>Participant 23</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td>True 1 2 3</td>
<td>True 1 2 3</td>
</tr>
<tr>
<td>1 0.26 0.56 0.17</td>
<td>1 1 0.81 0.19</td>
</tr>
<tr>
<td>2 0.04 0.4 0.56</td>
<td>2 0.81 0.19 0</td>
</tr>
<tr>
<td>3 0 0.15 0.85</td>
<td>3 0.43 0.43 0.14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant 10</th>
<th>Participant 24</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td>True 1 2 3</td>
<td>True 1 2 3</td>
</tr>
<tr>
<td>1 0.73 0.27 0</td>
<td>1 0.94 0.06 0</td>
</tr>
<tr>
<td>2 0.08 0.61 0.31</td>
<td>2 0.69 0.31 0</td>
</tr>
<tr>
<td>3 0 0.3 0.7</td>
<td>3 0.26 0.48 0.26</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant 11</th>
<th>Participant 25</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td>True 1 2 3</td>
<td>True 1 2 3</td>
</tr>
<tr>
<td>1 0.36 0.64 0</td>
<td>1 0.98 0.03 0</td>
</tr>
<tr>
<td>2 0 0.79 0.21</td>
<td>2 0.69 0.27 0.04</td>
</tr>
<tr>
<td>3 0 0.73 0.27</td>
<td>3 0 0.05 0.95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant 12</th>
<th>Participant 26</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td>True 1 2 3</td>
<td>True 1 2 3</td>
</tr>
<tr>
<td>1 0.77 0.18 0.05</td>
<td>1 0.76 0.19 0.05</td>
</tr>
<tr>
<td>2 0.23 0.5 0.27</td>
<td>2 0.35 0.54 0.12</td>
</tr>
<tr>
<td>3 0 0.38 0.62</td>
<td>3 0.05 0.38 0.57</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant 13</th>
<th>Participant 27</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td>True 1 2 3</td>
<td>True 1 2 3</td>
</tr>
<tr>
<td>1 0.67 0.33 0</td>
<td>1 0.01 0.99 0</td>
</tr>
<tr>
<td>2 0.09 0.91 0</td>
<td>2 0.99 0.01 0</td>
</tr>
<tr>
<td>3 0 0.43 0.52</td>
<td>3 0 0 0.99</td>
</tr>
</tbody>
</table>

| Participant 14 | | |
|---------------|----------------|
| **Observed**  | **Observed**   |
| True 1 2 3    | True 1 2 3    |
| 1 0 1 0       | 1 0 1 0       |
| 2 0.09 0.91 0 | 2 0.99 0.01 0 |
| 3 0 0.43 0.52 | 3 0 0 0.99    |

### Table A2. The estimated probabilities for the Gs.

<table>
<thead>
<tr>
<th>Building ID</th>
<th>Damage Types</th>
<th>Building ID</th>
<th>Damage Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 0 0</td>
<td>1</td>
<td>1 0 0</td>
</tr>
<tr>
<td>2</td>
<td>0.002 0.998</td>
<td>2</td>
<td>0.024 0.976</td>
</tr>
<tr>
<td>3</td>
<td>0 0 1</td>
<td>3</td>
<td>0 0 1</td>
</tr>
<tr>
<td>4</td>
<td>1 0 0</td>
<td>4</td>
<td>0.003 0.997</td>
</tr>
<tr>
<td>5</td>
<td>1 0 0</td>
<td>5</td>
<td>0 0 1</td>
</tr>
<tr>
<td>6</td>
<td>0.006 0.994</td>
<td>6</td>
<td>0.956 0.044</td>
</tr>
<tr>
<td>7</td>
<td>1 0 0</td>
<td>7</td>
<td>1 0 0</td>
</tr>
<tr>
<td>8</td>
<td>0.996 0.004</td>
<td>8</td>
<td>0.999 0.001</td>
</tr>
<tr>
<td>9</td>
<td>0 0 1</td>
<td>9</td>
<td>0 0 1</td>
</tr>
<tr>
<td>10</td>
<td>1 0 0</td>
<td>10</td>
<td>0 1 0</td>
</tr>
<tr>
<td>11</td>
<td>1 0 0</td>
<td>11</td>
<td>0 1 0</td>
</tr>
<tr>
<td>12</td>
<td>1 0 0</td>
<td>12</td>
<td>0 0 1</td>
</tr>
<tr>
<td>13</td>
<td>0.01 0.99</td>
<td>13</td>
<td>0 0 1</td>
</tr>
</tbody>
</table>

434
Table A2. Cont.

<table>
<thead>
<tr>
<th>Building ID</th>
<th>Damage Types</th>
<th>Building ID</th>
<th>Damage Types</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>0.072</td>
<td>0.928</td>
</tr>
<tr>
<td>26</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>27</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>28</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>29</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>31</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>33</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>34</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>35</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>36</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>0</td>
<td>0.998</td>
<td>0.002</td>
</tr>
<tr>
<td>38</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>39</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>41</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>42</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>43</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>44</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>46</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>47</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>48</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>49</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>51</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>52</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>53</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>54</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>55</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>56</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>57</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>58</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>59</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>60</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>61</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>62</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>63</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>64</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

References


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).
Detection of Urban Damage Using Remote Sensing and Machine Learning Algorithms: Revisiting the 2010 Haiti Earthquake

Austin J. Cooner *, Yang Shao and James B. Campbell

Remote Sens. 2016, 8, 868
www.mdpi.com/journal/remotesensing

Article

Abstract: Remote sensing continues to be an invaluable tool in earthquake damage assessments and emergency response. This study evaluates the effectiveness of multilayer feedforward neural networks, radial basis neural networks, and Random Forests in detecting earthquake damage caused by the 2010 Port-au-Prince, Haiti 7.0 moment magnitude (Mw) event. Additionally, textural and structural features including entropy, dissimilarity, Laplacian of Gaussian, and rectangular fit are investigated as key variables for high spatial resolution imagery classification. Our findings show that each of the algorithms achieved nearly a 90% kernel density match using the United Nations Operational Satellite Applications Programme (UNITAR/UNOSAT) dataset as validation. The multilayer feedforward network was able to achieve an error rate below 40% in detecting damaged buildings. Spatial features of texture and structure were far more important in algorithmic classification than spectral information, highlighting the potential for future implementation of machine learning algorithms which use panchromatic or pansharpened imagery alone.

Keywords: earthquake damage; machine learning; computer vision; Random Forests; neural networks

1. Introduction

Earthquakes accounted for over 60% of all natural disaster-related deaths from 2001 to 2011—a danger that will likely increase due to rapid global urbanization [1]. Immediately after an earthquake occurs, satellite imagery is a critical component of damage mapping. Hussain et al. noted that “information derived from remote sensing data greatly helps the authorities in rescue and relief efforts, damage assessment, and the planning of remedial measures to safeguard such events effectively” [2]. For immediate rescue operations, rapid damage maps derived from satellite imagery must be developed quickly. A study of the 1995 Kobe earthquake in Japan showed a drastic reduction of the total rescued and the proportion of survivors after the third day of recovery efforts [3,4]. However, because rapid mapping is required to balance immediacy with in-depth analysis, early mapping efforts often yield coarse damage assessments [5].

Remote sensing has been used widely to map the effects of major disasters such as earthquakes. Numerous studies have utilized electro-optical (EO), synthetic aperture radar (SAR), light detection and ranging (LiDAR), ancillary data, or a combination thereof for post-earthquake damage detection [1,5,6]. One technique for damage detection involves fusion of SAR and EO data in pixel-based damage detection. Stramondo et al. used a maximum likelihood (ML) classifier on SAR features derived from the European Remote Sensing mission in combination with EO data provided by the Indian Remote Sensing satellite in order to identify damaged structures following the 1999 Izmit, Turkey earthquake [7]. A similar approach combined SAR from COSMO/SkyMed mission and very high

Academic Editors: Roberto Tomas, Zhenhong Li and Prasad S. Thenkabail

Received: 5 August 2016; Accepted: 17 October 2016; Published: 20 October 2016

* Correspondence: cooner@vt.edu; Tel.: +1-210-870-7909
resolution (VHR) EO data from the Quickbird satellite to improve damage detection at block level after combining the two datasets in a pixel-based classification following the 2009 L'Aquila earthquake [8].

As early as 1998, object-based image analysis (OBIA) has been used to detect earthquake damage from remote sensing [9]. More recently, OBIA has been a continual focus in earthquake detection damage with many studies focusing on the use of unmanned aerial systems (UAS), LiDAR, and the popular image segmentation and classification software eCognition. Hussain et al. [2] fused GeoEye-1 VHR EO data and airborne LiDAR elevation models derived from the RIT-ImageCAT UAS for image segmentation using the Definiens (now eCognition) software suite. The data were classified using nearest neighbor and fuzzy membership sets to detect damaged buildings and rubble following the 2010 Haiti earthquake. Similarly, Pham et al. [6] used aerial VHR RGB composite and LiDAR data (also from the RIT-ImageCAT UAS) along with eCognition for object segmentation and damage detection.

The application of machine learning algorithms (MLAs) to earthquake damage detection is a relatively new area of study. MLAs actively adapt and learn the problem at hand, often mimicking natural or biological systems, instead of relying on statistical assumptions about data distribution [10]. In addition to overall improved accuracy [11,12], MLAs have several advantages compared to traditional classification and change detection methods. MLAs work with nonlinear datasets [11,13], learn from limited training data [12,14], and successfully solve difficult-to-distinguish classification problems [15].

Ito et al. [16] used learning vector quantization (LVQ), a type of artificial neural network (ANN) to classify SAR features signifying damage after the 1995 Kobe earthquake. Li et al. [17] used a two-class support vector machine (SVM) on pre- and post-earthquake Quickbird imagery along with spatial relations derived from the local indicator of spatial association (LISA) index to detect structures damaged by the Wenchuan earthquake of 2008. Haiyang et al. utilized a SVM approach in combination with eCognition image segmentation on the RIT-ImageCAT RGB and LiDAR data, as well as the textural features of contrast, dissimilarity, and variance derived from the gray level co-occurrence matrix (GLCM) to detect urban damage in Port-Au-Prince [18]. Kaya et al. [19] used OBIA in combination with support vector selection and adaptation (a type of SVM) on pansharpened Quickbird imagery to conduct damage detection for specific buildings within Port-au-Prince after the 2010 earthquake. While OBIA using SVMs have been researched extensively in the past, ANNs, particularly radial basis function neural networks (RBFNNs), and Random Forests (RF) have shown promise in pattern recognition and image classification [15,20] and have yet to be examined in the application of earthquake damage detection. All three algorithms require parameter-tuning process to achieve optimal performance and a cross-validation approach can be applied to automate the parameter-tuning. SVM has an advantage in dealing with small sample size problems due to its sparse characteristics. However, for applications where a large number of training samples are available, SVM often yields a large number of support vectors, resulting in unnecessary complexity and a long training time [21].

Evaluating structural dimensions such as the Laplacian of Gaussian (LoG) and object-based metrics in addition to spectral and textural information could greatly increase damage detection rates. LoG, a blob detection technique, has been used for medical applications in nuclei mapping [22] and for the detection of buildings in bitemporal images [23]. As discussed earlier, OBIA has shown strong results in urban scenes and earthquake damage detection. Huang and Zhang had success applying the popular mean-shift segmentation algorithm for urban classification in hyperspectral scenes while statistical region merging (SRM) is another segmentation approach which is robust to noise and occlusions [24,25]. Additionally, various metrics such as rectangular fit, morphological shadow index, and morphological building index can describe the structure of objects in the scene before or after segmentation [26,27]. Applying structural descriptors such as a LoG filter and segmentation derived metrics to high resolution satellite imagery as an additional input to an MLA could evince damage in difficult to detect scenarios such as a pancake collapse [22,28]. The robustness and generalizability
of RF and ANN along with the additional dimensions of texture and structure may provide higher accuracies in the face of imperfect input data.

In past disasters, by the time an automated change detection scheme is ready for implementation, a crowdsourced team of visual interpreters is already mapping damaged buildings [29]. Dong and Shan mention that while manual digitization of damaged structures requires trained image analysts and is unsuitable for large areas, “visual interpretation remains to be the most reliable and independent evaluation for automated methods” [1]. Additionally, many previous studies suggest detection schema which require ancillary data such as UAS products, LiDAR, or GIS databases. Many of these products are unavailable in developing regions where the death toll is highest [30]. Using MLAs (RF and ANN), a rapid damage map derived from readily available multispectral imagery could allow for a minimal compromise between time and accuracy and allow first responders to more rapidly allocate their resources in a crisis. RF and ANNs along with derived textural and structural features may provide improved balance between rapid and accurate damage detection. The main purposes of this study include:

- Assess the performance of neural networks (to include radial basis function neural networks), and Random Forests on very high resolution satellite imagery in earthquake damage detection
- Investigate the usefulness of structural feature identifiers to include the Laplacian of Gaussian and rectangular fit in identifying damaged regions

2. Materials and Methods

2.1. Study Area and Data

The earthquake on 12 January 2010 near Port-au-Prince, Haiti, was an exceptionally devastating event. The initial shock of 7.0 Mw caused an astounding death toll: between 217,000 and 230,000 were reported dead by the Haitian government [2] and the official estimate has grown to 316,000 [31]. Additionally, the earthquake and its subsequent aftershocks caused extensive damage in and around Port-au-Prince including numerous landmark buildings such as the National Palace [2]. Because of the high density of collapsed and damaged structures available for training and validation of MLAs, the 2010 Haiti earthquake is an ideal case study to evaluate automated damage detection methods.

For this research, we obtained high resolution multispectral and panchromatic remote sensing data from the DigitalGlobe Foundation. A pre-disaster panchromatic image was acquired in December of 2009 by the WorldView-1 satellite (accessed via DigitalGlobe’s EnhancedView system) and post-disaster multispectral and panchromatic images were acquired on 15 January 2010 by the Quickbird 2 satellite. All datasets (see Table 1) were resampled to 0.6 meters using the nearest neighbor technique. Images were atmospherically corrected to top of atmosphere (TOA) reflectance [32] and clipped to the study area of central Port-Au-Prince. Pre- and post-earthquake imagery were co-registered using 15 control points and a third order polynomial transformation with a root mean square error (RMSE) of 0.55 m.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Native Resolution</th>
<th>Acquisition Date</th>
<th>Look Angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>WorldView-1</td>
<td>0.5 m panchromatic</td>
<td>7 December 2009</td>
<td>27.62</td>
</tr>
<tr>
<td>QuickBird-2</td>
<td>2.4 m multispectral0.6 m panchromatic</td>
<td>15 January 2010</td>
<td>20.7</td>
</tr>
</tbody>
</table>

The study area (Figure 1) was divided into training and validation regions. Training samples for the algorithms were selected by manually digitizing polygons over damaged structures, ensuring that
training data represented the input space of the entire study area. A total of 1,214,623 undamaged and 134,327 damaged pixels were used for training. For validation purposes, over 900 buildings were digitized and marked as damaged or undamaged according to the UNOSAT Haiti damage assessment [33]. This dataset classifies damage as points classified according to the European Macroseismic Scale developed in 1998 [34]. This classification schema defines five damage levels ranging from minor/no damage to destruction [1,5,6,35]. In order to simplify the damage detection problem, the UNOSAT points belonging to the most severe damage class (EMS grade V) were extracted for the validation sets and used for the damaged building class. This action is required primarily because visual distinction between classes is difficult even with sub-meter pixel resolution [1]. While these validation points have been previously used in several studies [6,36], it is important to note that the UNOSAT/UNITAR dataset was derived through manual interpretation of satellite and aerial imagery and very few points were ground verified. However, the assessment remains the standard for the validation of damage that occurred in the 2010 Haiti earthquake.

![Figure 1. Study area in Port-au-Prince showing training, building test and kernel density test sites (satellite image courtesy of the DigitalGlobe Foundation).](image)

### 2.2. Texture and Structure

Because several studies have shown that classification and change detection performance can be increased with the addition of spatial information [11,20], textural and structural information was extracted from the pre- and post-earthquake panchromatic images. Figure 2 shows the impact of earthquake damage on two selected textural and structural features. Entropy, energy, dissimilarity, and homogeneity are all second order texture features derived from the GLCM that have been correlated with damage or used as proxies for damage in previous studies [35,37–39]. In order to reduce dimensionality and eliminate redundancy, the two consistently correlated GLCM features of Entropy (a measure of gray level randomness) and Dissimilarity (a measure of gray level difference (the square of contrast)) were chosen as texture inputs.

\[
\text{Entropy} = - \sum_{i,j} \text{GLCM}_{i,j} \times \log(\text{GLCM}_{i,j})
\] (1)
Dissimilarity = \sum_{i,j} GLCM_{i,j} \times |i - j| \tag{2}

In order to reduce noise, a Gaussian filter (\(\sigma = 1\)) was applied to the panchromatic images before measuring image Entropy and Dissimilarity. A 7 \times 7 sliding window was used to compute the 0° GLCM and the corresponding texture values were calculated for both the before and after panchromatic images.

To define structural features, a Laplacian of Gaussian (LoG) filter was applied as it is one of the most commonly implemented methods of blob detection. A 2-dimensional LoG filter of size \((x, y)\) can be constructed using:

\[
\text{LoG}(x, y, \sigma) = \frac{x^2 + y^2 + 2\sigma^2}{\pi\sigma^4} \times e^{-\frac{x^2 + y^2}{2\sigma^2}} \tag{3}
\]

Because the Laplacian computes the second derivative, abruptly changing regions of an image will be highlighted. When combined with a Gaussian smoothing filter, blobs can be detected at different scales defined by \(\sigma = (r - 1)/3\) where \(r\) is the radius of a blob of interest [22]. In order to detect buildings of different sizes, a multiscale approach is required. A total of 50 separate convolutions of the LoG filter were applied to the pre- and post-earthquake imagery with sigma adjusted at equal intervals between the values of 15 and 35. Additionally, the LoG filter size was increased from 10 \times 10 to a 25 \times 25 window in order to accommodate the larger sigma values. The minimum response in scale space was assigned to each pixel due to the fact that LoG filters produce negative responses for bright areas (the majority of buildings produced high reflectance in the panchromatic images).

Figure 2. This figure comparing building damage shows the effects of an entropy filter and a LoG filter on pre- and post-earthquake imagery considering two different collapse scenarios: (a) a normal structural collapse; (b) a pancake collapse. Pre-earthquake satellite images © copyright DigitalGlobe. Post-earthquake satellite images courtesy of the DigitalGlobe Foundation.

Other structural identifiers can be derived through object-based image analysis (OBIA). While a number of geometric indices are possible through OBIA, this study chose rectangular fit as an input.
feature due to its obvious possibility for correlation with building shape. The analysis of equal area rectangles, drawn according to object moment, has been used as a more robust version of minimum bounding rectangle comparison [26]. Rectangular fit is defined as

$$\text{RectFit} = \frac{(A_R - A_D)}{A_O}$$  \hspace{1cm} (4)

where $A_O$ is the area of the original object, $A_R$ is the area of the equal rectangle ($A_O = A_R$) and $A_D$ is the overlaid difference between the equal rectangle and the original object [26]. For this study, image segmentation was performed on both the before and after panchromatic images using statistical region merging (SRM) as posited by Nock and Nielsen because of its fast and simple implementation [40]. Setting the scale parameter $Q = 512$ enabled the many small buildings in the scene to be distinguished. Each object’s rectangular fit was calculated and included as an input dimension.

2.3. Machine Learning Algorithms

Once all preprocessing steps were taken and textural and structural features were derived, the training and validation datasets were transformed into 14-dimensional arrays consisting of pre-panchromatic; pre-entropy; pre-dissimilarity pre-LoG; pre-rectangular fit; post-panchromatic; post-entropy; post-dissimilarity; post-LoG; post-rectangular fit; blue; green; red, and near infrared multispectral layers (see Table 2). Values in these layers were normalized to fall between the values of −1 and 1 in order to standardize the input and validation layers. While cross-validation was implemented for each MLA, exhaustive parameter analysis was not performed due to the large number of training samples available. All MLA design, algorithm implementation, training and validation were performed using MATLAB release 2015a on a 3.5 GHz quad core processor with 64 GB RAM.

<table>
<thead>
<tr>
<th>Input Feature</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panchromatic (450–900 nm)</td>
<td>WV1 and QB2</td>
</tr>
<tr>
<td>Entropy</td>
<td>WV1 and QB2</td>
</tr>
<tr>
<td>Dissimilarity</td>
<td>WV1 and QB2</td>
</tr>
<tr>
<td>LoG</td>
<td>WV1 and QB2</td>
</tr>
<tr>
<td>Rectangular fit</td>
<td>WV1 and QB2</td>
</tr>
<tr>
<td>Blue (450–520 nm)</td>
<td>QB2</td>
</tr>
<tr>
<td>Green (520–600 nm)</td>
<td>QB2</td>
</tr>
<tr>
<td>Red (630–690 nm)</td>
<td>QB2</td>
</tr>
<tr>
<td>Near infrared (760–900 nm)</td>
<td>QB2</td>
</tr>
</tbody>
</table>

2.3.1. Neural Networks

Artificial neural networks (ANN) are of continued interest due to their ability to approximate any function given sufficient neurons in each layer, the flexibility of data distribution, and their reduced computational complexity compared with statistical classification methods [11,13,20,41]. A feedforward ANN (also known as multilayer perceptron) assigns small random multiplicative weights and additive biases to input neurons and iteratively adjusts them with each additional training data input, minimizing the surface of the performance function representing the error between the training data and the network’s output in order to make the best classification [11,13,41]. Neural network design is difficult due to the number of free parameters and ambiguous requirements for network depth and complexity which depend upon the problem at hand [41]. As a result, neurons were grown and pruned in various combinations within 1 to 3 hidden layers until the network performance was maximized. The resulting ANN consisted of two hidden layers of 20 neurons each with sigmoid
transfer functions, and a binary softmax output layer (see Figure 3). Training was accomplished using backpropagation and the scaled conjugate gradient algorithm to minimize cross-entropy and two-fold cross validation via early stopping (when decrease of cross-entropy in a validation subset ceased) was used to prevent overfitting [41,42]. Finally, variable importance was measured by retraining the network an additional 14 times with one of each of the input dimensions set to zero for all data points and recording the change in cross-entropy after a set number of iterations (400).

An additional ANN which has been used for classification is the radial basis function neural network (RBFN). The first layer of RBFNs consists of basis functions, centered throughout the input space and the second layer consists of a transfer function which combines the results from the basis functions and classifies them into the categories of interest [41]. Similar to the backpropagation network design, the number of basis functions was determined through trial and error and grown until performance was maximized. The radial basis layer consisted of 150 Gaussian functions centered using an unsupervised k-Means approach, which intelligently assigns cluster centers in areas of the input space where high activity exists [43,44]. The radial basis outputs fed into the second layer, which consisted of a softmax transfer function, trained by minimizing cross-entropy using the scaled conjugate gradient with early stopping as before.

2.3.2. Random Forests

Random Forests, which was originally proposed by Breiman [45], is an ensemble classifier consisting of a large number of classification and regression trees (CART), where final classification is performed by a winner-takes-all voting system. The algorithm trains each tree on an independently drawn subset of the original data (bootstrap aggregating or bagging) and determines the number of features to be used at each node by the evaluation of a random vector [45,46]. Because RF is an ensemble classifier, the Law of Strong Numbers dictates that RF will converge without overfitting the model, so that the computationally optimal number of trees can be found through testing the algorithm. Additional benefits include robustness to outliers and noise and built-in estimates of error and variable importance [45,47]. Using MATLAB’s TreeBagger function, training was accomplished using 400 (additional tree growth resulted in no further decrease in out-of-bag error) classification trees grown by selecting three variables (at random) for each node split. Additionally, the out-of-bag error was collected and variable importance was measured for comparison with the ANN approach.
2.4. Testing and Accuracy Assessment

After training, validation testing was performed on two different datasets. The first dataset consisted of an area including the National Palace just to the south of the training area where building footprints were previously digitized. For this testing area, output from the algorithms was converted to polygons and a building by building accuracy assessment was performed on the first dataset, resulting in a simple confusion matrix. This approach was taken in order to ensure that our accuracy measurements were based on actual objects in the scene rather than a pixel-by-pixel assessment. The second validation dataset included a larger area in order to test the algorithms’ performance when based on kernel density map rank matching. This secondary assessment followed the procedures used by Tiede et al. [36] and Pham et al. [6] in which the centroid points of the damage polygons are computed and kernel density raster maps are generated for both the test damage points as well as the UNOSAT/UNITAR control data points. The damage density for each dataset was projected onto a 20 m raster grid and the rank value of each cell was computed according to the quartile of density that the cell belonged to. The two maps were overlaid and accuracy was assessed by subtracting the UNOSAT/UNITAR kernel density map from the test kernel density maps, with final output values ranging from $-3$ (omission error) to $+3$ (commission error).

3. Results

Table 3 (supplemented by the confusion matrices in Appendix A) compares overall performance of three MLAs using building-by-building and kernel density accuracy assessment. Our findings showed that the multilayer feedforward ANN outperformed both the RBFNN and random forests with an omission error rate of 37.7%. Figure 4 matches the spatially explicit locations of damage detected by the ANN algorithm with the digitized buildings marked as damaged or undamaged (please refer to Appendix B for the other algorithms’ damage maps). Both RBFNN and RF had higher overall accuracies, yet drastically underestimated damage. RBFNN and RF created a high user’s accuracy for the damaged building class, but a lower producer’s accuracy. The feedforward ANN also had the shortest runtime, an advantage that was primarily gained through the extremely fast implementation of the ANN for testing. The kappa value for each of the algorithms indicated that the distribution of damaged and undamaged buildings could not be accounted for by random chance, however both of the ANNs clearly outperformed Random Forests in this measure as well.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Train + Test Runtime (s)</th>
<th>Overall Accuracy (%)</th>
<th>Building Class Omission Error (%)</th>
<th>Cohen’s Kappa</th>
<th>Kernel Density Match (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>623</td>
<td>74.14</td>
<td>37.69</td>
<td>0.402</td>
<td>87.41</td>
</tr>
<tr>
<td>RBFNN</td>
<td>1532</td>
<td>77.26</td>
<td>55.97</td>
<td>0.3951</td>
<td>90.25</td>
</tr>
<tr>
<td>RF</td>
<td>8692</td>
<td>76.14</td>
<td>71.27</td>
<td>0.3057</td>
<td>88.77</td>
</tr>
</tbody>
</table>

For the kernel density accuracy assessment, a good result was measured as a value in the comparison map that ranged between $-1$ and $1$ in accordance with the Tiede et al. [36] approach. RBFNN reached a 90% kernel density map match (shown in Figure 5), outperforming the standard ANN and RF, which were not far behind. This indicates that the radial basis function ANN may be able to generalize to a larger area with greater success than either a feedforward network or Random Forests. Even so, each of the algorithms performed at higher accuracies when generalizing the distribution of damage over a wide area instead of detecting individual, building level damage.
Figure 4. Results of the feedforward ANN (highest performer) on the building test dataset. Direct output of the algorithm (pink) overlays the building categories are represented by shaded polygons (satellite image courtesy of the DigitalGlobe Foundation).

Figure 5. Results of the RBFNN algorithm on the kernel density test dataset. Each 20 m cell is marked from −3 to 3, where an adequate density match falls between −1 and 1 (satellite image courtesy of the DigitalGlobe Foundation).

As well as examining an algorithm’s wide area generalizability, kernel density accuracy assessment also allowed for investigation into areas of common error of both omission and commission. One of the more interesting results was that these areas were common in all three algorithms. A common error of commission occurred in the north-center of the test area and coincided with
the development of an internally displaced persons (IDP) camp (see Figure 6). According to the algorithms, this camp broke up the “structure” of the underlying field and increased the randomness of the texture, which led it to misclassifying the area as a damaged building. Figure 7 shows a common area of omission error contained within the map in the central region of the testing area. The underlying cause of error in this region is the scene complexity and high density of small structures before the earthquake occurred. This preexisting randomness and highly variable structure and texture were difficult for the algorithm to interpret, leading to an error of omission. Even so, with kernel density matching occurring in nearly 90% of cells for each algorithm, a wide area damage density classification was successful for both of the algorithms tested.

![Figure 6](image-url)  
**Figure 6.** Results of the RBFNN on the kernel density test dataset overlaid before (a); and after (b) panchromatic imagery. This ad hoc IDP caused an error of commission among each of the algorithms. Pre-earthquake satellite image copyright © DigitalGlobe. Post-earthquake satellite image courtesy of the DigitalGlobe Foundation.

![Figure 7](image-url)  
**Figure 7.** Results of the RBFNN on the kernel density test dataset overlaid before (a); and after (b) panchromatic imagery. This area of complex small structures was a common error of omission among the algorithms. Pre-earthquake satellite image copyright © DigitalGlobe. Post-earthquake satellite image courtesy of the DigitalGlobe Foundation.
Finally, variable importance shows similar trends for ANN and RF algorithms. There were a few variables that showed rather different utilization between the algorithms. Figure 8 shows the change in error between each variable and was developed by averaging the assigned variable importance between the pre- and post-earthquake datasets. Overall, the multispectral variables had lower changes in out-of-bag error and cross entropy in comparison to the textural and structural values, however the panchromatic images and the near infrared band was useful to each of the algorithms. Of the two texture measures, entropy was utilized more than dissimilarity in all three algorithms. Rectangular fit was marked as important for both the ANN and RF however the Laplacian of Gaussian filter was more impactful on the RBFNN assessment and also had a moderately high impact on RF performance.

![Figure 8. Chart of variable importance for both ANNs and RF. Overall, structure and texture played a larger role in classification than spectral information. Random Forests use the ΔOOB Error measure while ANN and RBFNN use ΔCross Entropy.](image)

4. Discussion

It is difficult to determine outright which algorithm is better. While the multilayer ANN outperformed RF in building by building assessments and required slightly less overall training time, it required a large number of training samples in order to perform well. This is not necessarily the case for the RF algorithm. Also, it is rather easy to overfit an ANN to the training data, which can be avoided using RF due to its nature as an ensemble classifier [41,45]. Finally, ANNs can also become stuck at local minima in the performance surface without reaching the global solution, yielding an insufficient result [41]. However, in our study, the multilayer ANN had the lowest rates of error in detecting damaged buildings, without sacrificing much performance in wide area generalization or overall accuracy. SVMs, while not examined here, have shown promise in earthquake detection in previous studies [17–19]. While our study focused on ANNs and RF, as little research has been done on their applicability to earthquake detection problems, future studies may investigate the performance of these algorithms (to include SVM) with respect to training sample size.

Beyond damage detection performance, practical considerations require an investigation into time complexity, particularly when considering any kind of operationalization of an algorithm in automatic damage detection. RF took much longer than either of the ANNs to train and test the datasets. The time complexity of a single classification and regression tree is $O(mn \cdot \log n)$ where $m$ is the total number of variables and $n$ is the number of samples [48]. Because RF is an ensemble classifier, the overall time complexity of Random Forests can be summarized as $O(M(mn \cdot \log n))$ where $M$ is the number of trees grown. For a large number of samples with moderate dimensionality, this can be quite
time consuming. In contrast, neural network complexity is highly dependent on network architecture. Time complexity for the scaled conjugate descent algorithm is often polynomial, overall complexity is determined by the problem and the number of free parameters (weights, biases, or basis functions (in the case of RBFNN)) required to describe that problem [49]. As such, testing showed that the ANNs trained faster and tested faster, which is important to consider given the requirement to process a potentially large amount of imagery in an operational context.

As previously discussed, a number of preprocessing steps are required to develop each of the textural and structural dimensions. Also, a k-means unsupervised clustering was used as part of the RBFNN algorithm to intelligently center the basis functions before training. Each of these steps adds time and complexity to the final product. For future studies, the parallelization of many of these processes is one way to greatly reduce computational time. Our data were gathered using serial processing (primarily because a parallel implementation of k-means was not immediately available) in order to establish a baseline and fairly assess each algorithm, however parallel implementations (which include graphics processing units [GPUs]) of both ANN and RF training are readily available for use and will greatly speed up training and implementation of these machine learning algorithms.

The actual results from this study mirrored what was expected quite well; the areas of imagery where texture and structure were broken up were often identified as damage, as one would expect. As mentioned in the results section, one interesting finding was that each of the algorithms erroneously identified IDP camp areas as building damage. These IDP camps are ad hoc structures (tents, tarps, and shanties) built primarily on open spaces. As these IDP camps were erected, they broke up the coherent texture and structure of the underlying terrain, causing the algorithm to mark them as damage. While this is technically an error of commission, it is nevertheless a useful result in showing the power of MLAs in seeking out patterns as well as their ability to simultaneously detect damage and displaced persons. In an operational context, the MLA results in combination with a priori knowledge of building distribution via a GIS database would allow first responders and emergency planners to easily distinguish damaged structures from these IDP camps.

As the experiment on variable importance showed, the textural and structural features were some of the most important factors which allowed both ANNs and RF to detect damage and IDP camps. Stramondo et al. [7] also used important textural features for earthquake damage detection although in their study a maximum likelihood classifier was used. This line of thinking, paramount to computer vision applications, is expanded here by using more intelligent algorithms and readily available data. The importance of the panchromatic features along with the texture and structure products derived exclusively from that panchromatic imagery presupposes that future implementations of machine learning may be able to perform earthquake damage detection from panchromatic imagery alone. One reason that the multispectral imagery was not a driving variable is simply due to resolution; the native 2.4 m is too coarse to detect many of the features associated with earthquake damage. Interestingly, the only multispectral product which was found to be an important variable for each of the algorithms was the near infrared band, which may have resulted from a correlation between exposed rubble and a higher near infrared reflectance. These findings may guide future research in determining which variables to focus on in earthquake damage studies.

Our focus on simple panchromatic imagery is a departure from many previous earthquake studies. The state-of-the-art focuses on LiDAR, SAR, unmanned aerial vehicles, and the software suite eCognition [1]. However, the access and availability of these additional data requirements may be limited in the aftermath of a destructive earthquake in a less developed region such as Haiti. A return to easily accessible data products such as multispectral or even panchromatic imagery alone could allow a MLA (potentially even one that is pre-trained) to detect imagery without the requirement of ancillary data. One potential disadvantage of the reliance on bitemporal VHR imagery is the requirement for precise coregistration. Different look angles can cause problems in classification and change detection. While image registration is still important to our study, a small look angle difference may not be critical due to our use of textural and structural features rather than the VHR imagery.
alone. Additionally, registration errors can be seen as a source of noise in the system; each of the MLAs used has been shown to be robust to noise [41,45]. The difference in our look angles (~7°) did not appear to cause any damage detection errors in a visual survey of our results. Future research may investigate the limits of acceptable look angle differences or use a complex coregistration approach to eliminate the issue altogether [50].

The future of earthquake damage detection may lie in deep convolutional neural networks (DCNN) coupled with high performance computing and GPUs. DCNNs have already pushed the boundaries of artificial intelligence and image recognition; rather than being told which textural, structural, or spectral inputs should be used, these networks automatically learn and identify the defining features (convolutions) of the problem at hand in order to classify future samples [51,52]. Initial results are promising. Using MatConvNet (a deep learning library for Matlab), we experimented with training a DCNN with the VGG-F architecture following the approach and using the hyperparameters described by Chatfield et al. [52,53]. We segmented the post-earthquake pansharpened image using SRM and trained the DCNN on each labeled, extracted object. The DCNN was not only able to detect buildings at a comparable rate (>55% detection rate), it was able to distinguish between damage and IDP camps (>65% detection rate) and did so using an after-only pansharpened image, reducing data requirements and eliminating the need for coregistration. A pre-trained DCNN optimized specifically for earthquake detection may offer a robust and operationally implementable solution to the much studied topic of earthquake damage detection.

5. Conclusions

This study analyzed the use of machine learning algorithms to include feedforward neural networks, radial basis function neural networks, and Random Forests in detecting earthquake damage caused by the 12 January 2010 event near Port-au-Prince, Haiti. The algorithms’ efficacy was improved by providing coregistered 0.6 m multitemporal imagery, texture features (dissimilarity and entropy), and structure features (Laplacian of Gaussian and rectangular fit) as inputs. Detection results were assessed on a structure specific basis by digitizing more than 900 buildings and comparing each MLA’s response to the UNITAR/UNOSAT validation set. For a wide area generalization, a kernel density map comparison was performed between each of the algorithms’ classification results and the UN damage validation points.

The feedforward ANN consisting of two hidden layers had the lowest error rate (<40%) without sacrificing much performance in overall accuracy or generalization to wider area damage estimates. Additionally, textural and structural features derived from panchromatic imagery were shown to be more important than spectral variables in the algorithms’ classification process. Each algorithm had common errors of commission occurring around ad hoc IDP camps that were spontaneously formed in open spaces following the earthquake; this technically incorrect result can be easily integrated with a GIS layer containing building footprints.

The results of this study show that not only do MLAs have potential for use in earthquake damage detection, but that panchromatic or pansharpened imagery can be the exclusive data source for training and testing. Measures of variable importance found that the panchromatic derived texture and structure products are the main drivers behind the success of these “shallow” machine learning algorithms. Future research into an operationally implementable machine learning method is warranted. An attractive next step is to transition into deep learning where convolutional neural networks move beyond pixel-based or object-based paradigms and begin to detect remotely sensed features in ways akin to natural image recognition.

Acknowledgments: The authors would like to thank Stephen Prisley for his supporting comments and direction in the study. Also, we extend our gratitude to our peer-reviewers and thank them for their time and helpful comments and suggestions. We would like to thank Chris McCormick for his excellent tutorial on RBFNN and corresponding software which was modified for use here. The authors would also like to thank the DigitalGlobe Foundation for their critical support in providing QuickBird imagery and the Virginia Tech Open Access Subvention Fund.
for supporting this publication. Finally, we would like to thank Earl and Marion Nutter; without their entrusted scholarship this study would not be possible.

**Author Contributions:** Austin Cooner designed the experiment, coded the software, and wrote the manuscript. James Campbell provided direction and edited the manuscript. Yang Shao offered oversight, guidance, and edited the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest. The DigitalGlobe Foundation and scholarship support had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

**Appendix A**

**Table A1.** Confusion matrices for each of the algorithms, including user’s accuracy (UA) and producer’s accuracy (PA).

<table>
<thead>
<tr>
<th></th>
<th>Undamaged UNOSAT</th>
<th>Damaged UNOSAT</th>
<th>ANN UA</th>
<th>RBFNN UA</th>
<th>RF UA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undamaged ANN</td>
<td>498</td>
<td>101</td>
<td>83.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Damaged ANN</td>
<td>131</td>
<td>167</td>
<td>56.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANN PA</td>
<td>79.17</td>
<td>62.31</td>
<td></td>
<td>79.31</td>
<td></td>
</tr>
<tr>
<td>Undamaged RBFNN</td>
<td>575</td>
<td>150</td>
<td></td>
<td>54.00</td>
<td>76.04</td>
</tr>
<tr>
<td>Damaged RBFNN</td>
<td>54</td>
<td>118</td>
<td>68.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBFNN PA</td>
<td>91.41</td>
<td>44.03</td>
<td></td>
<td></td>
<td>77.00</td>
</tr>
<tr>
<td>Undamaged RF</td>
<td>606</td>
<td>191</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Damaged RF</td>
<td>23</td>
<td>77</td>
<td>77.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RF PA</td>
<td>96.34</td>
<td>28.73</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Appendix B**

**Appendix B.1. Algorithmic Damage Map for the RBFNN Algorithm**

![Figure A1](image-url)  

**Figure A1.** Results of the RBFNN algorithm on the building test dataset. Direct output of the algorithm (pink) overlays the building categories are represented by shaded polygons (satellite image courtesy of the DigitalGlobe Foundation).
Appendix B.2. Algorithmic Damage Map for the Random Forests Algorithm

Figure A2. Results of the RF algorithm on the building test dataset. Direct output of the algorithm (pink) overlays the building categories are represented by shaded polygons (satellite image courtesy of the DigitalGlobe Foundation).

References

Article

Earthquake-Induced Building Damage Detection with Post-Event Sub-Meter VHR TerraSAR-X Staring Spotlight Imagery

Lixia Gong, Chao Wang, Fan Wu, Jingfa Zhang, Hong Zhang and Qiang Li

Abstract: Compared with optical sensors, Synthetic Aperture Radar (SAR) can provide important damage information due to its ability to map areas affected by earthquakes independently from weather conditions and solar illumination. In 2013, a new TerraSAR-X mode named staring spotlight (ST), whose azimuth resolution was improved to 0.24 m, was introduced for various applications. This data source made it possible to extract detailed information from individual buildings. In this paper, we present a new concept for individual building damage assessment using a post-event sub-meter very high resolution (VHR) SAR image and a building footprint map. With the building footprint map, the original footprints of buildings can be located in the SAR image. Based on the building imaging analysis of a building in the SAR image, the features in the building footprint can be extracted to identify standing and collapsed buildings. Three machine learning classifiers, including random forest (RF), support vector machine (SVM) and K-nearest neighbor (K-NN), are used in the experiments. The results show that the proposed method can obtain good overall accuracy, which is above 80% with the three classifiers. The efficiency of the proposed method is demonstrated based on samples of buildings using descending and ascending sub-meter VHR ST images, which were all acquired from the same area in old Beichuan County, China.

Keywords: earthquake; damage assessment; building; Synthetic Aperture Radar; TerraSAR-X; high resolution

1. Introduction

Damage detection after earthquakes is an important issue for post-disaster emergency response, impact assessment and relief activities. Building damage detection is particularly crucial for identifying areas that require urgent rescue efforts. Remote sensing has shown excellent capability for use in rapid impact assessments, as it can provide information for damage mapping in large areas and in an uncensored manner, particularly when information networks are inoperative and road connections are destroyed in areas impacted by earthquakes. Compared with optical sensors, Synthetic Aperture Radar (SAR) can provide important damage information due to its ability to map affected areas independently from the weather conditions and solar illumination, representing an import data source for damage assessment.

Many approaches for damage detection with SAR imagery have been introduced [1,2]. Change detection with pre- and post-event SAR images is widely used for damage assessment,
Remote Sens. 2016, 8, 887

and it has become a standard procedure [3]. The differences between the backscattering intensity and intensity correlation coefficients of images before and after earthquakes indicate potential damage areas. In addition, the interferometric coherence of pre- and post-event SAR data can also provide important information for damage assessment [1]. In the case of change detection, the pre- and post-event data are needed for comparison. However, in many areas, almost no archived SAR images, especially for high-resolution SAR data, are available. Therefore, research on building damage assessment with only post-event SAR images is necessary.

The new generation of high-resolution SAR satellite systems was first deployed for earthquake monitoring during the Wenchuan Earthquake in 2008 [3]. Meter-resolution SAR images from TerraSAR-X and COSMO-SkyMed were acquired for damage assessment. Although the image resolution has improved, it is difficult to identify collapsed buildings using single meter-resolution SAR images [3,4]. After 2013, a new TerraSAR-X mode named staring spotlight (ST), whose azimuth resolution was improved to 0.24 m, was introduced. The sub-meter very high resolution (VHR) SAR data source provides new opportunities for earthquake damage mapping because more details can be observed in images, making it possible to focus an analysis on individual buildings.

Previous works have generally concentrated on the building block level of damage assessment with low- or medium-resolution SAR images. Few studies have focused on individual buildings due to lack of VHR SAR images. Balz et al. [3] analyzed the appearances of destroyed and partly destroyed buildings using high-resolution SAR images and proposed a technical workflow to identify building damage using post-seismic high-resolution SAR satellite data. Another method was developed to detect building damage in individual building units using pre-event high-resolution optical images and post-event high-resolution SAR data based on the similarity measurement between simulated and real SAR images [5]. Wang et al. [6] proposed an approach for post-earthquake building damage assessment using multi-mutual information from pre-event optical images and post-event SAR images. Brunner et al. [4] analyzed a set of decimeter-resolution SAR images of an artificial village with different types of destroyed buildings. The analysis showed that the decimeter-resolution SAR data are fine enough to classify building signatures according to basic damage classes. Kuny et al. [7] used simulated SAR images of different types of building damage for feature analysis and proposed a method for discriminating between debris heaps and high vegetation with simulated SAR data and TerraSAR-X images [8]. Wu et al. [9,10] analyzed different building damage types in TerraSAR-X ST mode data by visual interpretation, statistical comparison and classification assessment. The results showed that the data effectively separated standing buildings from collapsed buildings. However, the damaged building samples were obtained by expert interpretation and manual extraction. This method is time consuming and technique dependent because the exact edges of buildings and debris must be acquired.

In sub-meter VHR SAR images, more features, such as points and edges of objects, become visible, more details of individual buildings can be observed and new features of buildings are shown [11]. It is difficult to detect specific buildings in complicated surroundings in VHR images. Moreover, collapsed building debris is visually similar to other objects such as high vegetation [8]. This could lead to false alarms when detecting debris heaps.

Therefore, in this paper, we propose a method to assess the damage to individual buildings affected by earthquakes using single post-event VHR SAR imagery and a building footprint map. Given a building footprint map, which can be provided by a cadastral map or extracted from a pre-event VHR optical image, the original footprints of buildings can be located in the geometrically-rectified post-event VHR SAR image. Because collapsed and standing buildings exhibit different characteristics in their original footprints, features can be extracted from the image. Then, the buildings can be classified as different damage types. We demonstrate the feasibility of the proposed method in Qushan town (31.833°N, 104.459°E), old Beichuan County, China, which was heavily damaged in the Wenchuan Earthquake on 12 May 2008. The ruins in the area are preserved as the Beichuan Earthquake Memorial Park. Therefore, some damaged buildings in the area can be used for algorithm analysis. In our
experiments, the post-event descending and ascending TerraSAR-X ST data are used for building damage assessment. Post-event airborne optical images, light detection and ranging (LIDAR) data, as well as in situ photography are used as reference data. The main novelty of the paper is the use of the VHR TerraSAR-X ST data for individual damage building analysis and demonstrating the concept for collapsed building assessment with building footprints.

The remainder of this paper is structured as follows. In Section 2, the properties of damaged buildings in VHR SAR images and damage detection problems are stated. In Section 3, we introduce the basic principle of the proposed method and describe damage detection with sub-meter VHR SAR data in detail. The study area and dataset are introduced in Section 4. Section 5 gives the results and analysis of the experiments. Finally, the discussion and conclusions are presented in Sections 6 and 7, respectively.

2. Properties of Damaged Buildings in VHR SAR Images and Damage Detection Problems

Due to the side-looking image mode of SAR sensors, some unique features, such as layover, fore-shortening effects and multi-path propagation, can be observed in SAR images. Those effects make interpretation with SAR images difficult compared to using optical images. However, the effects are also the unique patterns of buildings, which can be differentiated from other objects in SAR images.

In general, buildings have very distinct appearances in SAR images due to their cube-like or regular shapes. Buildings in SAR images typically consist of four zones: a layover area, corner reflection, roof area and shadow area. When an earthquake occurs, buildings may be destroyed or collapse. Destroyed buildings have different scattering distributions. Strong double-bounce scattering may be missing because the vertical wall facing the sensor may be collapsed, and the wall-ground dihedral corner reflector may be destroyed. The layover area may also be missing. Furthermore, the shadow area of a building will be reduced and potentially missing, depending on the structure and slope of the ruins.

Figure 1 illustrates a standing building after an earthquake. Figure 1a is an in situ photograph; Figure 1b is an optical image; and Figure 1c,d show the building in TerraSAR-X ST mode SAR images, respectively. The azimuth direction of Figure 1c is from bottom to top, and the range direction is from left to right. The azimuth direction of Figure 1d is from top to bottom, and the range direction is from right to left. The building has a flat roof with slight damage. In the SAR image, the double-bounce line, layover area and shadow area are obvious. The layover areas can be detected as the areas facing the sensor. The shadow areas are distinguishable from their surroundings.

Figure 2 illustrates a totally collapsed building after an earthquake. Figure 2a is an in situ photograph; Figure 2b is an optical image; and Figure 2c,d show the building in TerraSAR-X ST mode SAR images, respectively. The azimuth direction of Figure 2c is from bottom to top, and the range direction is from left to right. The azimuth direction of Figure 2d is from top to bottom, and the range direction is from right to left. The damaged building was collapsed entirely,
leaving a heap of randomly-oriented planes mainly made of concrete. In the SAR image, the heap of debris of the collapsed building shows random scattering effects. Some bright scattering spots can be found, which are caused by corner reflectors, resulting from the composition of different planes. However, the double-bounce line, layover area and shadow area are missing in this case.

![Figure 2](image-url) A collapsed building: (a) in situ photograph; (b) optical image acquired in 2013; (c) TerraSAR-X ST ascending SAR image (the azimuth direction is from bottom to top, and the range direction is from left to right); (d) TerraSAR-X ST descending SAR image (the azimuth direction is from top to bottom, and the range direction is from right to left).

Thus, in Figures 1 and 2, we can see that the standing building and the totally collapsed building have their own unique features in the VHR SAR images. However, high vegetation shows features that are similar to those of the debris of the totally collapsed building and may cause false alarms [8]. Moreover, it is not easy to extract the exact edges of the debris due to the similarity between the debris and surroundings. Therefore, although the signature of the standing building shows different features compared to those of the collapsed building in the VHR SAR images, detecting the collapsed building debris using only post-event VHR SAR images by the traditional image processing method remains problematic. Overcoming this issue is the motivation of the paper.

3. Proposed Methodology for Damage Detection Using VHR SAR Images

3.1. Concept of Damage Detection

Figure 3 gives an ideal example of SAR imaging of a flat roof building. In Figure 3b, the layover, double-bounce line and shadow area can be observed. The shadow caused by the standing building covers the majority of the building's footprint (denoted by the red line). Generally, there is no return from the building or the ground in the shadow area. Therefore, the backscattering intensity in the area is lower than that in the surrounding area.

![Figure 3](image-url) An illustration of SAR imaging of a flat roof building: (a) Relationship between the sensor and the building; (b) building in the image plane.

For further investigation, Figures 4 and 5 show examples of the backscattering range profiles of flat roof and gable roof buildings under different boundary conditions [12]. In Figure 4, the scattering...
effects of flat roof buildings for three different situations can be observed according to the boundary conditions among the building height $h$, width $w$ and incidence angle $\theta$ (Figure 4a–c). The symbols $l$ and $s$ denote the lengths of the layover and shadow in the ground-projected image space, respectively; $b$ is the double bounce caused by the dihedral corner reflector formed by the vertical wall of the building and the surrounding ground; and $e$ represents the shadow area from which there is no return from the building or the ground. In Figure 4, some regularity can be observed. The double bounce always appears as a linear feature corresponding to the building’s front wall, and it can indicate the presence of a building. Near the double bounce, the shadow areas ($e$ in Figure 4b,c) usually appear in the image and cover the area of the building’s footprint.

In Figure 4a, there is backscattering return from the roof between the double bounce and shadow. If the roof is flat, minimal backscattering from the roof is received by the sensor. Thus, $d$ (backscattering from the roof) in Figure 4a usually appears as a dark region in SAR images. Figure 5 shows the scattering effects of a gable roof building. It shows three examples of backscattering profiles from a gable roof building with roof inclination angles $\alpha$ for different incidence angles $\theta$. Near the double-bounce line, the shadow areas ($e$ in Figure 5) also cover the majority of the building’s footprint.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Backscattering range profiles from a simple flat roof building model with width $w$ and different heights $h$. Ground scattering $a$, double bounce $b$, scattering from the vertical wall $c$, backscattering from the roof $d$ and shadow area $e$ are labeled. The gray values in the backscattering profiles correspond to the relative amplitudes. (a) $h < w \cdot \tan(\theta)$; (b) $h = w \cdot \tan(\theta)$; (c) $h > w \cdot \tan(\theta)$ (re-edited from [12]).}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Backscattering range profiles from a gable roof building with roof inclination angles $\alpha$. The legend is similar to that used for the flat roof buildings in Figure 3. The gray values in the backscattering profiles correspond to the relative amplitudes. (a) $\theta < \alpha$; (b) $\theta = \alpha$; (c) $\theta > \alpha$ (re-edited from [12]).}
\end{figure}

Generally, for a flat or gable roof building, no matter how the aspect angle is changed, the building profile incised by the incidence plane of the radar wave will keep its basic shape. Figure 6 shows some examples of the profiles of a flat roof and a gable roof building with different aspect angles. In Figure 6, $\phi_1$–$\phi_4$ show examples of four different aspect angles; $w_1$–$w_4$ are the widths of the four building profiles. The figures show that the building profiles keep the basic shapes, which are similar to the building profiles shown in Figures 4 and 5, if the aspect angle is changed. Therefore, a conclusion can be made...
that when the flat roof and gable roof buildings are illuminated by the radar in different aspect angles, the backscattering range profiles of the buildings can also be explained by Figures 4 and 5.

Figure 6. Illustrations of the building profiles of a flat roof building and a gable roof building with different aspect angles. (a) A flat roof building model illuminated by radar waves in four directions; (b) building profiles corresponding to the four different directions; (c) a gable roof building model illuminated by radar waves in four directions; (d) building profiles corresponding to the four different directions.

Based on the analysis above, we can conclude that if the building is still standing, the region of the building’s footprint in an SAR image usually has low backscattering intensity and is dark compared to surrounding areas under the condition of various incidence angles and aspect angles. However, if the building is totally collapsed, debris piles form. The debris exhibits brighter scattering spots caused by smaller corner reflectors resulting from the composition of different planes [4]. Therefore, if the footprint of a building can be obtained, standing buildings and collapsed buildings can be separated based on features derived from the original footprint of the building. Based on this concept, we proposed a new scheme for detecting totally destroyed buildings with post-event single polarization VHR SAR images (Figure 7).

Figure 7. Block scheme of the proposed methodology.
3.2. Preprocessing

The main inputs of the method are a sub-meter resolution SAR image and a building footprint map. Before analysis, the digital numbers of the TerraSAR-X SAR data were converted into radar sigma naught values as follows [13]:

\[ \sigma = (k_s \times |DN|^2 - \text{NEBN}) \times \sin \theta_{loc} \]  

(1)

where \( k_s \) is the calibration and processor scaling factor given by the parameter calFactor in the annotated file, DN is the pixel intensity value and NEBN is the noise equivalent beta naught. It represents the influences of different noise contributions on the signal. \( \theta_{loc} \) is the local incidence angle [13].

The building footprint map in Figure 7 can be provided by a cadastral map from the government. If there is no building footprint map of the affected area, the map can be obtained by interpreting high-resolution pre-earthquake optical images. Unfortunately, we have no pre-earthquake high-resolution remote sensing images covering the research area. Therefore, we obtained the building footprint map by careful interpretation of the post-earthquake optical image with information of the in situ investigation and LIDAR data of the research area in our experiment. To register the building footprint map, the SAR data should be geometrically rectified. The ground control points, which provide the relationships between pixel coordinates in the image and geo-coordinates, can be found in the XML file in SAR products. After geometric rectification, the SAR image can be matched to the building footprint map, which also has geo-coordinates information. To improve the registration precision, registration between the geo-rectified SAR images and optical image is implemented with ENVI software by manually selecting the control points in the images in our experiment. The SAR images we used are single look complex (SLC) products. For real operation, the geocoded ellipsoid corrected (GEC) or enhanced ellipsoid corrected (EEC) products of the SAR data can be ordered to avoid geometric rectification by users. Although a layover, shadow area of a building usually occurs in the very high resolution SAR image, the footprint indicates the real location of the building. Therefore, with the building footprint map, the original footprints of buildings can be located in the descending and ascending SAR images. Then, features can be extracted based on the footprints in the SAR image.

3.3. Features for Classification

Four first-order statistics and eight second-order image statistical measures are investigated for analysis and damage type classification.

3.3.1. First-Order Statistics

Statistical features including mean, variance, skewness and kurtosis are selected for first-order statistical analysis. Setting \( f_1 = \mu \) as the mean value and \( f_2 = \sigma^2 \) as the variance value, the skewness and kurtosis can be expressed as follows.

(1) Skewness:

\[ f_3 = \frac{E(x - \mu)^3}{\sigma^3} \]  

(2)

In Equation (2), \( E(\cdot) \) represents the expected value of variables. Skewness is a measure of the asymmetry of the data distribution around the sample mean. If skewness is negative, the data are spread out more to the left of the mean than to the right. If skewness is positive, the data are spread out more to the right. The skewness of the normal distribution is zero [14].

(2) Kurtosis:

\[ f_4 = \frac{E(x - \mu)^4}{\sigma^4} \]  

(3)
Kurtosis describes how peaked or flat a distribution is, with reference to the normal distribution. The kurtosis of the normal distribution is 3. If kurtosis has a high value, then it represents a sharp peak distribution in the intensity values about the mean and has a longer, fat tail. If kurtosis has a low value, it represents a flat distribution of intensity values with a short, thin tail [15].

Skewness and kurtosis are features that describe the position and shape of a data distribution. SAR data usually exhibit long-tail distributions, and standing and collapsed buildings show different pixel intensities based on the above analysis. Therefore, skewness and kurtosis are selected as features to describe these differences.

3.3.2. Second-Order Image Statistics

The second-order image statistics measure the relationships between pixels and their neighbors [16]. To calculate the second order statistics, the pixel values for a given scale of interest are first translated into a gray level co-occurrence matrix (GLCM). Texture measures derived from the GLCM have been widely used for land cover classification and other applications with radar and optical data.

GLCM describes the frequency of one gray tone appearing in a specified spatial linear relationship with another gray tone in the area under investigation [17]. The textural statistics are calculated based on GLCM. The GLCM-based textural features are second-order features widely used in texture analysis and image classification. Recently, GLCM features have been used to analyze the signatures of damaged buildings in real and simulated VHR SAR images [8,10,18].

Haralick et al. proposed 14 features based on GLCM for textural analysis [19]. These features measure different aspects of the GLCM, and some features are correlated. In this paper, eight texture features are chosen for analysis: the mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment and correlation. Let \( p(i,j) \) be the \((i,j)\)-th entry in a normalized GLCM. The features are as follows [19–21].

(1) Mean:

\[
g_1 = \sum_i \sum_j i \times p(i,j)
\]

(2) Variance:

\[
g_2 = \sum_i \sum_j (i - \mu)^2 p(i,j)
\]

(3) Homogeneity:

\[
g_3 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i,j)
\]

(4) Contrast:

\[
g_4 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \mid |i-j| = n \right\}
\]

In Equation (7), \( N_g \) is the number of distinct gray levels in the quantized image.

(5) Dissimilarity:

\[
g_5 = \sum_i \sum_j |i - j| \times p(i,j)
\]

(6) Entropy:

\[
g_6 = -\sum_i \sum_j p(i,j) \log(p(i,j))
\]

(7) Second moment:

\[
g_7 = \sum_i \sum_j (p(i,j))^2
\]
Correlation:

\[ g_{ij} = \frac{\sum_i \sum_j (i-j) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \]  

where \( \mu_x, \mu_y, \sigma_x \) and \( \sigma_y \) are the means and standard deviations of \( p_x \) and \( p_y \).

The GLCM mean measures the average gray level in the GLCM window. The GLCM variance measures the gray level variance and is a measure of heterogeneity. GLCM homogeneity, which is also called the inverse difference moment, measures the image homogeneity, assuming larger values for smaller gray tone differences in pair elements. GLCM contrast is a measure of the amount of local variation in pixel values among neighboring pixels. It is the opposite of homogeneity. GLCM contrast is related to first-order statistical contrast. Thus, high contrast values reflect highly contrasting textures. GLCM dissimilarity is similar to contrast and inversely related to homogeneity. GLCM entropy measures the disorder in an image. When the image is not texturally uniform, the value will be very large. The GLCM second moment is also called energy. It measures textural uniformity. If the image patch is homogeneous, the value of the second moment will be close to its maximum, which is equal to 1. GLCM correlation is a measure of the linear dependency of pixel values on those of neighboring pixels. A high correlation value (close to 1) suggests a linear relationship between the gray levels of pixel pairs. Therefore, we can see that the GLCM homogeneity and second moment are measures of homogeneous pixel values across an image, while the GLCM contrast, dissimilarity and variance are measures of heterogeneous pixel values. Based on the analysis of Section 2, we can observe that the standing building shows relative homogeneity in the footprint, while the collapsed building shows more heterogeneity. Therefore, these features are selected for analysis. In addition, the GLCM mean, entropy and correlation are also considered for analysis. In this paper, image texture measures are calculated for every pixel using ENVI software.

### 3.4. Classifier

Three machine learning methods involving random forest (RF) [22], support vector machine (SVM) [23] and K-nearest neighbor (K-NN) [24] classifiers are used to classify collapsed and standing buildings in our experiments.

Ensemble learning algorithms (e.g., random forest) have received increased interest because they more accurately and robustly filter noise compared to single classifiers [22,25,26]. RF is an ensemble of classification trees, where each tree contributes a single vote to the assignment of the most frequent class of the input data. An RF can handle a number of input variables without variable deletion, and it estimates the importance of variables in the classification.

Support vector machine (SVM) is a supervised learning model with associated learning algorithms that analyze data and recognize patterns based on the structural risk minimization principle in statistics.

The K-NN algorithm is a statistical instance-based learning method [27,28]. It decides that a certain pattern X belongs to the same category as its closest neighbors given a training set of \( m \) labeled patterns. The algorithm estimates the probabilities of the classes using a Euclidian metric.

The three machine learning methods are supervised classifiers, and they have been widely used for remote sensing image classification. We adopt the methods to evaluate the results of building damage detection using VHR SAR images.

### 4. Study Area and Materials

The study area is located in old Beichuan County. On 12 May 2008, the Wenchuan Earthquake with a magnitude of 8.0 occurred in Sichuan province, China [29]. The earthquake devastated a huge area in Sichuan province, killing and injuring a large number of people and causing major damage to buildings and other infrastructures. Beichuan County (centered at approximately 31.833°N and 104.459°E) was one of the most affected areas. The Beichuan County was originally located in Qushan Town. After the earthquake, the county was moved to Yongchang Town, which belonged to An County before the
earthquake. The old Beichuan County seat was abandoned, and some ruins were preserved as the Beichuan Earthquake Memorial Park. The study area belongs to the Beichuan Earthquake Memorial Park. Therefore, some damaged buildings are preserved in the area and can be selected for analysis.

Three decimeter level resolution TerraSAR-X ST mode SAR images acquired in 2014 and 2015 are used for analysis and experimentation. One is ascending SAR image, and two are descending SAR images. TerraSAR-X ST SAR is a new mode that was introduced in 2013 by the German Aerospace Center (DLR). With the new mode, a spatial resolution of approximately 0.24 m can be achieved in the azimuth direction of the system by widening the azimuth beam steering angle range [30]. With coarse spatial resolution data, damage detection is usually performed on the building group or residential area level. While in the VHR SAR image, much more building features like edges and point structures become visible. Therefore, it is possible to focus an analysis on individual buildings. Table 1 gives detailed information regarding the data.

<table>
<thead>
<tr>
<th>Number</th>
<th>Polarization</th>
<th>Orbit Direction</th>
<th>Date of Acquisition</th>
<th>Resolution (m)</th>
<th>Look Direction</th>
<th>Incidence Angle (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HH</td>
<td>Ascending</td>
<td>4 December 2014</td>
<td>0.86</td>
<td>0.23</td>
<td>Right</td>
</tr>
<tr>
<td>2</td>
<td>HH</td>
<td>Descending</td>
<td>8 December 2014</td>
<td>0.86</td>
<td>0.23</td>
<td>Right</td>
</tr>
<tr>
<td>3</td>
<td>HH</td>
<td>Descending</td>
<td>8 April 2015</td>
<td>0.86</td>
<td>0.23</td>
<td>Right</td>
</tr>
</tbody>
</table>

An optical multispectral image and LIDAR data covering the study area are used as reference data. In May 2013, 5 years after the earthquake, the Airborne Remote Sensing Center (ARSC) of the Institute of Remote Sensing and Digital Earth carried out a campaign for monitoring the areas affected by the Wenchuan earthquake in 2008. One remote sensing airplane carrying an optical ADS80 sensor collected a number of images of the areas, including the old Beichuan County. The data acquired by ADS80 are multi-spectral images, and the resolution of the image is approximately 0.6 m. The optical image has been rectified when we obtained it from ARSC. In situ investigations were carried out in 2014 and 2015, corresponding to the ST SAR data acquisition dates. Based on the post-earthquake optical image, LIDAR data and in situ investigation, a reference map of the standing and collapsed buildings can be generated by careful interpretation. The reference map of the standing and collapsed building can be regarded as the ground truth for analysis and validation. Figure 8 illustrates the research area, the SAR image and building footprint map. Figure 9 presents the optical image (Figure 9a) and the reference map (Figure 9b). Figure 9c,d illustrates a descending and ascending SAR image superimposed by the reference map, respectively. In order to illustrate the image more clearly, the values of the SAR image in Figure 9c,d have been stretched to the 0–255 gray level. The building footprint map and reference map are generated by optical image, LIDAR data and in situ investigations.

![Figure 8. The study area: (a) a post-earthquake descending SAR image (the azimuth direction is from top to bottom, and the range direction is from right to left); and (b) building footprint map.](image-url)
Figure 9. Reference data: (a) optical image acquired in 2013; (b) reference map; (c) a post-earthquake descending SAR image superimposed by the reference map (the azimuth direction of the SAR image is from top to bottom, and the range direction is from right to left); (d) a post-earthquake ascending SAR image superimposed by the reference map (the azimuth direction of the SAR image is from bottom to top, and the range direction is from left to right). The SAR images have been stretched so as to be illustrated more clearly.

5. Experimental Results and Analysis

Three sub-meter VHR TerraSAR-X ST SAR images are used in the experiments and algorithm evaluation. With the building footprint map, the original footprints of the buildings can be obtained in the SAR images. Then, the features are calculated, forming a feature vector \( FV = \{f_1, f_2, f_3, g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8\} \) for classification.

5.1. Parameters of GLCM-Based Textural Features

Generally, if the textural measures derived from the GLCM are used, some fundamental parameters should be defined, including the quantization levels of the image, the window size used to calculate the GLCM and the displacement and orientation values of the measurements.
5.1.1. Quantization Level

The quantization level is a fundamental parameter. A value that is too low can lead to substantial information reduction. However, fewer levels reduce noise-induced effects and directly affect the computing cost because the number of bins determines the size of the co-occurrence matrix [7]. This tradeoff was discussed by Soh et al. [20], with the conclusion that 64 levels are suggested for sea ice in SAR images. Additionally, they have that a 256-level representation is not necessary and an eight-level representation is undesirable. A level of 256 was applied in [16]; a level of 128 was adopted in [7]; and a 64-level quantization was applied in [8] to detect and distinguish debris in SAR images. Additionally, 16-level, 64-level and 32-level quantization was adopted in [31–33], respectively, for sea ice classification and other applications. Based on [20] and [7,8], a 64-level quantization using ENVI software is adopted in our experiments. Figure 10 illustrates the histograms of the sigma naught of the SAR images (Figure 10a) and the 64-level quantization images by ENVI (Figure 10b). It shows that the shapes of the histograms of the images are kept well after quantization.

![Figure 10. Histograms of (a) sigma naught of the SAR image and (b) 64-level quantization image by ENVI.](#)

5.1.2. Window Size

In order to calculate texture features for each pixel in an image, a moving window is usually used to define the neighborhood of a pixel, and the texture measurement calculated from the window is assigned to the center pixel [26,34]. The window size for texture analysis is usually related to image resolution and the contents within the image. However, it might be difficult to determine the window size through a formula. In the experiments, we chose 13 × 13 pixels as the window size to calculate the texture features. The analysis is as follows.

1) Estimation by classification result:

The overall accuracy or kappa coefficient of the classification results is usually applied for estimating the window size [26,34]. If the texture features can give the highest classification accuracy, then the window size will be chosen as the optimal size. Figure 11 shows the influence of window size on the overall accuracy and accuracy of the classification results of texture features for collapsed and standing buildings. The classifier is random forest. With the whole of 192 samples (36 collapsed buildings, 156 standing buildings), 1/3 of samples are for training, and 2/3 of samples are for testing. The figure shows that the classification results of small window size are not stable. The overall accuracy increases when the window size is larger than 7 × 7 pixels and decreases when the window size is larger than 15 × 15 pixels. The overall accuracy is stable when the window size is between 9 × 9 pixels and 15 × 15 pixels. It appears that the window sizes of 13 × 13 pixels and 15 × 15 pixels can be identified as appropriate to achieve best overall accuracy.
Figure 11. Influence of window size on the overall accuracy and the accuracy of the classification results of texture features for collapsed and standing buildings.

(2) Estimation by coefficient of variation of texture measures:

To facilitate the choice of an optimal texture window, the coefficient of variation of each texture measure for each class in relation to window size was also calculated [35]. The chosen optimal window size is usually that in which the value of the coefficient of variation starts to stabilize to the smallest value [35]. Figure 12 presents the evolution of the variance in relation to different window sizes for the textural measures. Figure 12a–c is the results of the second moment, homogeneity and entropy, respectively. The red starred line indicates the collapsed building, and the blue open-circled line represents the standing building values. The variances of textural features of the second moment and homogeneity decrease at larger window sizes; while the values of entropy show general increasing trends at small window sizes and decrease with increasing window size. The variation starts to stabilize at a $13 \times 13$ pixels window in Figure 12a–c. The results show that a $13 \times 13$ pixels window size can be a good choice for the textural measures.

Figure 12. Evolution of the variation coefficient in relation to different window sizes for textural measures of the (a) second moment; (b) homogeneity; and (c) entropy.

(3) Estimation by visual interpretation:

If the building is still standing, the footprint is usually covered by shadow area. If the building is collapsed, the footprint is usually covered by debris. Figure 13 shows examples of shadow and debris in building footprints in the SAR image. The red and yellow rectangles indicate the footprints of standing and collapsed buildings, respectively. The debris usually contains some heaps of randomly-oriented planes mainly made of concrete. The planes are usually not very large and may be 1–3 m in width or length. After geometric rectification, the SAR image is resampled to a 0.6-pixel spacing, which is the same as the optical image. Therefore, in the 0.6-pixel spacing SAR image, the planes may be usually 2–5 pixels (Figure 13). Consequently, we think the window size of $13 \times 13$ pixels may be proper for texture analysis.
Figure 13. Examples of shadows of standing buildings (red rectangle) and collapsed building (yellow rectangle) debris in building footprints in the SAR image.

Generally, the larger the window size is, the coarser the information that can be provided by textural features. Therefore, the window size used to compute textural features should not be too large. Based on the quantitative and qualitative analysis, a $13 \times 13$ pixels window size is chosen to calculate the GLCMs in the experiments.

5.1.3. Orientation and Displacement

The orientation parameter is less important compared to other factors in the co-occurrence matrix [20]. We investigate the values of the textural measures of collapsed (red starred line) and standing buildings (blue open-circled line) in four directions, including $0^\circ$, $45^\circ$, $90^\circ$ and $135^\circ$. Figure 14 gives the results of the values of GLCM-based textural features with different orientations. It shows that the features have relatively stable values when the orientation changes. However, we also found that the texture features of $45^\circ$ and $135^\circ$ have similar values. Additionally, they have some differences from the values in $0^\circ$ and $90^\circ$. In Figure 9, it can be observed that most of the buildings’ orientations are about $135^\circ$ or $45^\circ$. This may be the reason that the values of features in $45^\circ$ and $135^\circ$ have a similar tendency. In the experiments, the orientation of $135^\circ$ is chosen for the calculation of GLCM.

![Figure 14. GLCM features with different orientations.](image)

(a) Mean of second moment; (b) Mean of homogeneity; (c) Mean of entropy.

The GLCM describes the probability of finding two given pixel gray-level values in a defined relative position in an image [36]. Displacement indicates the inter-pixel distance. Figure 15 presents the investigation of GLCM features with different displacements. In the experiment, the features are calculated at an orientation of $135^\circ$, with a $13 \times 13$ window size. The results show that the values changed little when the displacement increased, indicating that the parameter displacement is not important in the co-occurrence matrix. The ideal distance of pixel offset depends on the level of detail of the texture that is analyzed. The GLCM cannot capture fine-scale textual information if a large pixel offset is used [7]. Accordingly, in our experiment, an orientation of $135^\circ$ and a displacement of one are selected.
5.2. Classification Results

5.2.1. Features for Classification

Classification experiments had been carried out for features analysis. Random forest (RF) is the classifier. Based on the reference map, we obtained original footprints of 12 totally collapsed buildings and 52 standing buildings (Figure 9b) in each image. For each test, 6 collapsed buildings and 17 standing buildings are used to train samples, and the corresponding numbers of test samples are six and 35, respectively. Table 2 shows the classification results with only first-order statistics. Table 3 is the classification results with only second-order statistics. Table 4 gives the classification results with all features.

Table 2. Classification results with only first-order statistics (CO = totally collapsed building, ST = standing building; Pa = producer’s accuracy, Ua = user’s accuracy).

<table>
<thead>
<tr>
<th>Type</th>
<th>Classification</th>
<th>8 April 2015</th>
<th>4 December 2014</th>
<th>8 December 2014</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CO  ST  Pa</td>
<td>CO  ST  Pa</td>
<td>CO  ST  Pa</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td>5  1  83.3%</td>
<td>4  2  66.7%</td>
<td>6  0  100%</td>
<td></td>
</tr>
<tr>
<td>ST</td>
<td>6  29  82.8%</td>
<td>6  29  82.8%</td>
<td>3  32  91.4%</td>
<td></td>
</tr>
<tr>
<td>Ua</td>
<td>45.5%  96.7%</td>
<td>40%  93.5%</td>
<td>66.7%  100%</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>82.9%</td>
<td>80.5%</td>
<td>92.7%</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Classification results with only second-order statistics (CO = totally collapsed building, ST = standing building; Pa = producer’s accuracy, Ua = user’s accuracy).

<table>
<thead>
<tr>
<th>Type</th>
<th>Classification</th>
<th>8 April 2015</th>
<th>4 December 2014</th>
<th>8 December 2014</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CO  ST  Pa</td>
<td>CO  ST  Pa</td>
<td>CO  ST  Pa</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td>5  1  83.3%</td>
<td>5  1  83.3%</td>
<td>5  1  83.3%</td>
<td></td>
</tr>
<tr>
<td>ST</td>
<td>3  32  91.4%</td>
<td>5  30  85.7%</td>
<td>6  29  82.9%</td>
<td></td>
</tr>
<tr>
<td>Ua</td>
<td>62.5%  97.0%</td>
<td>50%  96.8%</td>
<td>45.5%  96.7%</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>90.2%</td>
<td>85.4%</td>
<td>82.9%</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Classification results with all features (CO = totally collapsed building, ST = standing building; Pa = producer’s accuracy, Ua = user’s accuracy).

<table>
<thead>
<tr>
<th>Type</th>
<th>Classification</th>
<th>8 April 2015</th>
<th>4 December 2014</th>
<th>8 December 2014</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CO</td>
<td>ST</td>
<td>Pa</td>
<td>CO</td>
</tr>
<tr>
<td>CO</td>
<td>6</td>
<td>0</td>
<td>100%</td>
<td>5</td>
</tr>
<tr>
<td>ST</td>
<td>5</td>
<td>30</td>
<td>85.7%</td>
<td>3</td>
</tr>
<tr>
<td>Ua</td>
<td>54.5%</td>
<td>100%</td>
<td>62.5%</td>
<td>97.0%</td>
</tr>
<tr>
<td>Overall</td>
<td>87.8%</td>
<td>90.2%</td>
<td>90.2%</td>
<td></td>
</tr>
</tbody>
</table>

Figure 16 illustrates the overall accuracy of the classification results of different images with different features. The red line in Figure 16 is the results with only first-order statistics of the images. The green line indicates the results with only second-order statistics of the images. The blue line is the results with all features of the images. It shows that the classification result with all features is good and stable for all images, and the overall accuracies are all above 85%. However, the results are relatively unstable with only first-order statistics and only second-order statistics. As mentioned in Section 3.3, the mean and variance values can describe the difference of backscattering of the standing and collapsed buildings. The skewness and kurtosis reflect the difference of shapes of histograms of the standing and collapsed buildings. The textural measures can reveal textural characteristics of the two kinds of buildings. Therefore, all features are recommended for classification due to the good ability to discriminate the standing and collapsed buildings.

5.2.2. Classification Results

Based on the reference map, we obtained original footprints of 12 totally collapsed buildings and 52 standing buildings (Figure 8b) in each image. For each classifier, six collapsed buildings and 17 standing buildings are used to train samples, and the corresponding numbers of test samples are six and 35, respectively. In the experiments, 10, 20, 50, 100 and 200 trees are used for the RF classifier, and the classification result with the highest accuracy is recorded and reported. For the SVM classifier, the C-support vector classification (C-SVC), which was introduced by Cortes et al. [37] and presented by Chang and Lin [23], is adopted. As suggested by Chang, the radial basis function (RBF) is chosen as the kernel function to construct the classifier. Furthermore, the two parameters (the kernel width and penalty coefficient) of the RBF are determined by performing a grid search and cross-validation method in our study.

Tables 5–7 present the classification results of different classifiers in the three images. For Image 1, Table 5 shows that the collapsed buildings can be classified well by the three classifiers. Only one
collapsed building was misclassified as a standing building by the RF. The K-NN classifier gives the lowest overall accuracy of 80.5%, as there are six standing buildings misclassified as collapsed buildings. The overall accuracies of RF and SVM reach 90.2%. For Image 2, Table 6 shows that RF and SVM yield the same results, with overall accuracies of 90.2%. The overall accuracy of K-NN is 87.8%. For Image 3, Table 7 shows that SVM gives the best result, with an overall accuracy of 92.7%. The overall accuracies of RF and K-NN are 87.8% and 82.9%, respectively.

### Table 5. Classification results of Image 1 using different methods (CO = totally collapsed building, ST = standing building; Pa = producer’s accuracy, Ua = user’s accuracy).

<table>
<thead>
<tr>
<th>Type</th>
<th>Classification</th>
<th>RF</th>
<th>SVM</th>
<th>K-NN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CO</td>
<td>ST</td>
<td>Pa</td>
</tr>
<tr>
<td>CO</td>
<td>5</td>
<td>1</td>
<td>83.3%</td>
<td>6</td>
</tr>
<tr>
<td>ST</td>
<td>3</td>
<td>32</td>
<td>91.4%</td>
<td>4</td>
</tr>
<tr>
<td>Ua</td>
<td>62.5%</td>
<td>97.0%</td>
<td>60%</td>
<td>100%</td>
</tr>
<tr>
<td>Overall</td>
<td>90.2%</td>
<td>90.2%</td>
<td>80.5%</td>
<td></td>
</tr>
</tbody>
</table>

### Table 6. Classification results of Image 2 using different methods (CO = totally collapsed building, ST = standing building; Pa = producer’s accuracy, Ua = user’s accuracy).

<table>
<thead>
<tr>
<th>Type</th>
<th>Classification</th>
<th>RF</th>
<th>SVM</th>
<th>K-NN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CO</td>
<td>ST</td>
<td>Pa</td>
</tr>
<tr>
<td>CO</td>
<td>6</td>
<td>0</td>
<td>100%</td>
<td>6</td>
</tr>
<tr>
<td>ST</td>
<td>4</td>
<td>31</td>
<td>88.6%</td>
<td>4</td>
</tr>
<tr>
<td>Ua</td>
<td>60%</td>
<td>100%</td>
<td>60%</td>
<td>100%</td>
</tr>
<tr>
<td>Overall</td>
<td>90.2%</td>
<td>90.2%</td>
<td>87.8%</td>
<td></td>
</tr>
</tbody>
</table>

### Table 7. Classification results of Image 3 using different methods (CO = totally collapsed building, ST = standing building; Pa = producer’s accuracy, Ua = user’s accuracy).

<table>
<thead>
<tr>
<th>Type</th>
<th>Classification</th>
<th>RF</th>
<th>SVM</th>
<th>K-NN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CO</td>
<td>ST</td>
<td>Pa</td>
</tr>
<tr>
<td>CO</td>
<td>6</td>
<td>0</td>
<td>100%</td>
<td>6</td>
</tr>
<tr>
<td>ST</td>
<td>5</td>
<td>30</td>
<td>85.7%</td>
<td>3</td>
</tr>
<tr>
<td>Ua</td>
<td>54.5%</td>
<td>100%</td>
<td>66.7%</td>
<td>100%</td>
</tr>
<tr>
<td>Overall</td>
<td>87.8%</td>
<td>92.7%</td>
<td>82.9%</td>
<td></td>
</tr>
</tbody>
</table>

Based on the analysis above, collapsed buildings and standing buildings can be effectively separated by the three classifiers, as the overall accuracies of the three classifiers all exceed 80%. The results of RF and SVM are similar, and these methods outperform the K-NN classifier. Thus, collapsed buildings can be detected well by the proposed method.

5.2.3. Further Analysis about the Classification

In our experiments, the classifiers adopted are supervised classifiers. Furthermore, they are also machine learning techniques. The training dataset is very important for these classifiers. To obtain the training samples, the ground truth must be known. However, after the training dataset has been built, the classifier can be performed when a new SAR image is input. Table 8 shows the results of
three tests by the RF classifier. The aim of the tests is to evaluate the ability to classify a new image based on an existing training dataset. In Test 1 and Test 2, the training dataset is samples from Image 3 (8 April 2015) and Image 2 (8 December 2014), respectively. In Test 3, the training dataset is all samples from Images 2 and 3. Image 1 (4 December 2014) is the input image for the classification of all of the tests. As mentioned in Table 1, Images 2 and 3 are descending images, and Image 1 is an ascending image. In Test 1, the accuracy of the collapsed building is not good. However, the result of Test 2 is acceptable. In Test 3, the results are stable and improved slightly. The results indicate that it is possible to classify a new image if there is an existing training dataset for the classifier. The result will be better if the training dataset contains more training samples. In future work, more samples should be acquired, and images of other research sites should be tested for a more precise conclusion.

Table 8. RF classification results of different tests (CO = totally collapsed building, ST = standing building, Pa = producer’s accuracy, Ua = user’s accuracy. In Test 1 and Test 2, the training dataset is samples from Image 3 (8 April 2015) and Image 2 (8 December 2014), respectively. In Test 3, the training dataset is all samples from Images 2 and 3. Image 1 (4 December 2014) is the input for classification of all of the tests).

<table>
<thead>
<tr>
<th>Type</th>
<th>Classification</th>
<th>Test 1</th>
<th>Test 2</th>
<th>Test 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CO  ST  Pa</td>
<td>CO  ST  Pa</td>
<td></td>
<td>CO  ST  Pa</td>
</tr>
<tr>
<td>CO</td>
<td>6    6    50.0%</td>
<td>8    4    66.7%</td>
<td></td>
<td>8    4    66.7%</td>
</tr>
<tr>
<td>ST</td>
<td>1    51   98.1%</td>
<td>4    48   92.3%</td>
<td></td>
<td>3    49   94.2%</td>
</tr>
<tr>
<td>Pa</td>
<td>85.7% 89.5% 66.7%</td>
<td>92.3% 72.3% 92.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>89.1% 87.5%</td>
<td>89.1%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. Discussion

Compared with low- or medium-resolution SAR images, VHR SAR image can provide more detailed information of man-made objects, such as buildings. In 2013, the TerraSAR-X mission was extended by implementing two new modes including the staring spotlight. The azimuth resolution of this mode is significantly increased to approximately 0.24 m by widening the azimuth beam steering angle range [30]. The sub-meter VHR SAR data source provides new opportunities for earthquake damage mapping and makes it possible to focus an analysis on individual buildings. A new method is proposed in the paper to use these VHR SAR data for collapsed building detection. The method is based on the concept that the original footprints of collapsed and standing buildings show different features in sub-meter VHR SAR images. In an SAR image, the shadow of a standing building covers the majority of the building’s footprint; thus, the footprint of a standing building usually exhibits low backscattering intensity. For a collapsed building, debris piles up above the footprint. Some bright spots can be observed, which are caused by corner reflectors resulting from the composition of different planes. Therefore, the collapsed building displays different features than do standing buildings in their original footprints. For the proposed method, the features for classification are derived from building footprints. Therefore, the high image resolution is the key factor to ensure that the building footprint contains more pixels for classification. If the building size is large, it also will benefit building damage detection.

The proposed method avoids the difficulties of finding exact edges for building damage detection. In VHR SAR images, more features, such as the points and edges of objects, become visible. Debris exhibits similar features as the surroundings (e.g., vegetation). It is not easy to identify debris from surroundings such as vegetation with a single SAR image. To solve this problem, the proposed method uses a building footprint map to locate the original footprints of buildings. The key algorithm begins with the footprint of a building in an SAR image. The footprint of a building can be provided by cadastral maps or directly extracted from a pre-event SAR image or other VHR optical data.
The approach shows a good ability for isolated buildings. However, if the buildings are too close, the footprint of a low building may be affected by a nearby tall building’s layover in the image. Because the layover usually shows relatively high backscattering intensity, the low standing building may be misclassified as a collapsed building by the method. If the debris of a collapsed building is near a standing building, the incidence wave from the sensor may be obstructed by the standing building, as little of the backscatter wave from the collapsed building can be received by the sensor. In this case, the collapsed building also shows low backscattering intensity in the footprint, and it may be misclassified as a standing building.

The high resolution of an SAR image usually means small coverage of a single scene. The scene extent of a TerraSAR-X ST mode image usually varies between 2.5 and 2.8 km in the azimuth direction and 4.6 km and 7.5 km in the range [30,38]. The small coverage of ST mode data may not be adequate for damage detection over a wide area with a single image. Commercial TerraSAR-X services commenced in January 2008. A TerraSAR-X rebuild was launched as TanDEM-X in 2010 to fly in close formation with TerraSAR-X [39]. The almost identical Spanish PAZ satellite will be launched in 2016 [40] and added to the TerraSAR-X reference orbit. The three almost identical satellites will be operated collectively and deliver optimized revisit times, increase coverages and improve services [41]. When we monitor Earth surface deformation by earthquakes, SAR images with large coverages are preferable. When detecting building damage in a town or a city, the coverage requirements of the image may be smaller. If the study area is a large city, the data can be acquired from TerraSAR-X, TanDEM-X and PAZ to cover a large area. Therefore, in real applications, we can use large-coverage SAR images to locate the city or town influenced by an earthquake and perform preliminary damage assessments. The ST data from TerraSAR-X, TanDEM-X and PAZ can be obtained to monitor the region for a more detailed and accurate assessment. In addition, airborne VHR SAR images covering the study area can be used to detect building damage.

7. Conclusions

In this paper, we present a new damage assessment method for buildings using single post-earthquake sub-meter resolution VHR SAR images and original building footprint maps. The method can work at the individual building level and determines whether a building is destroyed after an earthquake or is still standing. First, a building footprint map covering the study area is obtained as prior knowledge. Then, an SAR image is geometrically rectified by the ground control points provided by the SAR product files. After rectification, the SAR image can be registered as the building footprint map. Thus, with the building footprint map, the original footprint of a building can be located in the SAR image. Then, features can be extracted in the image patch of a building’s footprint to form a feature vector. Finally, the buildings can be classified into damage classes with classifiers.

We demonstrated the effectiveness of the proposed approach using spaceborne post-event sub-meter VHR TerraSAR-X ST mode data from old Beichuan County, China, which was heavily damaged in the Wenchuan earthquake in May 2008. The results show that the method is able to distinguish between collapsed and standing buildings, with high overall accuracies of approximately 90% for the RF and SVM classifiers. We tested the method using both ascending and descending data, demonstrating the effectiveness of the proposed method. The results also show that the first-order statistics and second-order image features derived from building footprints have a good ability to discriminate standing and collapsed buildings. Furthermore, for the supervised classifiers in our experiments, after the training dataset has been built, the classifier can be well performed when a new SAR image is input. In future work, more samples should be acquired, and images of other research sites should be tested for a more precise conclusion.
Acknowledgments: This work was supported in part by the National Natural Science Foundation of China under Grant No. 41371413 and by the Key Program of the National Natural Science Foundation of China under Grant No. 41331176. The authors would like to thank the German Aerospace Center for providing the TerraSAR-X ST data from the TerraSAR-X AO project (LAN2456). The airborne optical image and LIDAR data are provided by the Airborne Remote Sensing Center of the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences.

Author Contributions: Chao Wang and Hong Zhang conceived of and designed the in situ investigations. Lixia Gong, Fan Wu and Qiang Li performed the experiments and analyzed the data. Jingfa Zhang and Qiang Li provided some in situ investigation photographs and data. Lixia Gong and Fan Wu wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

Remote Sens. 2016, 8, 887

26. Feng, Q.; Liu, J.; Gong, J. UAV remote sensing for urban vegetation mapping using random forest and texture analysis. Remote Sens. 2015, 7, 1074–1094. [CrossRef]
Building Earthquake Damage Information Extraction from a Single Post-Earthquake PolSAR Image

Wei Zhai 1,2,3,4, Huanfeng Shen 5, Chunlin Huang 2,4,* and Wansheng Pei 6

1 Gansu Earthquake Administration, Lanzhou 730000, China; zwxzzzdsyhq@163.com
2 Key Laboratory of Remote Sensing of Gansu Province, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou 730000, China
3 College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China
4 Heihe Remote Sensing Experimental Research Station, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou 730000, China
5 School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China; shenhf@whu.edu.cn
6 Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou 730000, China; peiwansheng@126.com

* Correspondence: huangcl@lzb.ac.cn; Tel.: +86-0931-496-7975

Academic Editors: Zhenhong Li, Roberto Tomas, Zhong Lu and Prasad S. Thenkabail

Received: 25 November 2015; Accepted: 16 February 2016; Published: 25 February 2016

Abstract: After an earthquake, rapidly and accurately obtaining building damage information can help to effectively guide the implementation of the emergency rescue and can reduce disaster losses and casualties. Using a single post-earthquake fully-polarimetric synthetic aperture radar (PolSAR) image to interpret building damage information not only involves a guaranteed data source but is also easy and can be rapidly implemented. This paper is focused on rapid building earthquake damage detection in urban areas using post-earthquake PolSAR data. In PolSAR images, the undamaged buildings parallel to satellite flight pass are different from the collapsed buildings, but the undamaged buildings divergent to satellite flight pass are very similar to collapsed buildings because of their volume scattering characteristics. In this paper, the method of polarization orientation angle (POA) compensation is employed to increase the scattering power of buildings divergent to satellite flight pass, and then Wishart supervised classification is implemented on the PolSAR data after POA compensation. In addition, the two parameters of normalized difference of the dihedral component (NDDC) and $\rho_{HHHV}$ are proposed to improve the classification accuracy of the Wishart supervised classification, and both the undamaged buildings and collapsed buildings are determined. The study was carried out after the “4.14” Yushu earthquake in Yushu County, Qinghai province, China. The three damage levels are set for the urban area at the city block scale according to the values of the BBCR building damage index. The experimental results confirm that the scheme proposed in this paper can greatly improve the accuracy of the extraction of building damage information.

Keywords: earthquake; buildings; damage assessment; PolSAR

1. Introduction

In recent years, earthquake disasters have become more frequent around the world. The earthquake is one of the most dangerous natural disasters for human beings, and tens of billions of dollars in property loss are caused by earthquakes every year. Unfortunately, earthquakes cannot be predicted accurately at the current scientific level. Rapid and accurate damage assessment can help to reduce the disaster loss and can provide decision support for the rescue and reconstruction efforts. Buildings are the places where people live, and most of the casualties and economic losses in an earthquake are...
caused by the damage to buildings [1]. Therefore, building damage assessment is one of the most important parts of earthquake damage assessment.

The earthquake damage information obtained from a ground survey is the most accurate, but this process is inefficient and takes a long period of time. Remote sensing, which is characterized by wide coverage and speediness, is very suitable for areas with poor transport infrastructure and where there is a risk of secondary disasters. Optical remote sensing images allow easy interpretation, but they are susceptible to illumination variation [2]. Radar, with its strong penetrating ability, can operate day and night, independent of weather conditions. As a result, radar has become an important means of disaster assessment [3–6], but most of the studies of disaster assessment are based on multi-source [7,8] or multi-temporal data [9–12]. However, obtaining the matched pre-earthquake data is difficult in some situations, and the registration of the pre- and post-earthquake data is tricky and time-consuming. Therefore, it is quicker and more convenient to undertake earthquake damage assessment using only post-event single-temporal data.

Balz [13], Dell’Acqua et al. [14], and Polli et al. [15] have all evaluated building earthquake damage using only post-earthquake single-polarization synthetic aperture radar (SAR) data. Nevertheless, PolSAR (fully-polarimetric SAR) data record the scattering amplitude and phase of the HH (horizontal/horizontal polarization), HV (horizontal/vertical polarization), VH (vertical/horizontal polarization), and VV (vertical/vertical polarization) polarizations four ways for ground objects, and can better assist with the understanding of scattering mechanisms [16] than single-polarization SAR imagery. As a result, building damage assessment using PolSAR imagery is more accurate and more reliable. Guo et al. [17] and Li et al. [18] introduced the $H-\alpha-\rho$ method to extract the spatial distribution of collapsed buildings in the Yushu urban area by using only a single post-earthquake SAR image. Subsequently, Zhao et al. [19] improved the $H-\alpha-\rho$ method and replaced the parameter of $\rho$ with the normalized circular-polar correlation coefficient (NCCC), and, at the same time, the homogeneity texture feature was employed to solve the problem of collapsed buildings and buildings divergent to satellite flight pass being mixed with each other. Shen et al. [20] extracted collapsed buildings based on feature template matching, using 13 polarimetric features. On account of the present research into building earthquake damage information extraction being rather limited, this work aims to undertake some new research in this area.

A new scheme for earthquake damage assessment using only a single post-earthquake PolSAR image is proposed in this study. This work explores the potential of using polarimetric information to estimate earthquake damage for urban regions. In full PolSAR imagery, the scattering power of collapsed buildings characterized by volume scattering is weak, and the undamaged buildings are mainly characterized by double-bounce scattering, for which the scattering power is strong. However, the buildings divergent to satellite flight pass, which are not parallel to the flight pass, with significant cross-polarization backscattering, are similar to the collapsed buildings. This ambiguity between the building types commonly results in overestimation of collapsed buildings in damage assessment.

The buildings divergent to satellite flight pass, rather than the buildings parallel to satellite flight pass, rotate the polarization basis and induce a polarization orientation angle (POA) shift from zero [21]. Therefore, we first implement POA compensation for the original PolSAR data in order to solve the scattering mechanism ambiguity between collapsed buildings and buildings divergent to satellite flight pass. Wishart supervised classification is then performed on the PolSAR data after POA compensation to extract the initial earthquake damage information. The parameters of the normalized difference of the dihedral component (NDDC) and the HH-HV correlation coefficient ($\rho_{HHHV}$) are then introduced to extract the buildings divergent to satellite flight pass, which are added to the undamaged buildings generated from the Wishart supervised classification. The $\rho_{HHHV}$ parameter is also used to improve the vegetation classification result of the Wishart supervised classification. When the undamaged buildings and collapsed buildings are acquired, the building collapse rate is quantized at the block level by the building block collapse rate (BBCR). Finally, a map with three building damage levels for the whole urban region is drawn according to the threshold value of the BBCR at the block scale.
2. Methodology

2.1. The Damage Assessment Procedures

There are five key procedures in the process flow of the damage estimation framework proposed in this study, as shown in Figure 1. Firstly, according to Section 2.2, the method of POA compensation is performed using the PolSAR data after preprocessing, and the new $[T3]$ matrix after POA compensation is obtained.

Secondly, the Wishart supervised classification is performed on the PolSAR data after POA compensation, and this procedure classifies the ground objects into the four classes of undamaged buildings, collapsed buildings, vegetation, and bare areas.

Thirdly, Yamaguchi four-component decomposition is performed on the PolSAR data before and after POA compensation, respectively. At the same time, the two dihedral components are respectively extracted to compute the NDDC, as described in Section 2.4. Next, the $\rho_{HHHV}$ parameter described in Section 2.4 is computed. The buildings divergent to satellite flight pass are then extracted using the two parameters of the NDDC and $\rho_{HHHV}$. The data items of buildings divergent to satellite flight pass are added to the undamaged buildings generated from the Wishart supervised classification, and they become the total undamaged buildings.

According to the third step, the undamaged buildings are extracted. The collapsed buildings are extracted in the fourth step. The vegetation class generated from the Wishart supervised classification is corrected by meeting the condition of NDDC < $\epsilon$. The final output of the class of bare areas is the same as the classification result of the Wishart supervised classification. After the three classes of
undamaged buildings, vegetation, and bare areas are determined, the remaining data items are the collapsed buildings.

Finally, the building collapse rate of each block is derived from the BBCR index described in Section 2.6, and the damage levels of all the blocks are divided into three levels according to the threshold values of the BBCR index. The earthquake damage assessment with three damage levels is then mapped out for the whole urban region. The methodology and parameters are introduced in detail in the next section.

2.2. Polarization Orientation Angle (POA) Compensation

For enhancing the contrast between buildings divergent to satellite flight pass and collapsed buildings, the scheme of POA compensation can be used to increase the double-bounce scattering power of buildings divergent to satellite flight pass.

The buildings divergent to satellite flight pass patch can induce the POA shift $\theta$, which can be estimated by Equation (1) based on the circular polarization method [21,22]:

$$\theta = \begin{cases} 
\theta_0, & \text{if } \theta_0 \leq \frac{\pi}{4} \\
\theta_0 - \frac{\pi}{2}, & \text{if } \theta_0 > \frac{\pi}{4}
\end{cases}$$  

(1)

where

$$\theta_0 = \frac{1}{4} \left[ \text{Arg} \left( \langle S_{RR}S_{LL}^* \rangle \right) + \pi \right]$$  

(2)

According to Lee [21], the data compensation of the orientation angle of $\theta$ can be achieved by:

$$T_\theta = R(\theta) T R(\theta)^T$$  

(3)

where the superscript $T$ denotes the matrix transpose, and the rotation matrix $R(\theta)$ is given by:

$$R(\theta) = \begin{bmatrix}
1 & 0 & 0 \\
0 & \cos 2\theta & \sin 2\theta \\
0 & -\sin 2\theta & \cos 2\theta
\end{bmatrix}$$  

(4)

2.3. Wishart Supervised Classification

The scattering power of PolSAR imagery can be greatly increased using the method of POA compensation. Therefore, in order to extract the undamaged buildings as completely as possible, Wishart supervised classification based on the complex Wishart distribution of the polarimetric coherency matrix [23] is performed on the PolSAR data after POA compensation. The classification algorithm proposed in [24] for polarimetric SAR images is the recommended method for supervised classification. Details of the Wishart supervised classification algorithm can be found in [23].

The ground objects are classified into four categories using the Wishart supervised classification: undamaged buildings, collapsed buildings, vegetation, and bare areas. Among the classification results, the buildings divergent to satellite flight pass will be mixed in both the collapsed buildings and vegetation classes, so the two classes obtained from the Wishart supervised classification will be inaccurate. The undamaged buildings class still lack some of the buildings divergent to satellite flight pass whose scattering power is not strong enough. However, the bare areas class has a high reliability. Therefore, the results of the Wishart supervised classification are considered as the initial classification results for extracting the building damage information. The initial extraction results are then improved using the following two indicators.
2.4. Building Divergent to Satellite Flight Extraction

Because the results of the Wishart supervised classification are not very accurate, the two parameters of the NDDC and $\rho_{HHHV}$ are proposed to correct the initial classification results. The buildings divergent to satellite flight pass are extracted using the two parameters of the NDDC and $\rho_{HHHV}$, which are introduced in the following.

The indicator of the difference of the dihedral component (DDC) is defined as the difference between the dihedral component obtained from the Yamaguchi four-component decomposition \cite{25,26} before and after POA compensation. It can be expressed as: the DDC equals the dihedral component after POA compensation minus the dihedral component before POA compensation. The DDC is normalized to a positive value range, which is named the normalized difference of the dihedral component, or the NDDC for short.

The scattering power of the buildings divergent to satellite flight pass is greatly increased after the POA compensation. Meanwhile, the scattering intensity of the dihedral component generated from the Yamaguchi four-component decomposition is also increased. The NDDC can measure the change in the double-bounce scattering power after the POA compensation. The double-bounce scattering power of the buildings divergent to satellite flight pass changes a great deal after the POA compensation, while that of the targets with reflection symmetry changes little. That is, the NDDC values of the buildings divergent to satellite flight pass are high and those of the targets with reflection symmetry are low. Hence, the NDDC can be introduced in the process of earthquake damage assessment to find the buildings divergent to satellite flight pass, which can be used to correct the classification result of the undamaged buildings generated from the Wishart supervised classification. The scattering intensity of some buildings divergent to satellite flight pass is not increased to as strong as the buildings parallel to satellite flight pass using the method of POA compensation. Therefore, the class of undamaged buildings obtained from the Wishart supervised classification is not complete. The main missing undamaged buildings are the buildings divergent to satellite flight pass whose scattering intensity is not significantly increased. The data items with high NDDC values correspond to the buildings divergent to satellite flight pass, which can be added to the undamaged buildings. The threshold value $\epsilon$ of the NDDC is set to distinguish the buildings divergent to satellite flight pass from the other ground objects:

\[ x \in \text{buildings divergent to satellite flight pass, if } \text{NDDC} (x) > \epsilon \]

The $\rho_{HHHV}$ parameter is computed using the following equation:

\[
\rho_{HHHV} = \frac{\langle S_{HH}S_{HV}^* \rangle}{\sqrt{\langle S_{HH}S_{HH}^* \rangle} \sqrt{\langle S_{HV}S_{HV}^* \rangle}}
\]

where the superscript $\ast$ denotes the complex conjugate.

In our experiments, we found that the distributions in the complex plane of the $\rho_{HHHV}$ parameter for the buildings divergent to satellite flight pass and collapsed buildings are different. The $\rho_{HHHV}$ parameter for the collapsed buildings is mainly distributed in the third quadrant of the complex plane, while the $\rho_{HHHV}$ parameter for the buildings divergent to satellite flight pass is mainly distributed in the other areas of the complex plane. Therefore, the buildings divergent to satellite flight pass can be extracted using the $\rho_{HHHV}$ parameter by the expression:

\[ x \in \text{buildings divergent to satellite flight pass, if } (\text{Re} (\rho_{HHHV} (x)) < \epsilon_1 \& \& \text{Im} (\rho_{HHHV} (x)) < \epsilon_2) = 0 \]

where $x$ is the data sample of the PolSAR imagery; “Re” and “Im” denote the real part and imaginary part of the complex number, respectively; and $\epsilon_1$ and $\epsilon_2$ are the two threshold values of the real part and imaginary part of $\rho_{HHHV}$, respectively.
The $\rho_{HHHV}$ parameter is computed using the PolSAR data without POA compensation, which can better distinguish collapsed buildings from buildings divergent to satellite flight pass than using the PolSAR data after POA compensation, based on experiments.

In this work, the parameters of $\rho_{HHHV}$ and the NDDC together determine the buildings divergent to satellite flight pass. The buildings divergent to satellite flight pass extracted using the NDDC include some collapsed buildings with the walls divergent to satellite flight, but the buildings divergent to satellite flight pass extracted by the $\rho_{HHHV}$ parameter contain a few collapsed buildings with the walls divergent to satellite flight. Therefore, the data items simultaneously satisfying the two conditions of $\rho_{HHHV}$ and the NDDC are determined as the buildings divergent to satellite flight pass, which can result in a higher extraction accuracy for the buildings divergent to satellite flight pass. Using the parameters of $\rho_{HHHV}$ and the NDDC to extract the buildings divergent to satellite flight pass can be expressed as:

$$x \in \text{buildings divergent to satellite flight pass},$$

\[\text{if } \text{NDDC}(x) > \epsilon \& (\text{Re}(\rho_{HHHV}(x)) < \epsilon_1 \& \text{Im}(\rho_{HHHV}(x)) < \epsilon_2) = 0\]

The buildings divergent to satellite flight pass extracted by the parameters of $\rho_{HHHV}$ and the NDDC, together with the undamaged buildings generated from the Wishart supervised classification, are the final output undamaged buildings.

2.5. Collapsed Building Extraction

The vegetation class generated from the Wishart supervised classification still needs to be corrected. The NDDC values of vegetation items with reflection symmetry are low, and this property can be used to correct the classification result of the Wishart supervised classification for the vegetation class. Therefore, the data items of the vegetation generated from the Wishart supervised classification should simultaneously satisfy the condition of NDDC $< \epsilon$, and can be determined as the final output of the vegetation class. The buildings divergent to satellite flight pass are determined according to Section 2.4, and the bare areas are obtained through the Wishart supervised classification. By excluding the data items of the above three classes, the remaining data items can be classified as the collapsed buildings.

2.6. Building Collapse Rate Calculation

The building collapse rate is calculated at the block scale, and is defined as the ratio of the collapsed building samples to the total number of building samples in one block. The damage level of one block can be indexed by the building collapse rate of the block. The building collapse rate of blocks is introduced to handle the damage assessment at the block scale, and it is termed the building block collapse rate (BBCR). The blocks separated by roads are regarded as individual areas of similar built-up patch structure [19]. Each block is assigned a BBCR to assess the damage level of the block. The BBCR is expressed as:

$$\text{BBCR}_j = \frac{\sum_i C_{ij}}{\sum_i U_{ij} + \sum_i C_{ij}}$$

where $\text{BBCR}_j$ is the BBCR of the $j$th block; $C_{ij}$ indicates whether pixel $i$ in the $j$th block belongs to a collapsed building or not, with values of 0 or 1; and $U_{ij}$ indicates whether pixel $i$ in the $j$th block belongs to a undamaged building or not, with values of 0 or 1.

3. Experimental Results and Analysis

3.1. Experimental Data

The study case is the “4.14” Yushu earthquake with magnitude 7.1 which occurred on 14th April, 2010. This earthquake severely affected the county of Yushu in Qinghai province of China. The location of the
epicenter was $33.1^\circ$N and $96.6^\circ$E, as shown in Figure 2. The area is characterized by complicated terrain, poor transport infrastructure, and harsh climate. The vegetation of the Yushu urban region is sparse and low level, and the buildings are mainly rural residential buildings. More than 240,000 people were affected by the earthquake, and more than 2600 people died as a result of the earthquake. There were many collapsed buildings and more than 22 billion CNY of direct economic losses. The experiments were carried out on the post-event airborne PolSAR imagery to validate the effectiveness of the proposed approach for earthquake damage assessment. The PolSAR imagery was acquired one day after the earthquake by the Chinese airborne SAR mapping system (SARMapper), which was developed by a group led by the Chinese Academy of Surveying and Mapping (CASM). The system collects P-band data, and both the range resolution and azimuth resolution are approximately 1 m. Some specific information about the PolSAR data used in this work is listed in Table 1. The Pauli RGB image is shown in Figure 3, formed as a color composite of $|HH - VV|$ (red), $|HV|$ (green), and $|HH + VV|$ (blue), with the size of $8192 \times 4384$ pixels. The mountains surrounding Yushu County account for a large part of the imagery. In order to focus on analyzing the buildings of the urban area, an urban area mask was applied to discard the mountains and preserve the urban area as the region of interest. To allow a comparison with the existing damage assessment maps of the 4.14 Yushu earthquake [19,27,28], the geographic information system (GIS) data layer was depicted manually along the major roads and by reference to these damage assessment maps. The urban area was divided into 72 blocks by the GIS data layer containing 72 polygons, and the layout and characteristics of the buildings within the city blocks were similar. The ground-truth map shown in Figure 4 was drawn with reference to the above-mentioned damage assessment maps [19,27,28] and the high-resolution QuickBird image of Yushu County, and the damage levels were grouped into three levels for the 72 city blocks.

Table 1. Information about the fully-polarimetric synthetic aperture radar (PolSAR) data used in this work.

<table>
<thead>
<tr>
<th>Date</th>
<th>Flight Direction</th>
<th>Illumination Direction</th>
<th>Incidence Angle</th>
<th>Band</th>
<th>Flight Altitude (m)</th>
<th>Spatial Resolution (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 April 2010</td>
<td>From right to left</td>
<td>Bottom</td>
<td>50°</td>
<td>P</td>
<td>10,079</td>
<td>1 (range); 1 (azimuth)</td>
</tr>
</tbody>
</table>

![Figure 2. Map of the location of Yushu earthquake.](image-url)
Figure 3. Pauli RGB color composite (|HH − VV| (red), |HV| (green), and |HH + VV| (blue)) image of Yushu County. The areas marked by red and blue rectangles are the samples for collapsed buildings and buildings divergent to satellite flight pass, respectively. The center coordinates of the image are 33°0′9″N and 97°0′11″E.

Figure 4. Reference map for the earthquake damage assessment, with three damage levels: slight damage, moderate damage, and serious damage. If more than half of the buildings collapsed after the earthquake, the city block was considered as serious damage. The city block with less than one-third buildings collapsing was considered as slight damage. The city block with the damage level between slight damage and serious damage was considered as moderate damage.

3.2. The Results of the Experiments

According to the process flow of the earthquake damage assessment shown in Figure 1, the POA shift was first estimated based on the circular polarization method, and the POA compensation described in Section 2.2 was carried out for the PolSAR data after speckle noise filtering. The two dihedral components before and after POA compensation were then extracted on the basis of Yamaguchi four-component decomposition. As can be seen from Figure 5, the dihedral component image brightness changed after the POA compensation, and there are great differences in some parts of the two images, such as the upper right part and the bottom left part. These areas are mainly undamaged buildings divergent to satellite flight pass and collapsed buildings divergent to satellite flight pass. The NDDC parameter described in Section 2.4 was calculated and is shown in Figure 6. According to the color bar in Figure 6, 190 can be easily chosen as the threshold value for distinguishing the buildings divergent to satellite flight pass from the other ground objects, which are mainly reflection symmetric. The areas corresponding to NDDC > 190 are mainly buildings divergent to satellite flight pass. Meanwhile, the PolSAR data after POA compensation were classified into undamaged buildings, collapsed buildings, vegetation, and bare areas using the Wishart supervised classification described
in Section 2.3. For vegetation, the NDDC was less than 190. Therefore, if a data item was classified as vegetation by the Wishart supervised classification, and simultaneously satisfied NDDC < 190, it was finally classified as vegetation.

![Figure 5. Images of the dihedral component before and after POA compensation. (a) before POA compensation; (b) after POA compensation.](image)

The $\rho_{HHHV}$ parameter was calculated according to Section 2.4 using the PolSAR data without POA compensation. The 5000 samples were randomly selected from the two regions of interest (ROIs) of buildings divergent to satellite flight pass and the two ROIs of collapsed buildings, respectively, and are shown in Figure 7. As can be seen from Figure 7, the collapsed building samples are mainly distributed in the third quadrant of the complex plane, and the buildings divergent to satellite flight samples are mainly located in other parts of the complex plane. Therefore, the buildings divergent to
satellite flight pass class can be finally determined by the condition that the complex number $\rho_{HHHV}$ is not in the third quadrant of the complex plane, and the NDDC is greater than 190. The buildings divergent to satellite flight pass and the undamaged buildings generated from the Wishart supervised classification, which are mainly undamaged buildings parallel to satellite flight pass, together form the final classification result of the undamaged buildings.

![Figure 7](image)

**Figure 7.** $\rho_{HHHV}$ of the building divergent to satellite flight samples and collapsed building samples. The red dots and blue dots represent the collapsed building samples and the building divergent to satellite flight samples, respectively.

The bare areas class was determined by the classification result of the Wishart supervised classification. The collapsed buildings were considered to be the remaining parts after removing the three classes of vegetation, undamaged buildings, and bare areas. The distribution map of the non-buildings and the three kinds of buildings is shown in Figure 8. The building damage level index of each block was implemented using the BBCR method described in Section 2.6, as shown in Figure 9. A greater BBCR value corresponds to a more seriously damaged block. Three damage levels were set for the building collapse degree of the blocks. When the BBCR was less than 0.3, the damage level was set as slight damage. A BBCR of greater than 0.5 was set as the serious damage level, and a BBCR value of between 0.3 and 0.5 was considered to be moderate damage. The results of the damage assessment are shown in Figure 10, where the numbered blocks are the misclassified blocks, and the color of the numbers denotes the correct damage level. For example, the no. 8 block should actually be slightly damaged (green), but is misclassified as moderate damage (blue).

![Figure 8](image)

**Figure 8.** Distribution map of the non-buildings and the three kinds of buildings at the block scale. The NB, PB, BB and CB represent the non-buildings, the buildings parallel to satellite flight pass, the buildings divergent to satellite flight pass and the collapsed buildings, respectively.
3.3. Analysis and Discussion

The whole procedure of the production of a damage assessment map has taken five to six hours, which can meet rapid acquisition for buildings damage information after an earthquake. The damage assessment accuracies of the proposed method and the method only using the PolSAR data without POA compensation for the Wishart supervised classification are listed in Table 2. As can be seen in Table 2, the method proposed in this study significantly improves the damage assessment accuracy. In addition, compared with the method proposed by Zhao et al. [19], the proposed method again greatly improves the damage assessment accuracy. For the seriously damaged blocks, there are actually 25 blocks in the ground-truth map, but there are 31 blocks classified by the Wishart supervised classification using the PolSAR data without POA compensation, while there are 26 blocks generated by the proposed method. Therefore, the method proposed in this study also reduces the over-assessment of the building collapse rate, which could help to save human, physical, and financial resources for the emergency rescue after an earthquake.

Table 2. Comparison of the damage assessment accuracy of the two methods.

<table>
<thead>
<tr>
<th>The Proposed Method</th>
<th>Wishart Supervised Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Experiment)</td>
</tr>
<tr>
<td></td>
<td>(No. of Blocks)</td>
</tr>
<tr>
<td>SLD</td>
<td>MOD</td>
</tr>
<tr>
<td>SLD (Reference)</td>
<td>10</td>
</tr>
<tr>
<td>MOD</td>
<td>1</td>
</tr>
<tr>
<td>SED</td>
<td>0</td>
</tr>
</tbody>
</table>

OA: 88.89%  OA: 72.22%

OA, SLD, MOD, and SED represent overall accuracy, slight damage, moderate damage, and serious damage, respectively.
For the results of the method proposed in this study with the Yushu earthquake data, there are eight blocks misclassified. Among them, the damage levels of both block no. 1 and block no. 4 are underestimated, while the damage levels of blocks no. 2 and no. 3 and blocks nos. 5 to 8 are overestimated. These blocks are mainly low-rise and small rural residential buildings with earth/wood structure or masonry structure. The remaining walls of the collapsed buildings present dihedral scattering characteristics and are easily misclassified as undamaged buildings. Thus, some remaining parts of the damaged buildings parallel to satellite flight pass and the damaged buildings divergent to satellite flight pass may be misclassified as undamaged buildings, which is the main reason for the underestimation of the block damage level. The buildings divergent to satellite flight pass cannot be completely extracted with just the two parameters of the NDDC and $\rho_{HHHV}$, so some ground objects with comparatively weak scattering power and characterized by volume scattering are still classified as collapsed buildings. This results in the overestimation of the building collapse rate at the block scale.

4. Conclusions

Collapsed buildings caused by an earthquake are one of the main causes of casualties, so rapid acquisition of the collapsed building information after the earthquake can play an extremely important role in saving lives. The use of only a single post-earthquake PolSAR image to extract the collapse information can meet the needs of rapid and accurate disaster information acquisition, and can assist with a rapid and effective emergency rescue operation. In this study of earthquake damage assessment, the two parameters of the NDDC and $\rho_{HHHV}$ were used to improve the classification results of Wishart supervised classification performed on the PolSAR data after POA compensation, with the aim of obtaining non-buildings, undamaged buildings (including the buildings parallel to satellite flight pass and the buildings divergent to satellite flight pass), and collapsed buildings.

This feasibility study was performed on the airborne PolSAR imagery acquired one day after the Yushu earthquake. The buildings divergent to satellite flight pass were extracted based on the conditions of NDDC > 190 and $\rho_{HHHV}$ not being in the third quadrant of the complex plane, and were included in the undamaged buildings. These operations improved the accuracy of the undamaged building extraction. Using the condition of NDDC < 190 to restrict the vegetation generated from the Wishart supervised classification improved the non-building extraction accuracy and reduced the numbers of collapsed buildings divergent to satellite flight pass mixed with the vegetation class. At the same time, the above operations excluded the non-buildings and undamaged buildings as much as possible, and also improved the extraction accuracy of the collapsed buildings. Finally, the earthquake damage assessment map of the Yushu urban region with three damage levels at the block scale was obtained according to the value of the BBCR index for each block. The damage assessment at the block scale can not only be flexibly applied to multiple-resolution radar images and can avoid some of the errors of damage assessment at the single-building scale, but could also be more effective in assisting with making comprehensive arrangements in the process of emergency rescue.

All in all, the method proposed in this study can greatly improve the accuracy of earthquake damage assessment. Nevertheless, some undamaged buildings divergent to satellite flight still cannot be extracted, and some remaining parts of collapsed buildings are easily misidentified as undamaged buildings, which is the main reason for the errors in the building earthquake damage information extraction. In our future work, state-of-the-art filtering methods [29] and multi-classifier fusion [30,31] will be considered to improve the current results.

Acknowledgments: This work was supported by the Program for Changjiang Scholars and Innovative Research Team in University (IRT1278); the Hundred Talent Program of the Chinese Academy of Sciences (29Y127D01); the Cross-Disciplinary Collaborative Teams Program for Science, Technology and Innovation of the Chinese Academy of Sciences; the Earthquake Science and Technology Development Fund Program of Lanzhou Earthquake Research Institute, the China Earthquake Administration (2015M02); the Object-Oriented High Trusted SAR Processing System of the National 863 Subject Program; and the Airborne Multiband Polariometric Interferometric SAR Mapping System of the National Major Surveying and Mapping Science and Technology Special Program. We would also like to thank the anonymous reviewers for their advice on improving the quality of this paper.
**Author Contributions:** Wei Zhai drafted the manuscript and was responsible for the research design, writing the source code, data analysis, and interpretation of the results. Huanfeng Shen designed the structures of the paper and guided the experiments. Chunlin Huang edited and reviewed the manuscript. Wansheng Pei collected literatures.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).