water

# Application of Remote Sensing and GIS in Droughts and Floods Assessment and Monitoring

Edited by
Yaohuan Huang, Yesen Liu, Runhe Shi and Hongyan Ren

Printed Edition of the Special Issue Published in *Water*

www.mdpi.com/journal/water

MDPI

# Application of Remote Sensing and GIS in Droughts and Floods Assessment and Monitoring

# Application of Remote Sensing and GIS in Droughts and Floods Assessment and Monitoring

Editors

**Yaohuan Huang**
**Yesen Liu**
**Runhe Shi**
**Hongyan Ren**

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin

*Editors*

Yaohuan Huang
Chinese Academy of Sciences
Beijing
China

Yesen Liu
China Institute of Water
Resources and Hydropower
Research
Beijing
China

Runhe Shi
East China Normal
University
Shanghai
China

Hongyan Ren
Chinese Academy of Sciences
Beijing
China

*Editorial Office*
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Water* (ISSN 2073-4441) (available at: https://www.mdpi.com/journal/water/special_issues/Remote_Sensing_GIS).

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

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

**ISBN 978-3-0365-7146-1 (Hbk)**
**ISBN 978-3-0365-7147-8 (PDF)**

# Contents

# About the Editors

**Yaohuan Huang**

Yaohuan Huang, Associate Professor. Received Ph.D. degree in hydrology and water resources from China Institute of Water Resources and Hydropower Research. Performed research in remote sensing and GIS-based applications in hydrological and ecological assessment and monitoring, UAV remote sensing objects detection, spatial analyst, and land use/cover change and their interactions with climate change.

**Yesen Liu**

Yesen Liu, Professor. Received Ph.D. degree in hydraulic engineering from Tianjin University in China. Performed research in urban hydrology, hydrological parameters simulation, flash floods analysis, rainfall impact on floods and digital twin watershed based on GIS over a 20 year career.

**Runhe Shi**

Runhe Shi, Associate Professor. Received Ph.D. degree in geography information system from Institute of Geographic Sciences and Natural Resources Research, CAS. Performed Research in wetland vulnerability, coastal monitoring, retrieval of land surface parameters, uncertainty analysis and improvement of vegetation index based on remote sensing and GIS.

**Hongyan Ren**

Hongyan Ren, Associate Professor. Received Ph.D. degree in geography information system from Institute of Geographic Sciences and Natural Resources Research, CAS. Performed research in water environment and public health, agriculture water resources based on hyperspectral remote sensing, geospatial analysis, hydrological harzards assessement and monitoring based on machine learning.

# Preface to "Application of Remote Sensing and GIS in Droughts and Floods Assessment and Monitoring"

The goal of this Special Issue is to discuss and address the applications of state-of-the-art techniques such as remote sensing and GIS in drought or flood monitoring and hydrological hazards assessment. Exhibiting diversity in topic areas, the research can generally be classified as follows: (1) remote sensing application in drought assessment; (2) the impact of rainfall characters on floods; (3) GIS technique application in flood forecasting and (4) water quality deterioration investigation based on GIS technique and in situ observation.

**Yaohuan Huang, Yesen Liu, Runhe Shi, and Hongyan Ren**
*Editors*

water

MDPI

*Editorial*

# Application of Remote Sensing and GIS in Drought and Flood Assessment and Monitoring

**Yaohuan Huang [1,2], Yesen Liu [3,*], Runhe Shi [4] and Hongyan Ren [1]**

[1] State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[2] College of Resource and Environment, University of Chinese Academy of Sciences, Beijing 100049, China
[3] State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing 100038, China
[4] Key Laboratory of Geographic Information Science, Ministry of Education, East China Normal University, Shanghai 200241, China
* Correspondence: liuys@iwhr.com

## 1. Introduction

Driven by global change and population pressure, droughts and floods have been two of the most serious natural hazards, leading to crop losses and economic havoc in many areas and ultimately affecting more people globally than any other natural hazard [1–4]. As mentioned in the 2021 by the Intergovernmental Panel on Climate Change (IPCC) Group 1 Report, the average global temperature increased by about 1.07 °C during 2010–2019. The report also warns of more frequent and intense extreme weather and climate events, describing a potential proliferation of droughts and floods [5]. Over the past few decades, extreme events of floods and droughts have increased [6], such as the severe heatwave and droughts of Europe in 2003 and 2018 [7,8], the flood that occurred in Pakistan in 2010 [9], the western Russian drought in 2010 [10], etc. Future changes are particularly drastic in regions that include many developing nations, societies which are especially vulnerable to global climate change. Under future climate conditions, the hydrological cycle will be forced to accelerate, and many areas of the world are projected to experience increased occurrences of extreme weather and climate events [11,12]. As droughts and floods are complex hydrological systems, they deserve a multidisciplinary monitoring effort in order to conduct appropriate and timely hazard assessments. Recently, remote sensing and GIS-based techniques have been widely applied to obtain synoptic and punctual overviews of basin-scale monitored areas [13,14]. It is clear that the application of remote sensing and GIS can potentially provide an extra contribution to drought and flood assessment and monitoring, for instance, in terms of accuracy of results, amount of information obtained, temporal availability, and so on.

In this Special Issue, we attempted to discuss and address the applications of remote sensing, GIS and other state-of-the-art techniques in drought or flood monitoring and hydrological hazards assessment. To fulfill these objects, we strongly invited contributions on various droughts or flood monitoring indexes from satellites and other data resources such as high time resolution and high-resolution imaging, or the Gravity Recovery and Climate Experiment (GRACE). Considering that the processes of hydrological hazards such as droughts and floods are complex, research based on machine learning and modeling was also included in this Special Issue. The investigative approach characterized by the integration of disciplines at different scales of vision and precision represents a modern effort to strive for a more complete understanding of drought and flood processes and, therefore, a better hazard evaluation.

**Citation:** Huang, Y.; Liu, Y.; Shi, R.; Ren, H. Application of Remote Sensing and GIS in Drought and Flood Assessment and Monitoring. *Water* **2023**, *15*, 541. https://doi.org/10.3390/w15030541

Received: 13 January 2023
Revised: 20 January 2023
Accepted: 27 January 2023
Published: 30 January 2023

## 2. Summary of This Special Issue

Amongst the papers in this Special Issue that represent examples of the state of the art of remote sensing application in drought assessment was that submitted by Cui et al. [15], in which the authors study the influence of climate background on drought events in mainland China. In fact, droughts are one of the most serious natural hazards in China, but they are also complexly affected by climate change. By considering water content as a whole index, the terrestrial water storage changes (TWSCs), derived from GRACE time-variable gravity fields, have constituted a useful dataset in hydrology research [13,14]. In this case, a drought severity index (denoted as GRACE-DSI) derived from TWSCs was applied, analyzing the role of the drought-related factors (e.g., precipitation, evapotranspiration) and extreme climate events (e.g., El Niño–Southern Oscillation (ENSO) and North Atlantic Oscillation (NAO) events) in the formation of droughts. The results of this study are valuable in the efforts to understand the formation mechanism of drought events.

Floods are also one of the most serious natural hazards, which can lead to crop losses and economic havoc in many areas, affecting more people globally than any other natural hazard. In this Special Issue, four interesting manuscripts based on the GIS technique were published, providing new insights about spatiotemporal patterns and modes of rainstorm, the impact of rainfall movement direction, and the modeling of flood forecasting. Liu et al. [16] used a manifold learning algorithm method of machine learning to analyze rainstorm patterns, which is considered to be essential for improving the precision and accuracy of flood forecasts and constructing flood disaster prevention systems. This research analyzed the spatial–temporal characteristics of heavy rain in Beijing and Shenzhen in China and found the key factors (topography and water vapor) to be diverse in different regions. The proposed method provided a possible way to analyze spatio-temporal distribution characteristics of rainfall, which can help stakeholders to establish strategies to reduce flood risks in different regions.

Liu et al. [17] proposed an approach to identify the characteristics of rainstorms of a short duration in urban areas from their temporal and spatial dimensions. This study case identified the typical spatiotemporal modes of rainfall and the reconstruction of the process of modes of Beijing in China. The result showed that there were three modes of rainstorms in the Beijing urban area, information which can be applied to rainstorm forecasting and flood prevention in inner urban areas. The authors stated that this approach provides more complete characteristics identification, including of its temporal and spatial dimensions, than traditional methods for considering a rainfall as one complete process.

Liu et al. [18] found rainfall movement direction to be a significant rainfall variability in urban floods, which is always ignored when comparing with rainfall intensity and duration. This study provided a very innovative insight into the impact of spatial–temporal rainfall variations on urban floods. In total, 1313 rainfall scenarios with different combinations of rainfall intensity and rainfall movement direction in the typically rainy city of Shenzhen in China were analyzed to investigate the effect of rainfall movement direction. They concluded that the impact of rainfall movement direction is almost symmetrical and is associated with the direction of the river. The closer rainfall movement direction is to the linear directional mean of rivers, the larger the peak runoff of section will be. The authors stated that rainfall movement direction is significant to urban peak runoff in the downstream reaches, something which should be considered in urban hydrological analysis.

Another example of monitoring floods is the research published by Liu et al. [19] about improvement of a floods elements correction model of the ridge estimation-based dynamic system response curve (DSRC-R). They proposed a new criterion called the balance and random degree criterion, considering the sum of squares of flow errors (BSR) to optimize the ridge coefficient in the DSRC-R method. The results indicated that the techniques can greatly shorten the search time of the ridge coefficient in optimization, which will improve operational efficiency and enhance the real-time flood forecasting performance.

Besides droughts and floods, water quality deterioration has become a serious hydrological hazard of late [20,21]. The rest of the two published papers were both related to

water quality. These two papers published here showing different study cases are very interesting. Chen et al. [22] studied the changes in the environmental quality of surface water during the "13th Five-Year Plan" period (2016–2020) in Heilongjiang Province, the location of the most important grain production base and the province with the highest latitude in China. They concluded that the population, the primary industry, the tertiary industry and forestry are the main factors affecting the change in water environment quality in Heilongjiang Province. This study provided a case analysis for water quality of Heilongjiang province in China, which will be helpful to regional water environment protection.

Zhang et al. [23] investigated the temporal and spatial patterns of surface water quality in China since the reform and opening-up program based on the monitoring of datasets. They indicated that the temporal change trend in surface water quality in China presented a "fluctuating changes stage–rapid deterioration stage–fluctuations stalemate stage–rapid improvement stage" pattern. They also concluded that the current regional surface water quality of China still has a polluted status. They sated that the potential for the continuous reduction in major pollutant discharges had become more challenging, and the marginal cost for pollution control had increased. This study provided a case analysis for the water quality of a whole country, which will be helpful to national water environmental protection.

**Author Contributions:** Conceptualization, Y.H. and Y.L.; writing—original draft preparation, Y.L.; writing—review and editing, Y.H. and R.S.; supervision, R.S. and H.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** Thanks to all of the contributions to the Special Issue, the time invested by each author, as well as to the anonymous reviewers and editorial managers who have contributed to the development of the articles in this Special Issue. All the guest editors are very happy with the review process and management of the Special Issue and offer their thanks.

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

## References

1. Liu, Y.; Yuan, X.; Guo, L.; Huang, Y.; Zhang, X. Driving Force Analysis of the Temporal and Spatial Distribution of Flash Floods in Sichuan Province. *Sustainability* **2017**, *9*, 1527. [CrossRef]
2. Liu, Y.; Huang, Y. Why Flash Floods Occur Differently across Regions? A Spatial Analysis of China. *Water* **2020**, *12*, 3344. [CrossRef]
3. Huang, Y.; Li, Z.; Chen, M.; Song, X.; Kang, P. Spatial Variability of Water Resources State of Regions around the "Belt and Road". *Water* **2021**, *13*, 2102. [CrossRef]
4. Liu, Y.; Huang, Y.; Wan, J.; Yang, Z.; Zhang, X. Analysis of Human Activity Impact on Flash Floods in China from 1950 to 2015. *Sustainability* **2021**, *13*, 217. [CrossRef]
5. IPCC. Summary for Policymakers. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. 2021. Available online: https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_SPM_final.pdf (accessed on 10 January 2023).
6. Li, S. Natural extreme events, government subsidies and corporate environment responsibility: Evidence from China's energy sector. *Energy Econ.* **2022**, *114*, 106278. [CrossRef]
7. Robine, J.; Cheung, S.; Le Roy, S.; Van Oyen, H.; Griffiths, C.; Michel, J.; Herrmann, F. Death toll exceeded 70,000 in Europe during the summer of 2003. *Comptes Rendus Biol.* **2008**, *331*, 2. [CrossRef]
8. Bastos, A.; Ciais, P.; Friedlingstein, P.; Sitch, S.; Zaehle, S. Direct and seasonal legacy effects of the 2018 heat wave and drought on European ecosystem productivity. *Sci. Adv.* **2020**, *6*, eaba2724. [CrossRef]
9. Hong, C.; Hsu, H.; Lin, N.; Chiu, H. Roles of European blocking and tropical-extratropical interaction in the 2010 Pakistan flooding. *Geophys. Res. Lett.* **2011**, *38*, 13. [CrossRef]
10. Shaposhnikov, D.; Revich, B.; Bellander, T.; Bedada, G.; Bottai, M.; Kharkova, T.; Kvasha, E.; Lezina, E.; Lind, T.; Semutnikova, E.; et al. Mortality related to air pollution with the moscow heat wave and wildfire of 2010. *Epidemiology* **2014**, *25*, 3. [CrossRef]
11. Coffel, E.; Keith, B.; Lesk, C.; Horton, R.; Bower, E.; Lee, J.; Mankin, J. Future hot and dry years worsen Nile Basin water scarcity despite projected precipitation increases. *Earth's Future* **2019**, *7*, 967–977. [CrossRef]
12. Ridder, N.; Ukkola, A.; Pitman, A.; Perkins-Kirkpatrick, S. Increased occurrence of high impact compound events under climate change. *NPJ Clim. Atmos. Sci.* **2022**, *5*, 3. [CrossRef]

13. Zhao, C.; Huang, Y.; Li, Z.; Chen, M. Drought Monitoring of Southwestern China Using Insufficient GRACE Data for the Long-Term Mean Reference Frame under Global Change. *J. Clim.* **2018**, *31*, 6897–6911. [CrossRef]
14. Liu, F.; Kang, P.; Zhu, H.; Han, J.; Huang, Y. Analysis of Spatiotemporal Groundwater-Storage Variations in China from GRACE. *Water* **2021**, *13*, 2378. [CrossRef]
15. Cui, L.; Zhang, C.; Yao, C.; Luo, Z.; Wang, X.; Li, Q. Analysis of the Influencing Factors of Drought Events Based on GRACE Data under Different Climatic Conditions: A Case Study in Mainland China. *Water* **2021**, *13*, 2575. [CrossRef]
16. Liu, Y.; Liu, Y.; Ren, H.; Du, L.; Liu, S.; Zhang, L.; Wang, C.; Gao, Q. Spatial and Temporal Pattern of Rainstorms Based on Manifold Learning Algorithm. *Water* **2023**, *15*, 37. [CrossRef]
17. Liu, W.; Chen, S.; Tian, F. Spatiotemporal Modes of Short Time Rainstorms Based on High-Dimensional Data: A Case Study of the Urban Area of Beijing, China. *Water* **2021**, *13*, 3597. [CrossRef]
18. Liu, Y.; Huang, Y.; Liu, Y.; Li, K.; Li, M. The Impact of Rainfall Movement Direction on Urban Runoff Cannot Be Ignored in Urban Hydrologic Management. *Water* **2021**, *13*, 2923. [CrossRef]
19. Liu, K.; Bao, W.; Hu, Y.; Sun, Y.; Li, D.; Li, K.; Liang, L. Improvement in Ridge Coefficient Optimization Criterion for Ridge Estimation-Based Dynamic System Response Curve Method in Flood Forecasting. *Water* **2021**, *13*, 3483. [CrossRef]
20. Nabi, B.G.; Mukhtar, K.; Arshad, R.N.; Radicetti, E.; Tedeschi, P.; Shahbaz, M.U.; Walayat, N.; Nawaz, A.; Inam-Ur-Raheem, M.; Aadil, R.M. High-Pressure Processing for Sustainable Food Supply. *Sustainability* **2021**, *13*, 13908. [CrossRef]
21. Ma, T.; Sun, S.; Fu, G.; Jim, H.; Ni, Y.; He, L.; Yi, J.; Zhao, N.; Du, N.; Du, Y.; et al. Pollution exacerbates China's water scarcity and its regional inequality. *Nat. Commun.* **2020**, *11*, 650. [CrossRef]
22. Chen, W.; Bai, Y.; Li, B.; Feng, C.; Zhou, M. Analysis of Water Environment Quality Changes and Influencing Factors during the "Thirteenth Five-Year Plan" Period in Heilongjiang Province. *Water* **2022**, *14*, 2367. [CrossRef]
23. Zhang, F.; Lin, L.; Li, W.; Fang, D.; Lv, Z.; Li, M.; Ma, G.; Wang, Y.; Wang, L.; He, L. Long-Term Study of Monitoring History and Change Trends in Surface Water Quality in China. *Water* **2022**, *14*, 2134. [CrossRef]

water

MDPI

*Article*

# Analysis of the Influencing Factors of Drought Events Based on GRACE Data under Different Climatic Conditions: A Case Study in Mainland China

**Lilu Cui [1,2], Cheng Zhang [1], Chaolong Yao [3,*], Zhicai Luo [4], Xiaolong Wang [5] and Qiong Li [6]**

1 School of Architecture and Civil Engineering, Chengdu University, Chengdu 610106, China; cuililu@cdu.edu.cn (L.C.); lilucui@whu.edu.cn (C.Z.)
2 School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China
3 College of Natural Resources and Environment, South China Agricultural University, Guangzhou 510642, China
4 MOE Key Laboratory of Fundamental Physical Quantities Measurement & Hubei Key Laboratory of Gravitation and Quantum Physics, PGMF and School of Physics, Huazhong University of Science and Technology, Wuhan 430074, China; zcluo@hust.edu.cn
5 Nanning Exploration & Survey Geoinformation Institute, Nanning 530022, China; wwxyzwxl@whu.edu.cn
6 School of Civil Engineering and Geomatics, Southwest Petroleum University, Chengdu 610500, China; qiongli@swpu.edu.cn
* Correspondence: clyao@scau.edu.cn; Tel.: +86-136-4075-1027

**Abstract:** The occurrence of droughts has become more frequent, and their intensity has increased in mainland China. With the aim of better understanding the influence of climate background on drought events in this region, we analyzed the role of the drought-related factors and extreme climate in the formation of droughts by investigating the relationship between the drought severity index (denoted as GRACE-DSI) based on the terrestrial water storage changes (TWSCs) derived from Gravity Recovery and Climate Experiment (GRACE) time-variable gravity fields and drought-related factors/extreme climate. The results show that GRACE-DSI was consistent with the self-calibrating Palmer Drought Severity Index in mainland China, especially for the subtropical monsoon climate, with a correlation of 0.72. Precipitation (PPT) and evapotranspiration (ET) are the main factors causing drought events. However, they play different roles under different climate settings. The regions under temperate monsoon climate and subtropical monsoon climate were more impacted by PPT, while ET played a leading role in the regions under temperate continental climate and plateau mountain climate. Moreover, El Niño–Southern Oscillation (ENSO) and North Atlantic Oscillation (NAO) events mainly caused abnormalities in PPT and ET by affecting the strength of monsoons (East Asian and Indian monsoon) and regional highs (Subtropical High, Siberian High, Central Asian High, etc.). As a result, the various affected regions were prone to droughts during ENSO or NAO events, which disturbed the normal operation of atmospheric circulation in different ways. The results of this study are valuable in the efforts to understand the formation mechanism of drought events in mainland China.

**Keywords:** GRACE; drought; mainland China; extreme climate; climatic conditions

**Citation:** Cui, L.; Zhang, C.; Yao, C.; Luo, Z.; Wang, X.; Li, Q. Analysis of the Influencing Factors of Drought Events Based on GRACE Data under Different Climatic Conditions: A Case Study in Mainland China. *Water* **2021**, *13*, 2575. https://doi.org/10.3390/w13182575

Academic Editors: Yaohuan Huang, Yesen Liu, Runhe Shi and Hongyan Ren

Received: 20 August 2021
Accepted: 15 September 2021
Published: 18 September 2021



## 1. Introduction

Drought is a severe natural hazard event on a global scale characterized by terrestrial water deficit. It has a negative impact on socioeconomic development, crop failure, ecosystems, and the lives of people [1–3]. Therefore, research on the influencing factors of drought events are of great significance for establishing early warning, strengthening water resources management and reducing disaster losses [4]. The drought index has always been used as a quantitative indicator to characterize the drought events due to its simple and easy-to-understand characteristics [5]. At present, the commonly used drought indices

are mainly the palmer drought severity index (PDSI) [6], the standardized precipitation index (SPI) [7] and the standardized runoff index (SRI) [8]. These traditional drought indices are mainly calculated based on the long-term accumulation of drought-related data such as precipitation (PPT), evapotranspiration (ET), temperature, etc. The drought-related data is provided by the meteorological stations. However, an insufficient number and uneven distribution of these stations in some regions leads to the inability to obtain surface data with high spatial resolution [9]. Furthermore, the traditional technical approaches not only require a great deal of construction and maintenance, and they can only observe parts of the hydrological component in the terrestrial water cycle [10]. Therefore, it is impossible to explain the cause of drought from the perspective of the entire terrestrial water cycle. These problems also appear in the traditional drought indices.

Since 2002, the Gravity Recovery and Climate Experiment (GRACE) mission [11] has provided monthly data on Earth's gravity field to infer the total terrestrial water storage change (TWSC) including surface and subsurface hydrological components [12]. Many research works have proved that the GRACE data can detect regional drought events and assess drought-associated losses, in regions such as the Amazon River [10,13,14], Yangtze River [9,15–17], Southeastern China [18–20] and State of Texas [21]. Therefore, some scholars have used GRACE TWSC data to construct drought indexes to achieve more accurate detection and assessment of local drought. Yirdaw et al. [22] derived the total storage deficit index (TSDI) to characterize the drought events in the Katchewan River during 2002 and 2003. Wang et al. [23] used the GRACE TWS anomaly index (TWSI), PPT anomaly index and vegetation anomaly index to detect drought events in the Haihe River basin from January 2003 to January 2013. The results indicate that TWSI is more suitable than traditional indices to monitor these drought events. Yi et al. [24] constructed the GRACE-based hydrological drought index (GHDI) to monitor the drought events in the United States from 2003 to 2012. The results indicate that the GHDI has a good correlation with PDSI over the United States. Sinha et al. [25] used the water storage deficit index (WSDI) to assess drought events in India. The results illustrate the validity and reliability of WSDI in quantifying the characteristics of large-scale drought events. Zhao et al. [26] developed the GRACE-based drought severity index (GRACE-DSI) to capture the major drought events worldwide, and GRACE-DSI showed good temporal and spatial agreement with PDSI and the standardized precipitation evapotranspiration index (SPEI). The above studies show that the GRACE-based drought index is a valuable tool for the detection and assessment of hydrological drought.

When people conduct in-depth research on drought-related data, they have a certain understanding of the driving mechanisms of drought. Li et al. [19] and Wu et al. [27] indicate that the extreme drought in Southwest China in 2010 was mainly caused by insufficient PPT, and that excessive ET played a secondary role. Panisset et al. [28] explain that anomalous PPT deficit and solar radiation anomalies were the main factors leading to the three drought events in the Amazon basin in 2005, 2010 and 2015. Zhang et al. [17] studied the two drought events that occurred in the Yangtze River basin in 2006 and 2011. The results showed that there was a certain connection with droughts and El Niño–Southern Oscillation (ENSO) in this region, and the TWSC in the lower reaches was more sensitive to the change in ENSO than the TWSC in the upper and middle reaches. However, the above studies mainly focused on a certain local drought event, and did not consider the influence of the regional climate background on the drought event.

Mainland China (MC) is the region with the most frequent drought disasters worldwide; these disasters have brought huge losses to the region, and local drought disasters occur almost every year in the region [29]. Therefore, we used MC as a research region to study the drought events that occurred from April 2002 to June 2017 in four different climate regions based on GRACE-DSI data. We calculated the correlation coefficients between PPT, runoff, ET soil water storage and GRACE-DSI in the four different climate regions to analyze the impact of the different climate backgrounds on drought events. We also carried out a statistical analysis of the PPT, ET and GRACE-DSI during ENSO and

North Atlantic Oscillation (NAO) events to discuss the influence of extreme climate on drought events in different climate regions. These research results can help to understand the driving mechanisms of drought events and provide early warning of drought disaster.

## 2. Study Area

MC is located approximately within 19° N–53° N and 73° E–135° E, and has an area of about 9.6 million km$^2$. Its digital elevation model shows that the terrain is like a ladder, gradually descending from west to east (Figure 1).

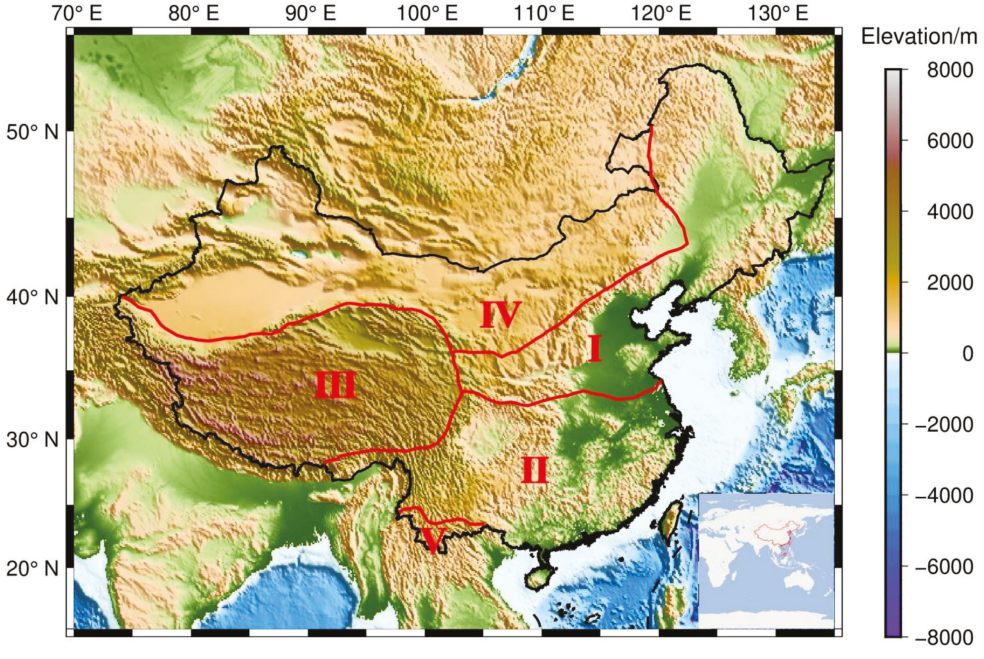



**Figure 1.** Digital elevation model of mainland China. The different climate regions are marked on the map: temperate monsoon climate (I), subtropical monsoon climate (II), plateau mountain climate (III), temperate continental climate (IV) and tropical monsoon climate (V).

China is a vast territory with a wide span of latitudes and many areas that are far from the sea. In addition, the terrain types and mountain directions are diverse, resulting in a diversity of temperature, PPT and climates formed. The Eastern region has a monsoon climate, the Northwest region has a temperate continental climate and the Qinghai-Tibet Plateau has a plateau mountain climate. There are also humid regions, semi-humid regions, semi-arid regions and arid regions. Furthermore, it is one of the countries with the most rivers in the world. There are more than 1500 rivers with an area larger than 1000 km$^2$ in MC, including the Yangtze River, Yellow River, Pearl River, Huaihe River, Liaohe River and Songhua River. Most rivers are located in Eastern and Southern China, and a few rivers are located in Northwest China. The ET gradually decreases from Southeast to Northwest China. The distribution of precipitation decreases sharply from the Southeastern (>3000 mm) to the Northwestern region (<50 mm) [30]. The occurrence of droughts and floods has become more frequent, and the intensity of drought and floods has also increased [31]. According to the climate type, MC can be divided into five parts (Figure 1), namely temperate monsoon climate (I), subtropical monsoon climate (II), plateau mountain climate (III), temperate continental climate (IV) and tropical monsoon climate (V). As the

area of region V is relatively small, it is not suitable for GRACE detection; we focus on regions I–IV in this study.

## 3. Data and Methods

### *3.1. GRACE TWSC*

The GRACE RL06 monthly spherical harmonic (SH) coefficients product (truncated to degree and order 60) was provided by the Center for Space Research at the University of Texas at Austin, and was used to calculate TWSCs in MC for the period from April 2002 to December 2016.

The GRACE data were preprocessed as follows: C20 coefficients were replaced with those derived by satellite laser ranging [32]. Degree-1 coefficients were replaced using Swenson's results [33]. Filter processing combining a 300 km fan filter [34] and a PM36 de-correlation filter [35] was performed to weaken high-frequency and correlated errors. Due to the influence of order truncation and filter processing, there are leakage errors in the hydrological signals derived from GRACE data. The single scale factor method was used to calibrate GRACE-based TWSC results to restore loss signals [36].

### *3.2. GLDAS Model*

The GLDAS is a global high-resolution land surface model published jointly by the Goddard Space Flight Center at NASA and the National Centers for Environmental Prediction at the National Oceanic and Atmospheric Administration (NOAA). It incorporates space- and ground-based observation and uses data assimilation techniques [37]. The monthly soil moisture (SM) and runoff data with a $1^\circ \times 1^\circ$ spatial resolution were provided by three GLDAS-2.1 models (Noah, the variable infiltration capacity model and the catchment land surface model). The SM and runoff data were the average of these three models' data, and the time span was from April 2002 to December 2016.

### *3.3. In Situ Precipitation (PPT) Data*

Monthly gridded precipitation data for the time period between April 2002 and December 2016 provided by the China National Meteorological Science Data Center and sorted by the National Meteorological Information Center with spatial resolution $0.5^\circ \times 0.5^\circ$ were used for the analysis.

### *3.4. ET Data*

ET is estimated according to a water balance equation [38,39]. Its expression is as follows:

$$\mathrm{ET} = \mathrm{P} - \mathrm{R} - \mathrm{TWSC} \quad (1)$$

where P is PPT, R is the averaged runoff derived from GLDAS-2.1, and TWSC is derived from GRACE data.

### *3.5. Self-Calibrating Palmer Drought Severity Index (SCPDSI) Data*

The SCPDSI data [40], a meteorological drought index, can evaluate the water loss caused by the imbalance of surface water supply and demand [41,42], and is provided by the Climate Research Unit at University of East Anglia. In this study, we extracted the relevant gridded data from April 2002 to December 2016 with spatial resolution $0.5^\circ \times 0.5^\circ$ in MC. The severity of drought events can be classified as shown in Table 1 [43].

**Table 1.** The grades of SCPDSI drought classification.

| Type | SCPDSI | Type | SCPDSI |
|---|---|---|---|
| Extreme Drought | $\leq -4.0$ | Light Drought | $-1.0 \sim -2.0$ |
| Heavy Drought | $-3.0 \sim -4.0$ | No Drought | $\geq -1.0$ |
| Moderate Drought | $-2.0 \sim -3.0$ | | |

### *3.6. Extreme Weather Index Data*

The ENSO is an abnormal phenomenon characterized by ocean surface warming or higher sea surface temperatures occurring in the Equatorial Eastern and Middle Pacific, and is able to influence the global atmospheric circulation, causing abnormal temperature and PPT [44,45]. The monthly Niño 3.4 index data indicate the magnitude of ENSO, which is provided by the NOAA. An El Niño event is designated as the occurrence of an ENSO index greater than or equal to 0.5 for 5 consecutive months, while a La Niña event is defined as an ENSO index less than or equal to −0.5 for 5 consecutive months [44,46].

The NAO is a kind of atmospheric circulation change that occurs in the middle and high latitudes of the Northern Hemisphere in winter, which reflects the atmospheric mass changes between Iceland Depression and Azores High in the North Atlantic. The NAO index can reflect the changes in the Iceland Depression and Azores High [47], and is also provided by the NOAA.

### *3.7. Calculation of GRACE-DSI Data*

Based on the gridded GRACE TWSCAs estimated in Section 3.1, GRACE-DSI is the standardized GRACE-based TWSC as follows [26]:

$$\text{GRACE-DSI}_{i,j} = \frac{\text{TWSC}_{i,j} - \text{TWSC}_j^{mean}}{\sigma_j} \quad (2)$$

where $\text{TWSC}_{i,j}$ is TWSC in the $i$th year and $j$th month. $i$ is a specific year from 2002 to 2016, and $j$ is a specific month from January to December. $\text{TWSC}_j^{mean}$ and $\sigma_j$ are the average and standard deviation of the TWSC in month $j$, respectively. This index can be used to detect drought and abnormally wet events. According to the size of GRACE-DSI values, drought events can be classified as shown in Table 2 [26]. Due to the truncation degree and filtering effect, the spatial resolution of GRACE-DSI grid data is 350 km [48].

**Table 2.** GRACE-DSI drought grades classification.

| Type | GRACE-DSI | Type | GRACE-DSI |
|---|---|---|---|
| Exceptional Drought | $\leq -2.0$ | Moderate Drought | $-1.3$~$-0.8$ |
| Extreme Drought | $-2.0$~$-1.6$ | Light Drought | $-0.8$~$-0.5$ |
| Severe Drought | $-1.6$~$-1.3$ | No Drought | $\geq -0.5$ |

### *3.8. The Extraction of Anomaly Signal*

To discuss the relationship between GRACE-DSI and hydrological components, it is necessary to extract the anomaly signal from the original signal of each hydrological component. The time series of the original data can be decomposed into a long-term trend change term, seasonal term and anomaly term. The expression is as follows [36]:

$$Data(t) = a_0 + a_1 t + a_2 \cos(2\pi t) + a_3 \sin(2\pi t) + a_4 \cos(4\pi t) + a_5 \sin(4\pi t) + \varepsilon \quad (3)$$

where $Data(t)$ is the original data; $t$ is the time; $\varepsilon$ is the residual signal; and $a_0, a_1, a_2, a_3, a_4, a_5$ are the pending parameters. $a_0$ is a constant; $a_1$ is the long-term trend change term; and $a_2, a_3, a_4, a_5$ are the seasonal signals. Therefore, the expression of the anomaly signals is:

$$Data_{anomaly}(t) = Data(t) - [a_1 t + a_2 \cos(2\pi t) + a_3 \sin(2\pi t) + a_4 \cos(4\pi t) + a_5 \sin(4\pi t)] \quad (4)$$

## 4. Results and Analysis

### *4.1. Comparison of GRACE-DSI and SCPDSI*

To verify the drought detection ability of GRACE-DSI, we compared the temporal and spatial distribution of GRACE-DSI and SCPDSI (Figures 2 and 3). From Figure 2, it can be seen that the two drought indices show good consistency in general, with a correlation of 0.66 (Table 3). There were four episodes with long-term negative GRACE-DSI. The

first time period was from August 2002 to March 2003, with the minimum value (−0.64) occurring on March 2003. Although SCPDSI showed similar changes during this period, the negative value (−0.07) was small and only appeared for one month (October 2002). The values of GRACE-DSI were mostly negative during the second period from June 2006 to February 2010, coinciding with SCPDSI. The minimum values of these indices appeared on April 2008 (GRACE-DSI, −0.62) and August 2006 (SCPDSI, −0.71), respectively. During the third period from March 2011 to April 2012, the minimum value of GRACE-DSI (−0.40) was smaller than that of SCPDSI (−0.31), which appeared on January 2012 and July 2011, respectively. For the fourth dry period from September 2013 to July 2015, the minimum value (−0.43) of GRACE-DSI occurred in January 2015 and the one of SCPDSI was −0.35 in April 2014.

Figure 3 shows the spatial distribution of GRACE-DSI and SCPDSI from October 2009 to September 2010. It can be seen in Figure 3 that the two drought indices had similar spatial distribution. In October 2009, most parts of MC were in the dry state, and only parts of Qinghai Province and the northern part of Northeast China were in the humid state. From November 2009 to January 2010, the arid regions were gradually decreasing, and the degree of dryness was also decreasing. January 2010 was the period with the fewest arid regions. By February 2010, the arid regions suddenly expanded, which was mainly concentrated in Southwest China, eastern Xinjiang, southern Northeast China and northern North China. The humid regions were concentrated in Qinghai Province, Northeast China and the Southeast Coastal region. Subsequently, the arid regions gradually decreased and the humid regions gradually increased. This trend continued until September 2010. The temporal and spatial changes of GRACE-DSI and SCPDSI in Southwest China are consistent with the severe drought event that occurred in the same period in the region [19].

From the perspective of different climate types, the highest correlation (0.72) between GRACE-DSI and SCPDSI was found in region II under a subtropical monsoon climate (see Table 3). In this region, the precipitation accounted for a relatively large proportion of the entire water cycle. The lowest correlation (0.29) was found in region III under the plateau mountain climate, implying a complex mechanism of droughts in this region. Since vegetation, snow and other terrestrial surface hydrological components are not explicitly processed in the SCPDSI [40], the correlations (0.42) between the two drought indices were the same in the other three regions. By comparing the temporal and spatial distribution of two drought indices and their correlation coefficients (Table 3, Figures 2 and 3), it can be seen that GRACE-DSI could detect the drought events.

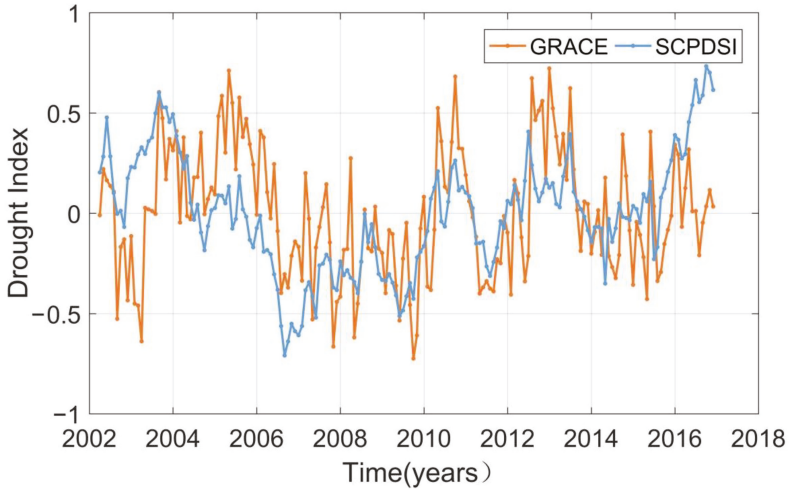



**Figure 2.** Time series of GRACE-DSI and SCPDSI.

**Table 3.** The correlation coefficients between GRACE-DSI and SCPDSI.

| Area | I | II | III | IV | V | MC |
|---|---|---|---|---|---|---|
| Correlation Coefficient | 0.42 | 0.72 | 0.29 | 0.42 | 0.42 | 0.66 |

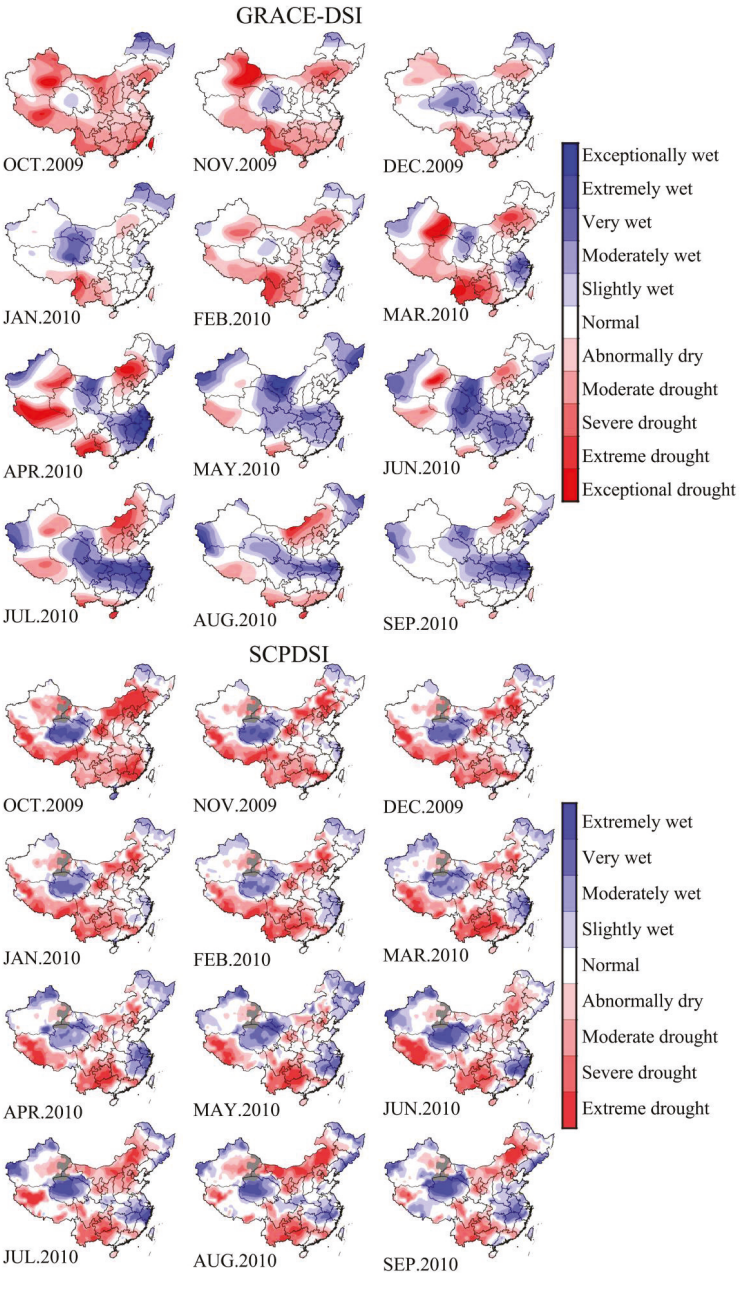



**Figure 3.** Spatial distribution of GRACE-DSI and SCPDSI concerning the October 2009–September 2010 period.

*4.2. The Analysis of Factors Affecting GRACE-DSI*

To study the impact of SM, PPT, ET and runoff on drought events, we calculated the correlation coefficients between GRACE-DSI/SCPDSI and SM, PPT, ET and runoff anomaly. The anomaly signals were extracted according to Equation (4). For a more intuitive comparison and analysis, the following will elaborate on different climate regions. As the tropical monsoon climate region is very small, it is not conducive to the detection of GRACE satellites; therefore, it is considered here.

4.2.1. Temperate Monsoon Climate (Region I)

Figure 4 shows the time series of GRACE-DSI and SM, PPT, ET and runoff anomaly. The SM and runoff had a relatively consistent change trend with GRACE-DSI and the consistency of SM with GRACE-DSI was higher. However, the time series of GRACE-DSI and SM had opposite change trends. This is also supported by the correlation coefficient results (Table 4). The above results indicate that the SM is the most important factor affecting the GRACE-DSI. This is consistent with the definition of hydrological drought—that is, an imbalance between the supply and demand of soil water storage causes a drought event [49].

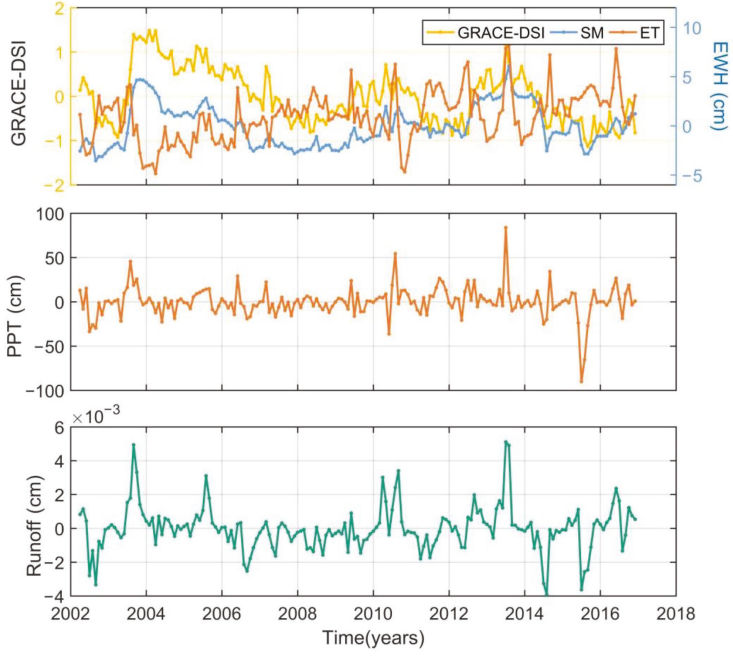



**Figure 4.** The time series of GRACE-DSI, PPT, SM, runoff and ET anomaly under the temperate monsoon climate.

**Table 4.** The correlation coefficients between GRACE-DSI and hydrological components.

| Hydrological Component | PPT | SM | Runoff | ET |
|---|---|---|---|---|
| GRACE-DSI | 0.18 | 0.75 | 0.37 | −0.61 |

To analyze the interaction between the terrestrial hydrological components under the temperate monsoon climate, the correlation coefficients between SM, PPT, ET and runoff were calculated (Table 5). Table 5 shows that there was a strong correlation between PPT

and runoff, and between runoff and SM, implying an obvious water transport channel between PPT, SM and runoff—that is, PPT affects runoff, runoff affects SM. Considering this with Table 4, we can see that the connection between PPT and GRACE-DSI is not strong, but PPT exerts the effect on GRACE-DSI through runoff and SM. Considering this alongside the strong correlation between ET and GRACE-DSI, it explains that PPT and ET are the main factors affecting the occurrence of drought under this climate. Since the correlation coefficient between SM and GRACE-DSI was larger than that between ET and GRACE-DSI, it can be said that PPT has a larger impact on drought than ET. In a word, PPT is the mainstay and ET is the supplement during the formation of drought events.

**Table 5.** The correlation coefficients between different hydrological components.

| Hydrological Component | PPT | SM | Runoff | ET |
|---|---|---|---|---|
| PPT | 1 | 0.29 | 0.61 | 0.32 |
| SM | 0.29 | 1 | 0.54 | −0.31 |
| Runoff | 0.61 | 0.54 | 1 | 0.13 |
| ET | 0.32 | −0.31 | 0.13 | 1 |

Usually, PPT and ET are vulnerable to extreme weather, so it is necessary to consider the impact of extreme weather on drought events. Figures 5 and 6 show the performance of the time series of GRACE-DSI, PPT and ET during the ENSO and NAO events, respectively. From Figure 6, a total of five El Niño events and five La Niña events occurred during the study period. The five El Niño events caused abnormal decreases in PPT, while three El Niño events also led to abnormal increases in ET. Previous studies [50,51] indicate that because of the abnormal decrease of sea surface temperature in the Western Pacific (El Niño event), the East Asian summer monsoon weakened, causing the Western Pacific Subtropical High and rain belt to move southward, resulting in less PPT and higher temperatures in Northern China. The results in this paper provide scientific support for these results. Due to less PPT and more ET, severe drought events occurring during the El Niño events from June 2002 to February 2003 and from April 2015 to April 2016. While the two El Niño events from July 2004 to January 2005 and from September 2006 to January 2007 caused an abnormal decrease in PPT, the ET also showed an abnormal decrease. The drought did not occur under the mutual offset of PPT and ET. The El Niño events from July 2009 to March 2010 led to less PPT and more ET, but perhaps because of the degree of PPT reduction and the minor increase in ET, the combined effect was not enough to cause a drought. According to the above analysis, we found that an El Niño event can indeed cause an abnormally low PPT and an increase in ET, but its impact on PPT is slightly greater than that on ET. Whether an El Niño event will cause drought is the result of its combined effect on PPT and ET. The intensity of this effect needs to reach a certain level in order to induce drought.

Three La Niña events led to an abnormal increase in PPT, and two La Niña events caused an abnormal reduction in ET. Ma [52] indicates that when a La Niña event occurs, the effects are simply the opposite of an El Niño event. At the time, the Western Pacific Subtropical High and rain belt moved northward with the strengthening of the East Asian Monsoon. Northern China showed higher PPT and higher temperature. It can be seen from Figure 6 that the La Niña event mainly affected PPT. Drought events occurred during the La Niña events from July 2007 to June 2008 and from August 2016 to December 2016. According to the previous studies, the cause of drought from 2007 to 2008 was that the northern part of China was in an interdecadal climate with high temperature and low PPT at that time, which caused a high probability of drought in Northeast China and North China [53]. The other drought event was the result of the interaction between the abnormal high pressure in Baikal Lake and Central Siberia and the abnormal low pressures in North China [54]. Due to the interaction of the above-mentioned high and low pressure, the central and eastern regions of China were controlled by the extremely strong dry and cold air flow. Under the control of the air flow, the water vapor transport was reduced, which in

turn caused an abnormal decrease in PPT. This explains that the two drought events above had little to do with the La Niña events.

Under the temperate monsoon climate, the ENSO cycle mainly affects the location of the Western Pacific Subtropical High through the strength of the East Asian Monsoon. The location of Western Pacific Subtropical High determines the amount of PPT and the temperature. The temperature affects the amount of ET. Less PPT and more ET will increase the probability of drought during an El Niño event. The situation is the opposite during a La Niña event.

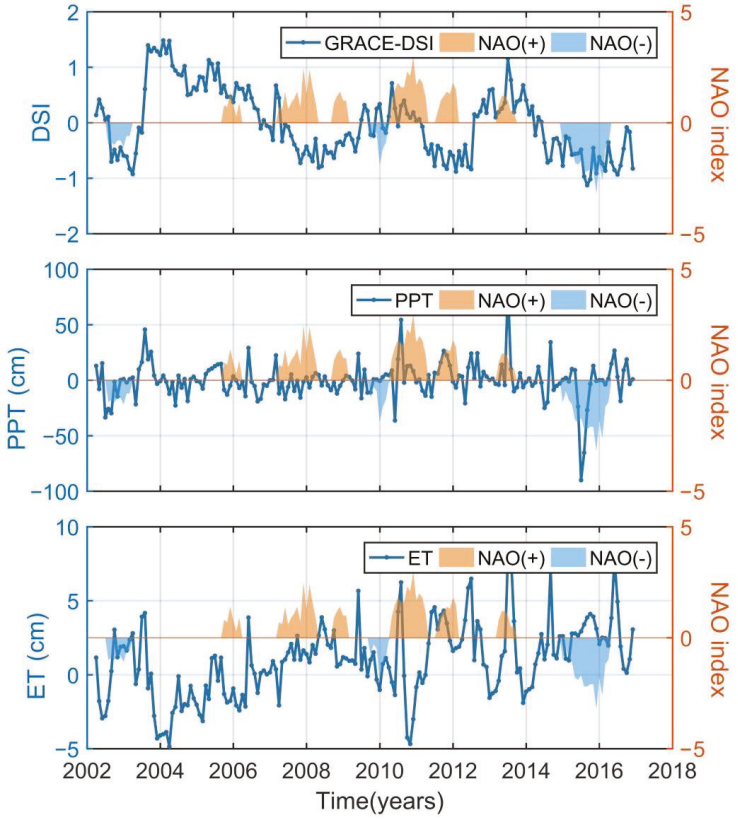



**Figure 5.** GRACE-DSI, PPT anomaly and ET anomaly during NAO events.

The relationship between NAO and GRACE-DSI, PPT and ET were analyzed in this paper (see Figure 5). There were three negative NAO events and seven positive NAO events during the study period. We found that the drought events occurred during the negative NAO events, and they were all caused by less PPT and more ET. Wu et al. [55] indicate that the NAO index has an inverse relationship with the range of Siberian High. When the NAO index was abnormally low, the Siberian High enhanced and its impact scope expanded, which cause the rain belt to move southward and there was less PPT and more ET in Northern China. This coincides with the results in this paper.

When the NAO index was abnormally high, the situation was the opposite. However, from Figure 5, it can be observed that there were three drought events during the positive NAO events. Among them, the drought event from August 2007 to August 2008 was caused by the background of climate, as explained in the previous section, and the one from October 2008 to February 2009 was affected by an El Niño event, indicating that the impact of the El Niño event may have exceeded the positive NAO event. However, the

specific formation mechanism of the above drought event is relatively complicated, and conclusions cannot be drawn regarding this as of yet. The drought event from August 2011 to February 2012 was mainly caused by abnormal atmospheric circulation. At the same time, the Western Pacific Subtropical High was located to the south, which was not conducive to the transportation of water vapor. In addition, there was a strong sinking movement and low humidity in this region, which is conducive to the development and continuation of drought [56].

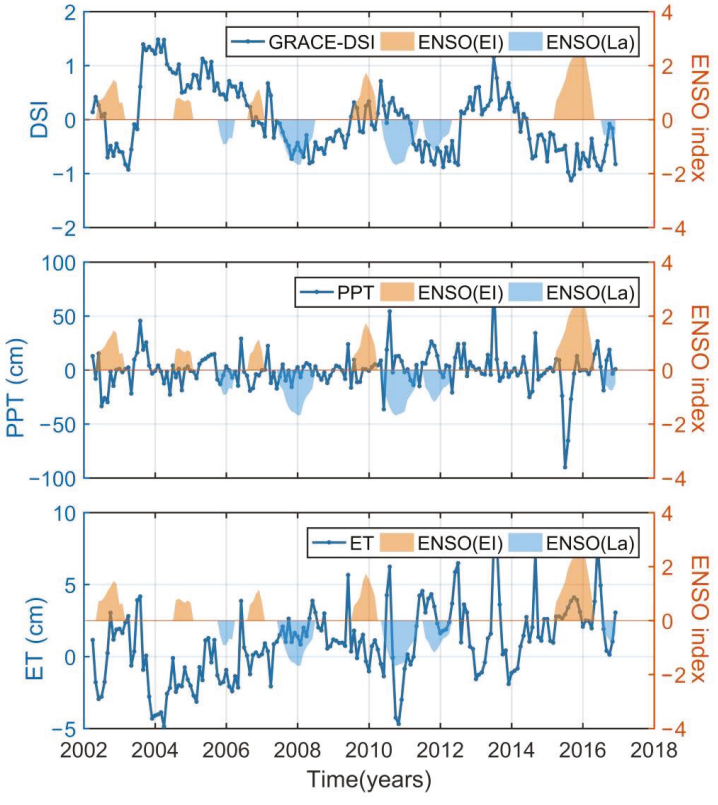



**Figure 6.** GRACE-DSI, PPT anomaly and ET anomaly during ENSO events (El, El Niño, orange; La, La Niña, blue).

Ineson et al. and Graf et al. [57,58] indicate that ENSO signal is propagated to the stratosphere by upward movement, and then transmitted to the North Atlantic region through the "subtropical bridge" mechanism in the stratosphere, which causes the NAO response. El Niño is a negative NAO event, while La Niña is a positive NAO event. This is consistent with the significant negative correlation (−0.88) between ENSO and NAO indices in this paper. Comparing Figures 5 and 6, we can see that there were seven ENSO events accompanied by NAO events. Among the seven events mentioned above, there were three El Niño events, and the negative NAO events occurred at the same time; additionally, four La Niña events and positive NAO events occurred together. The above results provide strong data support for Chen et al. [59]. However, there were three ENSO events that did not cause corresponding NAO events. This may be because these three ENSO events were not Central Pacific (CP) events. According to the study results of Zhang et al. [60], there is a significant relationship between CP ENSO events and NAO events.

4.2.2. Subtropical Monsoon Climate (II)

Figure 7 shows the time series of GRACE-DSI, PPT, SM, runoff and ET anomaly under the subtropical monsoon climate. We found that GRACE-DSI, SM and runoff had similar change trends, which was also confirmed by the correlation coefficient results (Table 6). Unlike the results under the temperate monsoon climate, the correlation coefficients indicate that there was no significant correlation between GRACE-DSI and ET, but GRACE-DSI had a close connection with runoff. This may be related to the sufficient PPT and numerous rivers in this region. We calculated the correlation coefficients between the four hydrological components (Table 7). It can be seen that the way in which PPT affected SM was the same as in the temperate monsoon climate, and the impacts of PPT on runoff, and of runoff on SM, were much greater than was observed in the temperate monsoon climate. This indicates that PPT plays a leading role in drought events, and the impact of ET is small under the subtropical monsoon climate.

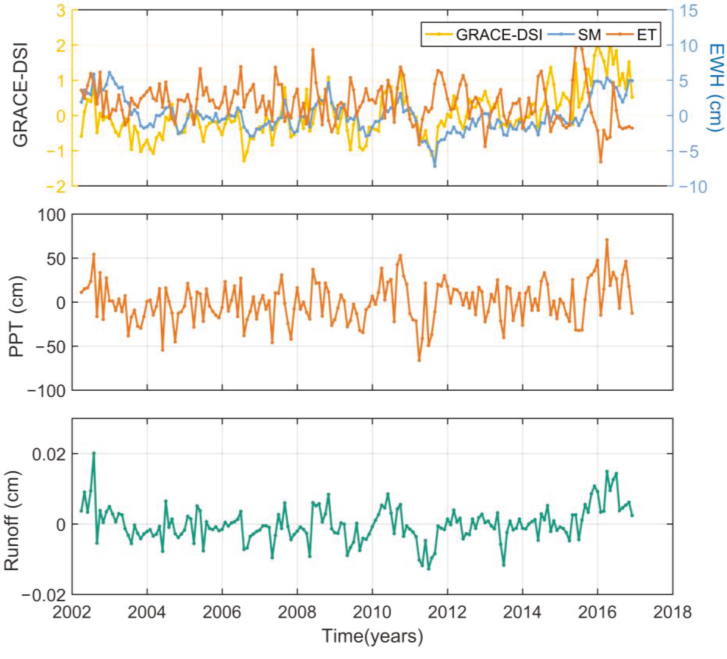



**Figure 7.** The time series of GRACE-DSI, PPT anomaly, SM anomaly, runoff anomaly and ET anomaly under the subtropical monsoon climate.

**Table 6.** The correlation coefficients between GRACE-DSI and hydrological components.

| Hydrological Component | PPT | SM | Runoff | ET |
|---|---|---|---|---|
| GRACE-DSI | 0.47 | 0.65 | 0.62 | −0.24 |

**Table 7.** The correlation coefficients between different hydrological components.

| Hydrological Component | PPT | SM | Runoff | ET |
|---|---|---|---|---|
| PPT | 1 | 0.40 | 0.72 | 0.29 |
| SM | 0.40 | 1 | 0.73 | −0.23 |
| Runoff | 0.72 | 0.73 | 1 | 0.11 |
| ET | 0.29 | −0.23 | 0.11 | 1 |

Similarly, to discuss the influence of extreme climate on drought events and PPT, the change condition of GRACE-DSI and PPT anomaly time series during ENSO and NAO events are shown in Figures 8 and 9. Figure 8 shows that there were drought events during three La Niña events. These drought events were caused by low PPT. Previous studies have shown that the PPT in Southeastern China was greater than normal due to the southward shift of the PPT belt during La Niña events, while the situation was the opposite during the El Niño events [50,61]. There was no reduction in PPT during the other two La Niña events, and so drought events did not appear. This can be attributed to the large amount of PPT brought by typhoons [62]. However, drought events occurred during three El Niño events, and these three drought events occurred from July 2004 to January 2005, from September 2006 to January 2007 and from July 2009 to May 2010. The first drought event was mainly due to a lack of PPT caused by a lack of tropical cyclones [63]. The second drought event was mainly caused by the control of the Subtropical High in Southern China, and the increase and continuation of the Subtropical High in 2006 were closely related to the strengthening of atmospheric convection in the South China Sea and the abnormal heating field in the Bengal Bay [64]. The main reason for the drought from 2009 to 2010 was abnormal circulation. The Western Pacific Subtropical High was stronger than usual, and as a result the Indian Ocean water vapor was not transported to Southwest China. Therefore, there was less PPT in the region [65].

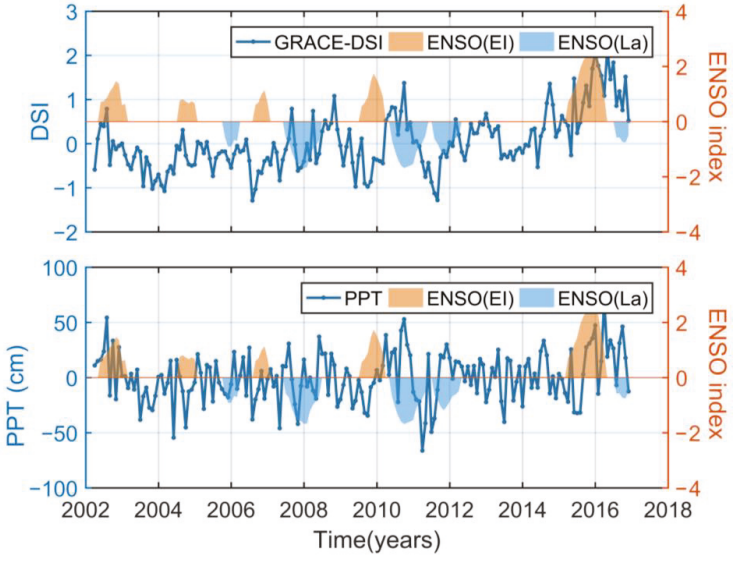



**Figure 8.** GRACE-DSI and PPT anomaly during ENSO events (El, El Niño; La, La Niña).

From Figure 9, it can be observed that the drought events occurred during four positive NAO events, which were caused by low PPT. A previous study indicates that a positive NAO event caused the rain belt to move north, so there was less PPT in this region [55]. The positive NAO event from October 2008 to February 2009 caused a reduction in PPT, but the peak and average values of GRACE-DSI were greater than 0 because of two large-scale waves in the winter 2008, which made the temperature lower than usual [66]. Therefore, severe drought did not appear under the interaction of reduced PPT and low temperature. There was only one drought during a negative NAO event from October 2009 to March 2010, which was caused by the Western Pacific Subtropical High. The specific reason is explained in the previous paragraph.

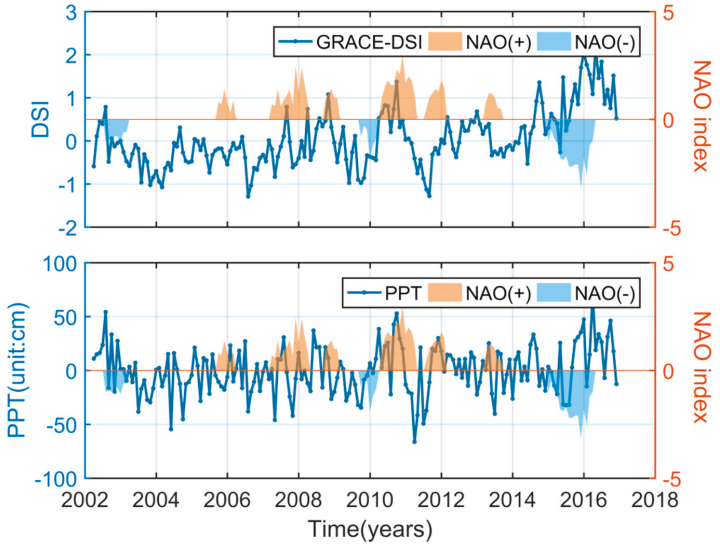



**Figure 9.** GRACE-DSI and PPT anomaly during NAO events.

4.2.3. Plateau Mountain Climate (III)

The time series of GRACE-DSI, PPT, SM, ET and runoff anomaly under the plateau mountain climate are shown in Figure 10. Only the ET and GRACE-DSI had a significant correlation, and it was negative. This is also confirmed by the results in Table 8. The change trends of runoff and PPT anomaly were relatively smooth. Table 9 shows that the correlations between the four hydrological components (SM, PPT, SM and runoff) were not strong, indicating that ET is the main factor causing drought events under the plateau mountain climate.

Figures 11 and 12 show the time series of GRACE-DSI and ET anomaly during ENSO and NAO events. Figure 11 shows that drought events occurred during all five El Niño events. There were four drought events caused by higher ET. Xu et al. indicated that ENSO events affect the surface temperature of the Tibetan Plateau by adjusting the strength of the Indian Ocean Monsoon [67]. In an El Niño year, the Indian Ocean Monsoon weakens and the surface temperature rises, leading to an increase in the possibility of drought events. In La Niña years, the opposite is true. However, drought events occurred during two La Niña events. The first drought event from July 2007 to June 2008 was due to a stronger Middle East Subtropical High. The warm and humid airflow from Bengal Bay could not reach the Tibetan Plateau under the control of this high [68]. The second drought from August to December 2016 was caused by the southward movement of an abnormal continental warm high and the northward movement of the Western Pacific Subtropical High. Under the influence of the above two high pressures, there was long-term, large-scale, sunny and hot weather in this region [54].

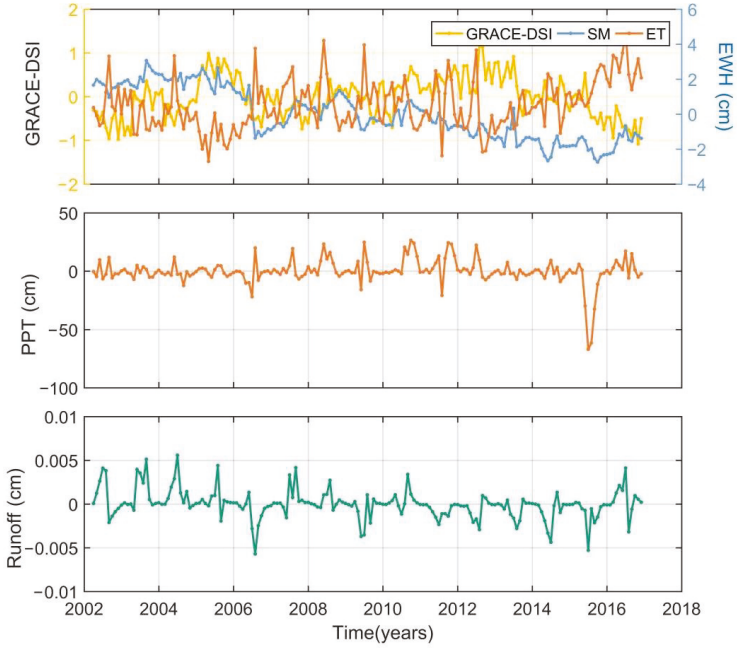



**Figure 10.** The time series of GRACE-DSI/SCPDSI, PPT anomaly, SM anomaly, runoff anomaly and ET anomaly under the plateau mountain climate.

**Table 8.** The correlation coefficients between drought indices and hydrological components.

| Hydrological Component | PPT | SM | Runoff | ET |
|---|---|---|---|---|
| GRACE-DSI | 0.12 | 0.29 | −0.01 | −0.65 |

**Table 9.** The correlation coefficients between different hydrological components.

| Hydrological Component | PPT | SM | Runoff | ET |
|---|---|---|---|---|
| PPT | 1 | 0.10 | 0.17 | 0.30 |
| SM | 0.10 | 1 | 0.43 | −0.25 |
| Runoff | 0.17 | 0.43 | 1 | −0.01 |
| ET | 0.30 | −0.25 | −0.01 | 1 |

Figure 12 shows that drought events occurred during all three negative NAO events, resulting from the southward movement of the Siberian High [55]. The opposite was true during positive NAO events. However, the drought event from April 2007 to April 2008 appeared during a positive NAO event. The reason for this drought was explained in the previous paragraph.

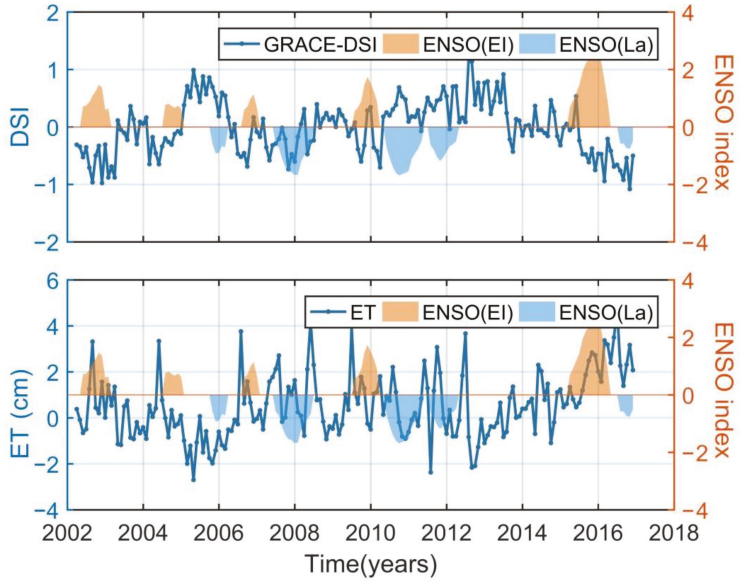



**Figure 11.** GRACE-DSI and ET anomaly compared with ENSO index (El, El Niño; La, La Niña).

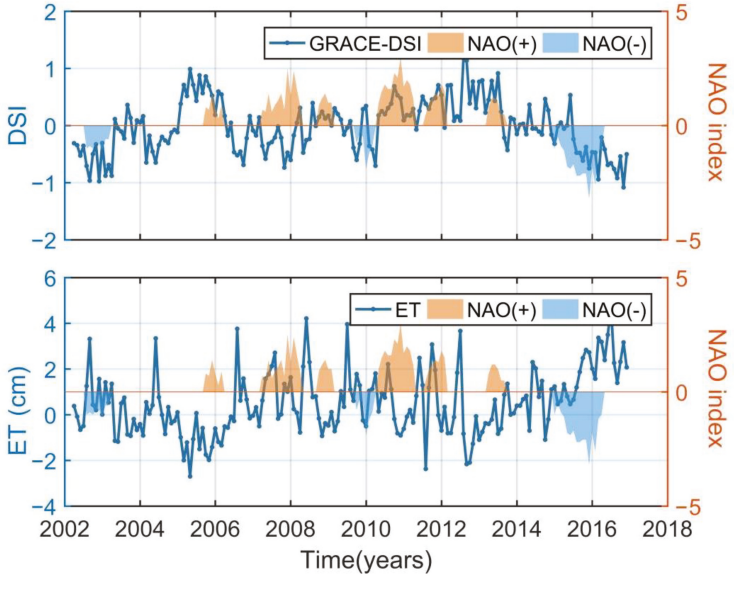



**Figure 12.** GRACE-DSI and ET anomaly compared with NAO index.

4.2.4. Temperate Continental Climate (IV)

Figure 13 compares the time series of GRACE-DSI and PPT, SM, ET and runoff anomaly under the temperate continental climate. We found that the GRACE-DSI had a significant correlation with both ET and SM. Additionally, the change trend of the PPT anomaly was relatively stable. The absolute value of the correlation coefficient between ET and GRACE-DSI (0.76) was greater than that of GRACE-DSI and PPT (0.57) (Table 10). This indicates that the impact of ET on drought events is greater than that of PPT in the

temperate continental climate in this climate. We calculated the correlation coefficients between the four hydrological components (SM, ET, PPT and runoff), as shown in Table 11. The results also show the same relationships between PPT, runoff and SM.

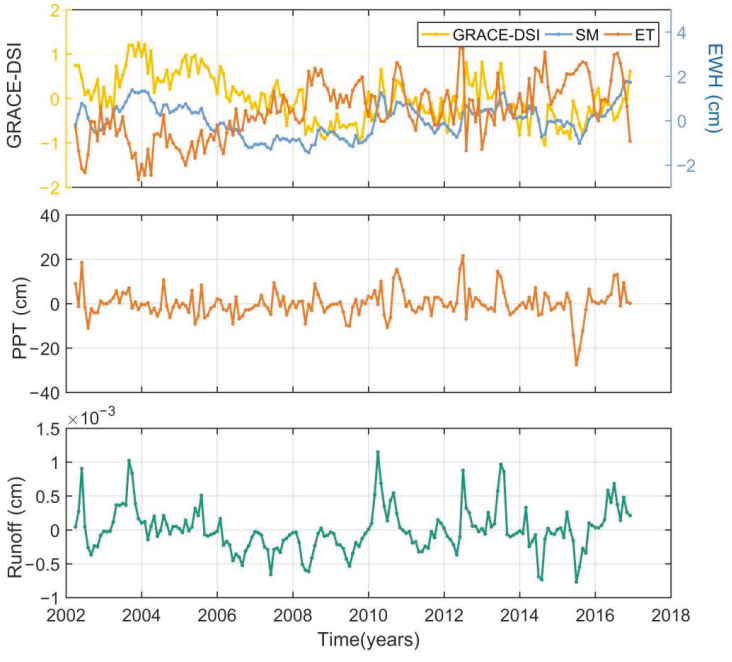



**Figure 13.** The time series of GRACE-DSI/SCPDSI, PPT anomaly, SM anomaly, runoff anomaly and ET anomaly under the temperate continental climate.

**Table 10.** The correlation coefficients between drought indices and hydrological components.

| Hydrological Component | PPT | SM | Runoff | ET |
|---|---|---|---|---|
| GRACE-DSI | 0.12 | 0.57 | 0.37 | −0.76 |

**Table 11.** The correlation coefficients between different hydrological components.

| Hydrological Component | PPT | SM | Runoff | ET |
|---|---|---|---|---|
| PPT | 1 | 0.29 | 0.53 | 0.21 |
| SM | 0.29 | 1 | 0.66 | −0.35 |
| Runoff | 0.53 | 0.66 | 1 | −0.15 |
| ET | 0.21 | −0.35 | −0.15 | 1 |

The time series of GRACE-DSI, PPT and ET anomaly during ENSO and NAO events are shown in Figures 14 and 15. Figure 14 shows that drought events occurred during three El Niño events. The previous studies show that the influence of ENSO events in the temperate continental climate region is basically the same as that in the temperate monsoon climate region [50,51]. However, there was no drought during the other two El Niño events, because the degree of anomalous decrease in PPT was small and the ET was less than usual. Drought events occurred during two La Niña events. The Subtropical High that moved eastward from Western or Central Asia under the influence of atmospheric circulation controlled this region and was the main factor causing the drought event from July 2007 to June 2008 [69]. The drought event from August 2011 to April 2012 was mainly

caused by abnormal atmospheric circulation. As the Subtropical High was weaker than usual, the region was mainly controlled by cold air, and warm and humid air currents could not reach the area. This led to reduced PPT and drought [56].

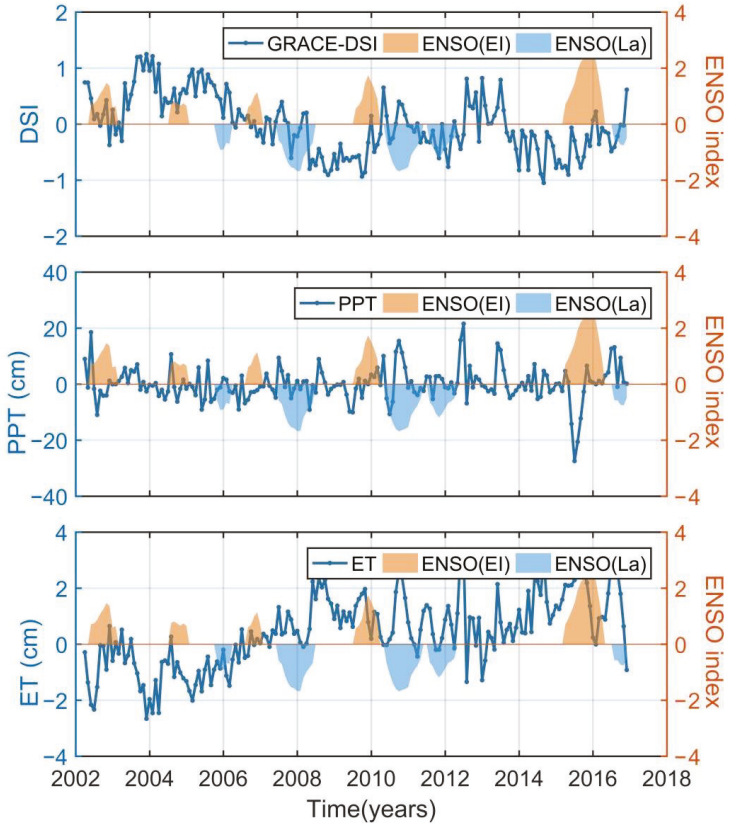



**Figure 14.** GRACE-DSI, PPT and ET anomaly compared with the ENSO index (El, El Niño; La, La Niña).

From Figure 15, two drought events appeared during negative NAO events. Although the PPT was less than usual, ET was also less; this is why there was no drought during the negative event from August 2008 to March 2003. Drought only appeared during two positive NAO events. The drought from July 2015 and April 2016 was affected by El Niño events. As the region was affected by the Subtropical High, the cold air flowed southward. At the same time, due to the influence of the sinking airflow, the transportation of water vapor in the south was blocked. This caused a drought event from October 2008 to February 2009 [70].

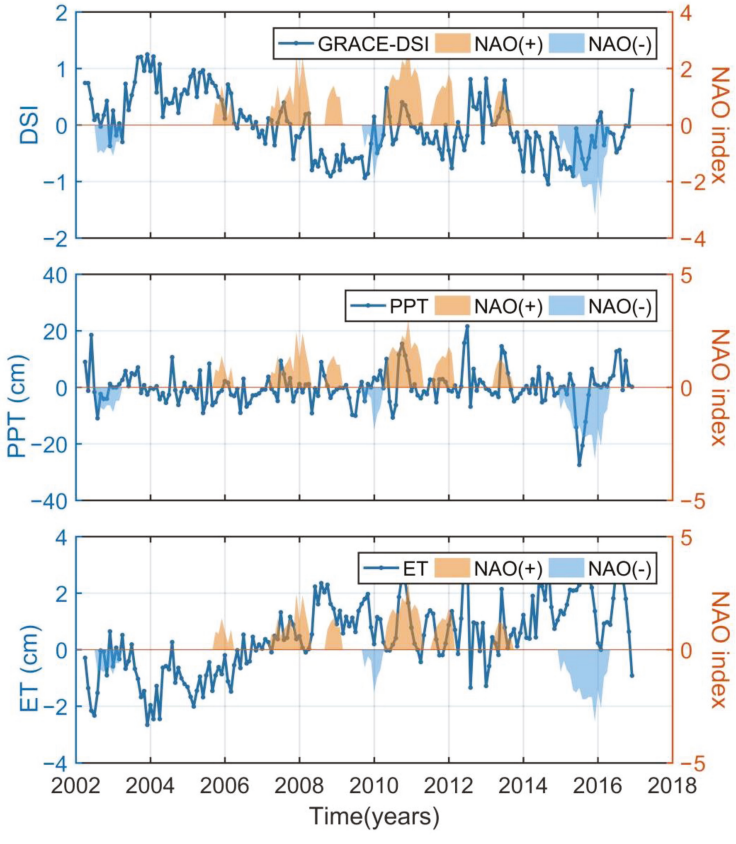



**Figure 15.** GRACE-DSI, PPT and ET anomaly compared with the NAO index.

## 5. Discussion

PPT and ET are the two decisive factors leading to drought events. Extreme climate is one of the factors causing abnormal PPT and ET. The purpose of our study is to discuss the influence of extreme climate on GRACE-DSI, PPT and ET in different climate regions. According to the results in Section 4, extreme climate led to drought by affecting PPT and temperature. Table 12 shows relationship between extreme climate and both PPT and ET anomaly in four climate types. From Table 12, it can be seen that ENSO index was negatively correlated with PPT anomaly under the temperate continental climate and temperate monsoon climate. Meanwhile, there was a positive correlation between ENSO index and PPT anomaly under the subtropical monsoon climate (see the analysis in Section 4). The previous studies show that East Asian Monsoon weakens during El Niño events, which causes the Western Pacific Subtropical High to weaken and move to the south. When this occurs, the boundary between cold air and warm and humid air in MC moves further southward than in regular years. This increases the probability of excessive PPT in South China, while it is likely to lead to reduced PPT in North China. During La Niña events, the situation is the opposite [51,61,71].

**Table 12.** The relationship between ENSO and NAO indices and GRACE-DSI, PPT and ET anomaly.

| Correlation Coefficient | I | II | III | IV |
|---|---|---|---|---|
| ENSO vs. PPT anomaly | −0.30 | 0.27 | - | −0.30 |
| ENSO vs. ET anomaly | 0.08 | - | 0.22 | 0.08 |
| NAO vs. PPT anomaly | 0.23 | −0.29 | - | 0.09 |
| NAO vs. ET anomaly | −0.11 | - | −0.25 | −0.10 |

Table 12 shows that ENSO index and ET anomaly were positively correlated in the temperate continental climate and temperate monsoon climate. Zuo indicated that the influence of ENSO events on the temperature in China is mainly concentrated in North China. The average maximum temperature was observed to be higher in the summer and autumn than normal during El Niño events, while one was lower during La Niña events [66]. Abnormal temperature changes are directly reflected in the ET. The PPT and ET are the two important hydrological components in the terrestrial water cycle. Therefore, abnormal changes in PPT and ET lead to abnormalities in TWSC, which in turn affects the GRACE-DSI. According to Table 12, the impact of extreme climate on PPT was greater than that on ET.

Due to the geographical location, the climate of the plateau mountain climate region is more strongly affected by the Indian Ocean Monsoon. Although blocked by the Himalaya, the Indian Ocean Monsoon still brings some warm and humid air to this region. In El Niño years, the surface temperature of the Qinghai-Tibet Plateau is relatively high due to the weakening of the Indian Ocean Monsoon, resulting in an increase in ET. In La Niña years, there is less ET in this region [65]. Therefore, the results in Table 12 also confirm the above conclusion. As the region with temperature continental climate is located in the hinterland of the continent and far from the ocean, it is difficult for warm humid air to reach this region, causing the region to be mostly semi-arid and arid. Therefore, the biggest difference between this region and the temperature monsoon climate is that ET plays a key role in the entire terrestrial water cycle.

## 6. Conclusions

To discuss the cause of drought events under different climatic conditions, we took four climatic regions in MC as an example to study the influences of various hydrological components and extreme climate on drought events using GRACE-DSI data. Firstly, we compared the temporal and spatial distribution of GRACE-DSI and SCPDSI and calculated the correlation coefficient between these two indices. The results indicate that GRACE-DSI and SCPDSI have similar temporal and spatial distribution and are strongly correlated (0.66) in MC. This proves that GRACE-DSI can detect drought events in MC. Secondly, we studied the influence of drought-related factors and extreme climate on drought events by analyzing the relationship between GRACE-DSI and anomalous changes of SM, PPT, ET and runoff during ENSO and NAO events. The results show that PPT and ET are the main causative factors of drought events. However, they play different roles in different climate regions. PPT plays a major role in temperate monsoon climate and subtropical monsoon climate regions, while ET plays a dominant role in the other two climate regions. ENSO and NAO events first affected the monsoon and regional high, and changes in the monsoon and regional high directly affected the changes of PPT and temperature in a given region. The changes of temperature affected the amount of ET. Therefore, extreme climate has a very important influence on regional drought events. However, the influence of extreme climate on different climatic regions is different depending on the different geographical locations and influence mechanisms.

Our research helps to further reveal the causes of drought events, and provides a reference for drought research in other similar and different climate regions. We mainly focused on the influence of climatic factors on drought events in this paper. Therefore, future work will mainly study the influence of human factors on drought events and the

formation mechanisms of drought in key regions, such as the North China Plain, Northwest China, the Ganges River basin, the Middle East, etc.

**Author Contributions:** Conceptualization, Z.L.; methodology, L.C.; software, C.Z. and L.C.; validation, L.C. and C.Y.; data curation, L.C. and Z.L.; writing—original draft preparation, L.C. and Q.L. writing—review and editing, L.C. and X.W.; funding acquisition, C.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Key R&D Program of China (Grant No. 2018YFC1503503), the National Natural Science Foundation of China (Grant No. 41931074, 42061134007, 42004013), Foundation of Young Creative Talents in Higher Education of Guangdong Province (Grant No. 2019KQNCX009) and Open Fund of Guangxi Key Laboratory of Spatial Information and Geomatics (19-050-11-03).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** GRACE RL06 data: ftp://icgem.gfz-potsdam.de/01_GRACE/CSR/CSRRelease06 (accessed on 16 September 2021); GLDAS model data: https://disc.gsfc.nasa.gov/datasets?keywords=gldas&page=1 (accessed on 16 September 2021); In-situ PPT data: http://data.cma.cn (accessed on 16 September 2021); SCPDSI data: (https://crudata.uea.ac.uk/cru/data/drought/ (accessed on 16 September 2021); ENSO index data: http://www.cpc.ncep.noaa.gov/data/indices/ (accessed on 16 September 2021); NAO index data: https://www.cpc.ncep.noaa.gov/products/pre-cip/CWlink/pna/nao.shtml (accessed on 14 September 2021).

**Acknowledgments:** We are grateful to the Center of Space Research (CSR) for providing the monthly GRACE gravity field solutions, and to the Goddard Space Flight Center for providing the monthly GLDAS-2.1 data, to the China National Meteorological Science Data Center for providing the monthly precipitation and temperature products, to the Climatic Research Unit at university of East Anglia for providing the SCPDSI data and to the National Oceanic and Atmospheric Administration for providing the Extreme weather Index Data.

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

## References

1. Wilhite, D.A. The enigma of drought. In *Drough Assessment, Management, and Planning: Theory and Case Studies*; Kluwer Academic Publishers: Norwell, MA, USA, 1993; pp. 3–15.
2. Breshears, D.D.; Cobb, N.S.; Rich, P.M.; Price, K.P.; Allen, C.D.; Balice, R.G.; Romme, W.H.; Kasten, J.H.; Floyd, M.L.; Belnap, J.; et al. Reginal vegetation die-off in response to global-change-type drought. *Proc. Natl. Acad. Sci. USA.* **2005**, *102*, 15144–15148. [CrossRef] [PubMed]
3. Khan, A.A.; Zhao, Y.J.; Khan, J.; Rahman, G.; Rafiq, M.; UI Moazzam, M.F. Spatial and temporal analysis of rainfall and drought condition in Southwest Xinjiang in Northwest China, using various climate indices. *Earth Syst. Environ.* **2021**, *5*, 201–216. [CrossRef]
4. Houborg, R.; Rodell, M.; Li, B.; Reichle, R.; Zaitchik, B.F. Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) Terrestrial water storage observations. *Water Resour. Res.* **2012**, *48*, W07525. [CrossRef]
5. Tsakiris, G.; Loukas, A.; Pangalou, D.; Vangelis, H.; Tigkas, D.; Rossi, G.; Cancelliere, A. Drought characterization. In *Drought Management Guidelines Technical Annex*; CIHEAM-IAMZ: Zaragoza, Spain, 2007; pp. 85–102.
6. Palmer, W.C. *Meteorological Drought*; US department of Commerce, Weather Bureau: Washington, DC, USA, 1965.
7. McKee, T.B.; Doesken, N.J.; Kleist, J. The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology*; American Meteorological Society: Boston, MA, USA, 1993.
8. Shukla, S.; Wood, A.W. Use of a standardized runoff index for characterizing hydrologic drought. *Geophy. Res. Lett.* **2008**, *35*, L20405. [CrossRef]
9. Sun, Z.L.; Zhu, X.F.; Pan, Y.Z.; Zhang, J.S.; Liu, X.F. Drought evaluation using the GRACE terrestrial water storage deficit over the Yangtza River Basin, China. *Sci. Total Environ.* **2018**, *634*, 727–738. [CrossRef] [PubMed]
10. Thomas, A.C.; Reager, J.T.; Famiglietti, J.S.; Rodell, M. A GRACE-based water storage deficit approach for hydrological drought characterization. *Geophys. Res. Lett.* **2014**, *41*, 1537–1545. [CrossRef]
11. Taply, B.D.; Bettadpur, S.; Ries, J.C.; Thompson, P.F.; Watkins, M.M. GRACE measurements of mass variability in the Earth system. *Science* **2004**, *305*, 503. [CrossRef]
12. Wahr, J.; Swenson, S.C.; Zlotnicki, V.; Velicogna, I. Time-variable from GRACE: First results. *Geophy. Res. Lett.* **2004**, *31*, L1123785. [CrossRef]

13. Chen, J.L.; Wilson, C.R.; Tapley, B.D.; Yang, Z.L.; Niu, G.Y. 2005 drought event in the Amazon River basin as measured by GRACE and estimated by climate models. *J. Geophys. Res.* **2009**, *114*, B05404. [CrossRef]
14. Frappart, F.; Paga, F.; da Sliva, J.S.; Ramillien, G.; Prigent, C.; Seyler, F.; Calmant, S. Surface freshwater storage and dynamics in the Amazon basin during the 2005 exceptional drought. *Environ. Res. Lett.* **2012**, *7*, 44010. [CrossRef]
15. Yao, C.; Luo, Z.; Wang, H.; Li, Q.; Zhou, H. GRACE-derived terrestrial water storage changes in the inter-basin region and its possible influencing factors: A case study of the Sichuan Basin, China. *Remote Sens.* **2016**, *8*, 444. [CrossRef]
16. Long, D.; Yang, Y.; Wada, Y.; Hong, Y.; Liang, W.; Chen, Y.N.; Yong, B.; Hou, A.H.; Wei, J.F.; Chen, L. Deriving scaling factors using a global hydrological model to restore GRACE total water storage changes for China's Yangtze river basin. *Remote Sens. Environ.* **2015**, *168*, 177–193. [CrossRef]
17. Zhang, Z.Z.; Chao, B.F.; Chen, J.L.; Wilson, C.R. Terrestrial water storage anomalies of Yangtze river basin droughts observed by GRACE and connections with ENSO. *Glob. Planet. Chang.* **2015**, *126*, 35–45. [CrossRef]
18. Tang, J.S.; Cheng, H.W.; Liu, L. Assessing the recent drought in southwest China using satellite gravimetry. *Water Resour. Res.* **2014**, *50*, 3030–3038. [CrossRef]
19. Li, Q.; Luo, Z.C.; Zhong, B.; Wang, H.H. Terrestrial water storage changes of the 2010 southwest China drought detected by GRACE temporal gravity field. *Chin. J. Geophys.* **2013**, *56*, 1843–1849. (In Chinese)
20. Chao, B.F.; Gross, R. Changes in the Earth's rotation and Love-degree gravitational field introduced by earthquakes. *J. Geophys. Res.* **1987**, *91*, 569–596.
21. Long, D.; Scanlon, B.R.; Longuevergne, L.; Sun, A.Y.; Fernando, D.N.; Save, H. GRACE satellite monitoring of large depletion in water storage in response to the 2011 drought in Texas. *Geophys. Res. Lett.* **2013**, *40*, 3395–3401. [CrossRef]
22. Yirdaw, S.Z.; Snelgrove, K.R.; Agboma, C.O. GRACE satellite observation of terrestrial moisture changes for drought characterization in the Canadian Prairie. *J. Hydro.* **2008**, *356*, 84–92. [CrossRef]
23. Wang, J.H.; Jiang, D.; Huang, Y.H.; Wang, H. Drought analysis of the Haihe river basin based on GRACE terrestrial water storage. *Sci. World J.* **2014**, *2014*, 578372. [CrossRef]
24. Yi, H.; Wen, L.H. Satellite gravity measurement monitoring terrestrial water storage and drought in the continental United States. *Sci. Rep.* **2016**, *6*, 19909. [CrossRef]
25. Sinha, D.; Syed, T.H.; Famiglietti, J.S.; Reager, J.T.; Thomas, R.C. Charaterizing drought in India using GRACE Observations of terrestrial water storage deficit. *J. Hydrometeor.* **2017**, *18*, 381–396. [CrossRef]
26. Zhao, M.; Gerou, A.; Velicogna, I.; Kimball, J.S. A global gridded dataset of GRACE drought severity Index for 2002–14: Comparison with PDSI and SPEI and a case study of the Australia Millennium drought. *J. Hydrometeor.* **2017**, *18*, 2117–2129. [CrossRef]
27. Wu, T.T.; Zheng, W.; Yin, W.J.; Zhang, H.W. Spatiotemporal characteristics of drought and driving factors based on the GRACE-derived total storage deficit index: A case study in Southwest China. *Remote Sens.* **2021**, *13*, 79. [CrossRef]
28. Panisset, J.S.; Libonati, R.; Gouveia, C.M.P.; Machado-Silva, F.; França, D.A.; França, J.R.A.; Peres, L.F. Contrasting patterns of the extreme drought episodes of 2005, 2010 and 2015 in the Amazon Basin. *Int. J. Climatol.* **2018**, *38*, 1096–1104. [CrossRef]
29. Wang, S.; Zhang, Z. Effects of climate change on water resources in China. *Clim. Res.* **2011**, *47*, 77–82. [CrossRef]
30. Cui, L.L. Research on Drought Events in Mainland China Using Satellite Time-Variable Gravity Field. Ph.D. Thesis, Wuhan University, Wuhan, China, 2021. (In Chinese)
31. Ye, M.; Qian, Z.H.; Wu, Y.P. Spatiotemporal evolution of the droughts and floods over China. *Acta Phys. Sin.* **2013**, *62*, 139203.
32. Cheng, M.; Tapley, B.D. Variations in the Earth's oblateness during the past 28years. *J. Geophys. Res.* **2004**, *109*, B09402.
33. Swenson, S.; Chambers, D.; Whar, J. Estimating geocenter variations form a combination of GRACE and ocean model output. *J. Geophys. Res: Solid Earth.* **2008**, *113*, 194–205.
34. Swenson, S.; Wahr, J. Post-processing removal of correlated errors in GRACE data. *Geophys. Res. Lett.* **2006**, *33*, L025285. [CrossRef]
35. Chambers D, P. Evaluation of new GRACE time-variable gravity data over the ocean. *Geophys Res Lett.* **2006**, *33*, L17603. [CrossRef]
36. Cui, L.L.; Song, Z.; Luo, Z.C.; Zhong, B.; Wang, X.L.; Zou, Z.B. Comparison of terrestrial water storage changes derived from GRACE/GRACE-FO and Swarm: A case study in the Amazon River Basin. *Water* **2020**, *12*, 3128. [CrossRef]
37. Rodell, M.; Houser, P.; Jambor, U.E.A.; Gottschalck, J.; Mitchell, K.; Meng, C.; Arsenault, K.; Cosgrove, B.; Radakovich, J.; Bosilovich, M. The global land data assimilation system. *Bull. Am. Meteorol. Soc.* **2004**, *85*, 381–394. [CrossRef]
38. Rodell, M.; Famiglietti, J.S.; Chen, J.; Seneviratne, S.I.; Viterbo, P.; Holl, S.; Wilson, C.R. Basin scale estimates of evapotranspiration using GRACE and other observations. *Geophys. Res. Lett.* **2014**, *31*, L20504. [CrossRef]
39. Ahi, G.O.; Jin, S.G. Hydrologic mass changes and their implications in Mediterranean-Climate Turkey from GRACE measurements. *Remote Sens.* **2019**, *11*, 120.
40. van der Schrier, G.; Barichivich, J.; Briffa, K.R.; Jones, P.D. A scPDSI-based global data set of dry and wet spells for 1901–2009. *J. Geophys. Res. Atmos.* **2013**, *118*, 4025–4048. [CrossRef]
41. Dai, A.; Lamb, P.J.; Trenberth, K.E.; Hulme, M.; Jones, P.D.; Xie, P. The recent Sahel drought is real. *Int. J. Climatol.* **2004**, *24*, 1323–1331. [CrossRef]
42. Heim, R.R. A review of twentieth-century drought indeices used in the Unite State. *Bull. Am. Meteorol. Soc.* **2002**, *83*, 1149–1165. [CrossRef]

43. Grades of meteorological drought (GB/T 20481-2017). In *National Standards of People's Republic of China*; China Standard Press: Beijing, China, 2017.
44. Chen, W.; Zhong, M.; Feng, W.; Zhong, Y.L.; Xu, H.Z. Effects of two strong ENSO events on terrestrial water storage anomalies in China from GRACE During 2005–2017. *Chin. J. Geophys.* **2020**, *63*, 141–154. (In Chinese)
45. Xi, H.; Zhang, Z.Z.; Lu, Y.; Li, Y. Mass sea level variation in the South China Sea from GRACE, altimetry and model and the connection with ENSO. *Adv. Space Res.* **2019**, *64*, 117–128. [CrossRef]
46. Huang, B.Y.; Banzon, V.F.; Freeman, E.; Lawrimore, J.; Liu, W.; Peterson, T.C.; Smith, T.M.; Thorne, P.W.; Woodruff, S.D.; Zhang, H.M. Extended reconstructed sea surface temperature version 4 (ERSST.v4). Part I: Upgrades and intercomparisons. *J. Clim.* **2015**, *28*, 911–930. [CrossRef]
47. Hurrell, J.W.; Kushnir, Y.; Ottersen, G.; Visbeck, M. An overview of the north atlantic oscillation. In *Geophysical Monograph Series*; American Geophysical Union: Washiongton, DC, USA, 2003.
48. Velicogna, I.; Wahr, J. Measurements of time-variable gravity show mass loss in Antarctica. *Science* **2006**, *311*, 1754–1756. [CrossRef]
49. Zargar, A.; Sadiq, R.; Naser, B.; Khan, F.I. A review of drought indices. *Env. Rev.* **2011**, *19*, 333–349. [CrossRef]
50. Liu, Y.Q.; Ding, Y.H. Influence of ENSO events on weather and climate of China. *J. Appl. Meteorolo. Sci.* **1992**, *3*, 473–481.
51. Ma, B.S. The research on the impact of two types of ENSO events on precipitation in Eastern China. Master's Thesis, Lanzhou University, Lanzhou, China, 2019. (In Chinese)
52. Three reason lead to the drought in Northeast China and North China. Available online: http://www.cma.gov.cn/2011xwzx/.2011xqxxw/2011xylsp/201110/t20111026_115636.html (accessed on 15 May 2008). (In Chinese)
53. Available online: 2011xqxxw/2011xylsp/201110/t20111026_115636.html (accessed on 15 May 2008). (In Chinese)
54. Wang, S.S.; Wang, S.P.; Feng, J.Y. Drought events and its influence in Autumn of 2016 in China. *J. Arid Meteorol.* **2016**, *34*, 1099–1104. (In Chinese)
55. Wu, B.Y.; Huang, R.H. Effects of the extremes in the North Atlantic Oscillation on East Asia Winter Monsoon. *Chin. J. Atmos. Sci.* **1999**, *23*, 641–651. (In Chinese)
56. Duan, H.X.; Wang, S.P.; Feng, J.Y. Drought events and its influence in Autumn of 2011 in China. *J. Arid Meteorol.* **2011**, *30*, 136–146. (In Chinese)
57. Ineson, S.; Scaife, A.A. The role of the stratosphere in the European climate response to EI Niño. *Nat. Geosci.* **2008**, *2*, 32–36. [CrossRef]
58. Graf, H.F.; Zanchettin, D. Central Pacific El Niño, the "subtropical bridge", and Eurasian climate. *J. Geophys. Res.* **2012**, *117*, D01102. [CrossRef]
59. Chen, W.; Lan, X.Q.; Wang, L.; Ma, Y. The combined effects of the ENSO and the Arctic Oscillation on the winter climate anomalies in East Asia. *Chin. Sci. Bull.* **2013**, *58*, 1355–1362. [CrossRef]
60. Zhang, W.J.; Wang, L.; Xiang, B.Q.; He, J.H. Impact of two types of La Niña on the NAO during boreal winter. *Clim. Dyn.* **2015**, *44*, 1351–1366. [CrossRef]
61. Zhang, J. Research on the Impact of ENSO Events on the Climate of China. Master's Thesis, Capital Normal University, Beijing, China, 2001. (In Chinese)
62. China Ministry of Water Resource (CMWR). *China Flood and Drought Disaster Bulletin*; China Water Conservancy and Hydropower Press: Beijing, China, 2016. (In Chinese)
63. Guan, M.Y. Cause analysis of serious drought events in 2004. *J. Meteorol. Res. Appl.* **2007**, *28*, 60–63. (In Chinese)
64. Peng, J.B.; Zhang, Q.Y.; Cholaw, B. On the characteristics and possible cause of a sever drought and heat wave in the Sichuan-Chongqing region in 2006. *Clim. Environ. Res.* **2007**, *12*, 464–474. (In Chinese)
65. Wang, S.P.; Duan, H.X.; Feng, J.Y. Drought events and its influence in Autumn of 2010 in China. *J. Arid Meteorol.* **2010**, *28*, 107–112. (In Chinese)
66. Ma, J. Less precipitation in most parts of China and the obvious impact caused by two cold waves. *Meteorol. Mon.* **2009**, *35*, 120–123. (In Chinese)
67. Xu, H.; Hong, Y.T.; Hong, B.; Zhu, Y.X.; Wang, Y. Influence of ENSO on multi-annual temperature variations at Hongyuan, NE Qinghai-Tibet Plateau: Evidence from of spruce tree rings. *Int. J. Climatol.* **2009**, *30*, 120–126. [CrossRef]
68. Tang, H.B.; Chen, Q. Analysis on the cause of drought in Qinghai province. *J. Water Res. Water Eng.* **2007**, *18*, 89–91. (In Chinese)
69. Tuo, Y. Research on the cause of formation and forecast method of draught in Inner Mongolian. Master's Thesis, Chinese Academy of Agricultural Sciences, Beijing, China, 2006. (In Chinese)
70. Chen, Q.L.; Hua, W.; Xiong, G.M.; Xu, H.; Li, X.R. Analysis on the cause of severe drought in North China. *Arid Zone Res.* **2010**, *27*, 182–187. [CrossRef]
71. Zuo, J.Q. Relationship between AO/NAO and ENSO and Their Impact on Climate Anomalies in China. Ph.D. Thesis, Lanzhou University, Lanzhou, China, 2011. (In Chinese)

water

MDPI

Article

# Spatial and Temporal Pattern of Rainstorms Based on Manifold Learning Algorithm

**Yuanyuan Liu [1,2], Yesen Liu [1,2], Hancheng Ren [1,*], Longgang Du [3], Shu Liu [1], Li Zhang [4], Caiyuan Wang [3] and Qiang Gao [3]**

[1] State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing 100038, China
[2] Key Laboratory of Digital Twin Watershed, Ministry of Water Resources, Beijing 100038, China
[3] Beijing General Hydrology Station, Beijing 100089, China
[4] Shenzhen National Climate Observatory, Shenzhen 518040, China
* Correspondence: rhc_iwhr@163.com

**Abstract:** Identifying the patterns of rainstorms is essential for improving the precision and accuracy of flood forecasts and constructing flood disaster prevention systems. In this study, we used a manifold learning algorithm method of machine learning to analyze rainstorm patterns. We analyzed the spatial–temporal characteristics of heavy rain in Beijing and Shenzhen. The results showed a strong correlation between the spatial–temporal pattern of rainstorms and underlying topography in Beijing. However, in Shenzhen, the spatial–temporal distribution characteristics of rainstorms were more closely related to the source of water vapor causing the rainfall, and the variation in characteristics was more complex and diverse. This method may be used to quantitatively describe the development and dynamic spatial–temporal patterns of rainfall. In this study, we found that spatial–temporal rainfall distribution characteristics, extracted by machine learning technology could be explained by physical mechanisms consistent with the climatic characteristics and topographic conditions of the region.

**Keywords:** manifold learning; machine learning; spatial–temporal distribution of rainstorms; feature extraction; Beijing; Shenzhen

**Citation:** Liu, Y.; Liu, Y.; Ren, H.; Du, L.; Liu, S.; Zhang, L.; Wang, C.; Gao, Q. Spatial and Temporal Pattern of Rainstorms Based on Manifold Learning Algorithm. *Water* **2023**, *15*, 37. https://doi.org/10.3390/w15010037

Academic Editor: Gwo-Fong Lin

Received: 21 October 2022
Revised: 6 December 2022
Accepted: 20 December 2022
Published: 22 December 2022



## 1. Introduction

Highly concentrated rainfall and large precipitation events may cause urban floods, mountain torrents, mud–rock flows, landslides, and other disasters, resulting in substantial economic losses and casualties [1,2]. The temporal and spatial distribution of hourly heavy precipitation has a good correspondence with flood disaster data. The intensity and spatial–temporal distribution of rainstorms show a clear correlation with geological disasters caused by flooding [3–5]. The movement direction of the rainfall center directly impacts the shape of the flood hydrograph and the change in flood peak discharge [6]. Under conditions with the same average rainfall and intensity, rainfall patterns with a rain peak in the middle or rear are more than 30% larger than flood peaks with uniform rain patterns [7,8]. A thorough understanding of the temporal and spatial variation in rainstorm patterns is essential for improving the accuracy of flood forecasting and building a flood disaster prevention system [9,10]. Previous research has contributed to advancements in the study of the spatial–temporal distribution of rainfall.

Some scholars have studied the temporal and spatial distribution of precipitation by using multiple analysis methods, combining rainfall data of different scales. Fung et al. [11] used inverse distance weighting (IDW) and ordinary kriging (OK), geographical weighted regression (GWR) and multi-scale geographical weighted regression (MGWR) to investigate spatiotemporal modeling of rainfall distribution in Peninsular Malaysia. Hitchens et al. [12] study the climatology of heavy rain events using two high-resolution precipitation datasets

that incorporate both gauge observations and radar estimates. Satya et al. [13] analyzed Tropical Rainfall Measuring Mission (TRMM) data with grid spatial resolution of 0.25° × 0.25° to obtain information about the characteristics of rainfall in South Sumatra. Ndiaye et al. [14] analyzes the spatio-temporal distribution of daily rainfall data from 13 stations in the country of Senegal located in the Northwest of Africa. Chaubey, P.K. et al. [15] examined the trend in extreme rainfall events using long-term observed high-resolution gridded rainfall data (1901–2019). Audu, M.O. et al. [16] analyzed the spatial distribution and temporal trends of precipitation and its extremes over Nigeria from 1979–2013 using climate indices in order to assess climatic extremes in the country. Yeung et al. [17] simulated rainfall in the New York area of the United States using meteorological station, radar, satellite, and other observation methods combined with a mesoscale meteorological model. Viglione et al. [18] quantified the temporal and spatial distribution of rainfall. Zoccatelli et al. [19] proposed an index system based on quantitative descriptions of the temporal and spatial distribution of rainfall. Wu et al. [20] analyzed the axially symmetrical precipitation characteristics of landfall typhoons in East China using radar data and historical precipitation data from ground stations. Ngongondo et al. [21] studied the spatial and temporal characteristics of rainfall in Malawi between 1960 and 2006.

These research results are of great significance for understanding the characteristics of rainstorms in various regions. However, static indicators, such as the total amount of rainstorm and distribution of rainfall in various regions, cannot describe characteristics of rainfall changes in time and space well. Understanding dynamic characteristics, such as moving path, range of rainstorm center, is important to predict rainstorm development process, forecast the flood risk area in advance and to effectively preventing the flood and geological disasters caused by rainstorms.

In recent decades, major advancements have been made in artificial intelligence (AI) technology [22]. AI has been applied in multiple areas, including the early identification of disaster risks and water resource management [23–25]. Furthermore, neural network models have been widely used in many fields [26,27]. A cluster-based Bayesian network has been used to predict the longitudinal dispersion coefficient in natural rivers [28], and a deep learning method has been used for spatio-temporal flood prediction [29]. In terms of urban flood management, the integration of AI and numerical simulation models has enabled the rapid prediction of urban floods [30]. In hydrology, machine learning technology and hydrological models have been combined to construct coupled simulations for watershed runoff and sediment [31]. Manifold learning algorithms are practical data processing methods in AI technology that play major roles in high-dimensional array reduction and transfer learning.

This study presents a novel application of these machine learning algorithms in the analysis of spatial–temporal distribution characteristics of rainstorms. We examined the spatial–temporal distribution characteristics of rainfall in Beijing and Shenzhen.

## 2. Materials and Methods

Beijing is located in the inland region of North China and is surrounded by mountains on three sides—in the west, north, and northeast. The annual distribution of precipitation varies seasonally. Rainfall during the flood season (from June to September) accounts for approximately 85% of the total annual precipitation. Heavy rain often occurs in late July and early August and is likely to cause flood disasters. Shenzhen is located in the coastal area of South China. The distribution of rainfall here varies, both seasonally and geographically, throughout the year. Rainfall during the flood season (April to September) accounts for 86% of the total annual rainfall. The uneven spatial and temporal distribution of rainfall contributes to alternating droughts and floods [32,33]. This study evaluated 5 min rainfall data from 122 automatic weather stations in Beijing from 1999 to 2020 and 63 automatic weather stations in Shenzhen from 2008 to 2021. We analyzed and extracted the spatial–temporal rainfall distribution data for the two regions. The locations of the stations are shown in Figure 1.

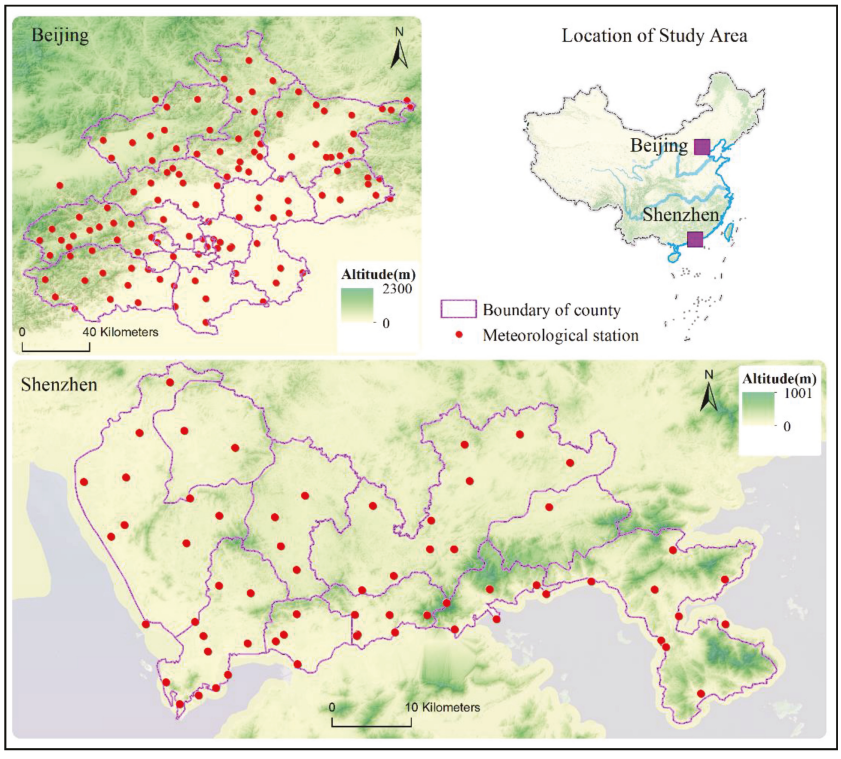



**Figure 1.** Study area and locations of stations whose data were used in this study.

In this study, 5 min rainfall measurements were used as samples, although many factors affected the accuracy of them, such as the wind conditions at the measurement site, which can be challenging during tropical rainfall. There are unreasonable values in the 5 min rainfall measurement data [34,35]. So, prior to analysis, the historical rainfall data were cleaned and screened to eliminate inaccurate data. The cleaning standards were as follows:

1. If the rainfall at a single station exceeded 10 mm in 5 min and existed in isolation, and there was no rainfall at the same station 30 min before and after the observation, it was considered an unreasonable record;
2. If the rainfall at a single station exceeded 10 mm in 5 min, but the observed data of other rain-measuring stations within a range of 5 × 5 km of the station was 0, it was considered an unreasonable record;
3. In the case of unreasonable records from a single station, the data were compared to the rainfall isosurface map of the period. If the data from the station were confirmed to be unreasonable, the interpolation results of surrounding stations within a range of 5 × 5 km were used to replace the unreasonable records of that station.

After cleaning, the data were divided into rainfall events to select samples for further analysis. The screening criteria were as follows:

1. Rainfall events were first identified. If the 5 min rainfall at all the stations was less than 0.1 mm over four consecutive hours, it was not considered effective rainfall. Two independent rainfall events were eliminated according to this standard.
2. Rainstorm samples were screened according to the yellow rainstorm warning standards of Beijing and Shenzhen. Rainstorm events were selected for further analysis.

According to the above criteria, we synthesized the historical rainfall data from Beijing and Shenzhen and categorized the rainstorm events based on duration (12 h, 24 h, and 72 h) in preparation for further analysis.

### *2.1. Methods and Procedures*

In this study, rainstorms were divided into discrete events, and a high-dimensional matrix was constructed to describe the spatial–temporal characteristics of rainfall. The spatial–temporal dynamic development characteristics of multiple rainfall events were described as shown in Equations (1)–(3):

$$\Omega = \{X_1, X_2, \ldots X_N\} \tag{1}$$

$$X_i = \begin{bmatrix} H_{11} & H_{21} \cdots & H_{s1} \\ H_{12} & H_{22} \ldots & H_{s2} \\ \vdots & \vdots & \vdots \\ H_{1m} & H_{2m} \ldots & H_{sm} \end{bmatrix}, \tag{2}$$

where $\Omega$ is the historical rainstorm samples set, including N rainstorm fields, $X_i$ is the proportion matrix for the $i^{th}$ rainfall event, and $H_{it}$ is the percentage of the rainfall at the $i^{th}$ rainfall station to the total rainfall at all the stations in the time $t$ during the $i^{th}$ rainfall event, i.e.,

$$H_{it} = \frac{R_{it}}{\sum_{k=1}^{s} R_{kt}} \tag{3}$$

where $R_{it}$ is the rainfall at time $t$ of rain measuring station $i$, $i = 1, 2, 3 \ldots S$, $t = 1, 2, 3 \ldots m$, $S$ is the number of rain-measuring stations, and $m$ is the number of time periods.

Subsequently, a manifold learning algorithm was used to process the rainfall data. Furthermore, the high-dimensional array of rainfall was projected to the low-dimensional space to achieve a dimensionality reduction. Clustering and feature extraction were then performed to obtain the spatial–temporal distribution in the low-dimensional space. Following feature selection and extraction, the information required for our research was extracted.

The manifold learning algorithm applied in this study is based on the consideration that the local linear relationship between the high-dimensional and the low-dimensional spaces remains unchanged. That is to say, the array features in the high-dimensional space remain constant in the low-dimensional space. Accordingly, feature reconstruction in the high-dimensional space was conducted based on the features in the low-dimensional space. This method facilitates the extraction of the spatial–temporal distribution features of rainfall in the high-dimensional space. The specific process is shown in Figure 2.

### *2.2. Manifold Learning Algorithm*

The high-dimensional sample database, describing the spatial–temporal distribution characteristics of rainstorms, is a nonlinear and high-dimensional data space. In this study, the locally linear embedding (LLE) algorithm was used to improve the efficiency of analysis and the accuracy of results [36]. This method is used for dimensionality reductions of nonlinear data and conducting dimensionality reduction analysis for high-dimensional data. The main features of the original data were extracted and expressed using the "effective" feature data with fewer dimensions, without reducing the intrinsic information contained in the original data.

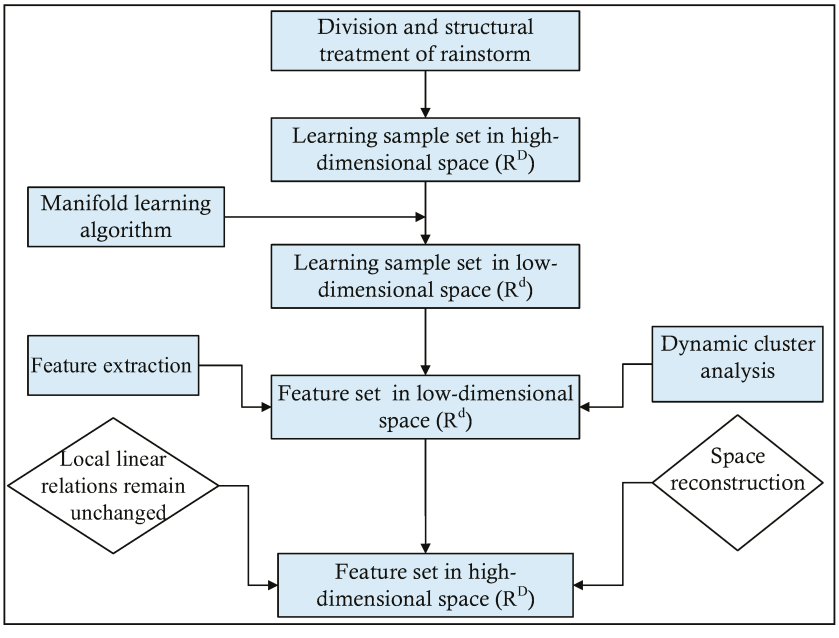



**Figure 2.** Schematic diagram.

The LLE algorithm is an unsupervised dimensionality reduction method for nonlinear data proposed by Roweis et al. [37]. It is a type of manifold learning algorithm that uses partial linearity to reflect the whole nonlinearity. This feature enables the dimensionality reduction data to maintain the original data topology. According to the LLE algorithm, each data point can be constructed using a linear-weighted combination of its neighboring points. The linear relation between $x_i$ and its surrounding samples in the high-dimensional space is the same as the partial linear relation between $y_i$ and its surrounding samples in the low-dimensional space. Therefore, the linear relationship between all samples and the classification results in the low-dimensional space $d$ are similar to those in the high-dimensional space. The algorithm comprises three main parts:

(1) In higher dimensional space, find the $K$ samples closest to sample $x_i$ by Euclidean distance measurement.

First, the Euclidean distance between each sample point $x_i$ and all other samples is calculated for $N$ data points $\{x_1, x_2, \cdots x_N\} \in R^D$. Each $x_i$ can be linearly expressed by $K$ samples $\{x_{i1}, x_{i2}, \cdots x_{ik}\}$ in its neighborhood with the nearest distance, as shown in Equation (4):

$$x_i \approx \overline{x}_i = \sum_{j=1}^{k} w_{ij} x_j \tag{4}$$

meets the condition:

$$\sum_{j=1}^{k} w_{ij} = 1 \tag{5}$$

(2) For each sample $x_i$, the linear relationship of K nearest neighbors in its neighborhood is first determined, and the weight coefficient of the linear relationship is obtained as follows:

The LLE algorithm is used to obtain the weight coefficient $W$ by solving the minimum value of Equation (4) in the constraint condition Equation (5).

$$W_i = \frac{Z_i^{-1} I}{I^T Z_i^{-1} I} \tag{6}$$

Subsequently, weights $W_i = (w_{i1}, w_{i2}, \cdots w_{ik})^T (i = 1, 2, \cdots, N)$ are obtained.

(3) It is assumed that the linear relation weight coefficient $W_i$ remains in high-dimensional and low-dimensional space within the K neighborhood. The weight coefficient $W_i$ is then used to reconstruct sample data in low dimensions and obtain $x_i \in R^D \rightarrow y_i \in R^d$, d $\ll$ D.

The LLE algorithm assumes that samples in high-dimensional space maintain local linear relationships in low-dimensional space, and that the weight coefficient remains unchanged.

Samples in high-dimensional space $\{x_1, x_2, \cdots x_N\} \in R^D$ $W$ are mapped to a low-dimensional space by weight coefficient and become samples in low-dimensional space, $\{y_1, y_2, \cdots y_N\} \in R^d$. High-dimensional sample point $x_i$, mapping $y_i$ in low-dimensional space, can also be obtained by solving the minimum mean-square deviation.

As the detailed solution process of this algorithm was not the focus of this study, it has not been described in detail.

### *2.3. Dynamic Cluster Analysis and Feature Extraction*

After the high-dimensional samples were dimensionally reduced by the LLE algorithm, the dynamic clustering algorithm was used to classify the dimension-reduced samples [38]. Using this algorithm, the population sample set was divided into $r$ subsets, where the samples in each subset were the most similar and the samples between each subset were the most different. The mean value of each subset was then extracted to obtain the features of the subset. In the analysis, $r$ sample points were randomly selected as the initial clustering center of $r$ subsets. The distance between all samples and initial clustering centers r, were calculated, and the samples were divided into the subset of the center nearest to them to obtain the number of initial classification categories and initial subsets. We calculated the mean value of all samples of each subset to obtain the new generation cluster center. We continuously iterated the values according to the above method. When the distance between the clustering centers of generation $p$ and generation $p + 1$ was within the threshold range, the calculation was considered to be convergent, and the final subset and the clustering centers of each subset were obtained.

Each subset C $= \{C_1, C_2, \cdots, C_r\}$ and the mean $Z_j^{p+1}$ of each subset obtained by the above clustering method were the feature spaces of the reduced dimension dataset. However, sample $x_i$ in a higher dimensional space had the same partial linear relationship with its surrounding samples as a mapping point $y_i$ in a lower dimensional space. The samples in the subsets C $= \{C_1, C_2, \cdots, C_r\}$ in the lower dimensional space belonged to the same subset B $= \{B_1, B_2, \cdots, B_r\}$ in the higher dimensional space. The mean value $S_j = \frac{1}{B_j} \sum_{x_{i \epsilon B_j}} x_i \epsilon R^D$ of all subsets in the high-dimensional space was obtained, which also belonged to the dynamic spatial–temporal distribution characteristics of the samples in the high-dimensional space.

## 3. Results and Discussion

### *3.1. Results*

As shown in Figure 3, the average duration time of heavy rain in Beijing was about 24 h, while in Shenzhen it was about 72 h. A heavy rainstorm in Beijing usually falls within one day, while a heavy rainstorm in Shenzhen typically lasts 2–3 days.

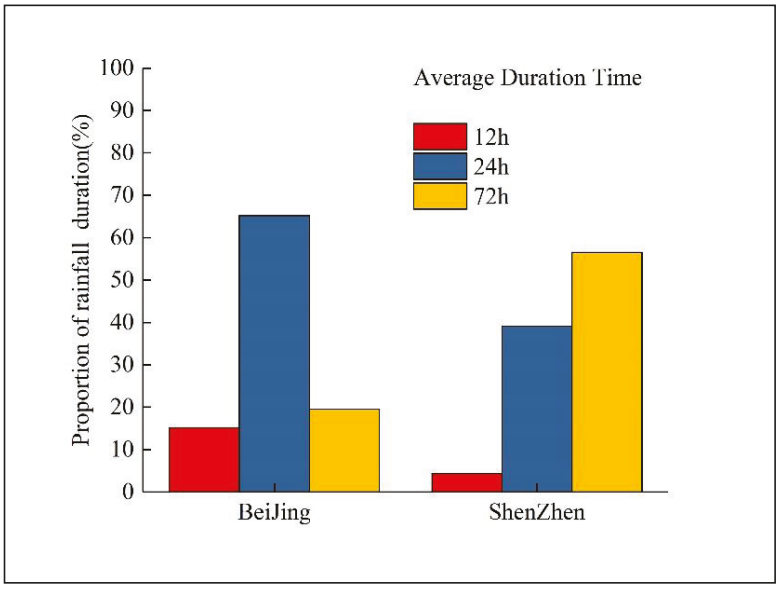



**Figure 3.** Comparison chart of rainfall with different duration.

In this study, we evaluated rainstorm events lasting 24 h in Beijing and 72 h in Shenzhen. The study samples included 32 events in Beijing and 76 in Shenzhen. The spatial–temporal distribution characteristics of the rainstorms in the two cities were analyzed and extracted, and the characteristic rainstorm processes were reconstructed based on the extracted results.

Our analysis revealed major differences between Beijing and Shenzhen in the spatial–temporal distribution of rainstorms. The results of the spatial–temporal distribution characteristics analysis of the 24 h duration rainfall in Beijing are shown in Figure 4. We observed that: (1) The rainstorm started from the mountain areas in the west and north and developed in the plain area. The temporal and spatial variation of the rainstorm was relatively stable, and the mobility of the rainstorm center was not strong during rainfall. (2) The spatial and temporal distribution of the rainstorm presented a line from southwest to northeast, and the trend of the rainfall belt was consistent with that of the underlying mountain plain. (3) The distribution of precipitation was not uniform, with more precipitation in the northeast and southwest, and less precipitation in the northwest and southeast.

Heavy rain occurred in Beijing on 21 July 2012 (Figure 5). During this event, the rainfall started from the mountainous areas in the southwest and moved to the northeast. The distribution of the rainstorm zone was consistent with the trend of the underlying mountain and plain and was similar to the spatial–temporal distribution characteristics of the extracted rainfall.

Compared with Beijing, which is located in the inland area of North China, the spatial–temporal distribution characteristics of rainfall in Shenzhen were diverse. Rainstorms in Shenzhen were divided into three types based on different spatial–temporal distribution characteristics. Figures 6–11 show the spatial–temporal distribution characteristics of each type of rainstorm in Shenzhen, along with the actual rainfall event of each type. As shown in Figure 6, this type of rain starts from the northwest and gradually moves to the southeast, with large rainfall in the north and northwest. Figure 7 shows the actual rainfall process of this type on 22 May 2009. As shown in Figure 7, the rainfall started from the northwest, and the rainfall in the western and northern regions was relatively heavier.

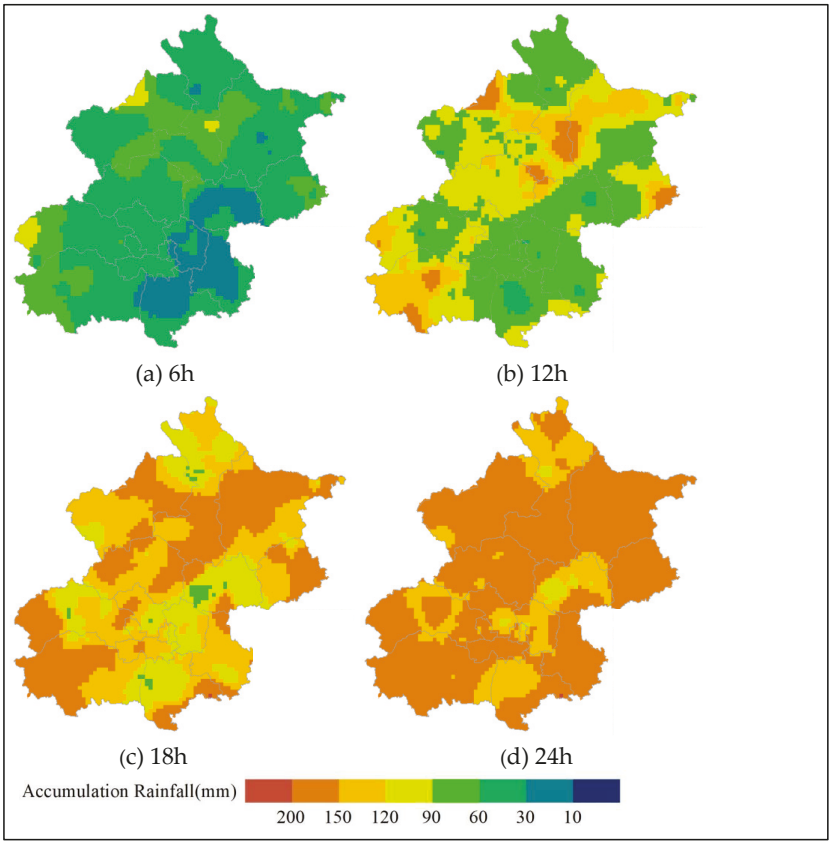



**Figure 4.** Temporal and spatial distribution pattern of 24 h rainfall in Beijing.

The second type of rainstorm, shown in Figure 8, starts from the southern and southeastern coastal areas and gradually moves to the north. The center of the rainstorm does not move significantly, and the rainfall in the southern and southeastern coastal areas is relatively heavier. Figure 9 shows the actual rainfall process of this type on 26 August 2017. As shown in the figure, the rainstorm started from the coastal areas and continued to move northward, with less rainfall in the northwest and more in the coastal areas.

The third type of rainstorm, shown in Figure 10, starts from the southeast coast and gradually moves to the north and northwest, with heavy rainfall in the east and southeast and relatively lighter rainfall in the west. Figure 11 shows the actual rainfall process of this type occurring on 12 June 2017. As shown in the figure, the rainstorm started from southeastern coastal areas and continued to move to the middle and west, with relatively lighter rainfall in the northwest and relatively heavier rainfall in the eastern and southeastern coastal areas.

These three types of rainfall, each with different spatial–temporal distribution patterns, can be explained by physical mechanisms. The first type of rainfall usually occurs from April through May. During this period, the cold air from the south is strong, while the warm and humid air masses in the southwest provide sufficient water vapor. This type of rainfall usually forms at the intersection of warm and cold air masses, and thus moves from northwest to southeast [39], and the rainfall in the northern and northwestern regions is relatively heavier. The second type of rainfall is usually caused by the southwest monsoon, which occurs from June through August. During this period, warm and humid air masses

from the sea strengthen, and a large amount of water vapor is continuously brought to the shore by the southwest monsoon [40]. Therefore, this type of rainfall usually moves from the southern and southeastern coastal areas to the northern areas, with more precipitation seen in the coastal areas. The third type of rainfall is generally caused by eastern waves or typhoons, usually from the northwest Pacific Ocean [41]. This type of rain makes landfall in coastal areas southeast of Shenzhen, and then continues to move to western areas.

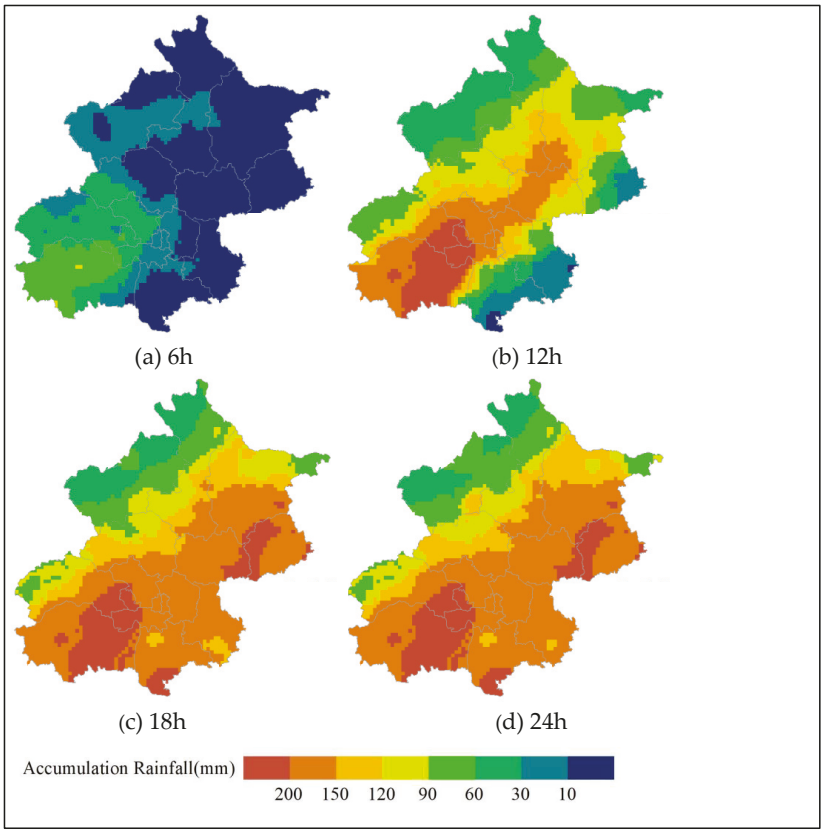



**Figure 5.** Rainfall accumulation of rainstorm occurred on 21 July 2012 in Beijing.

### *3.2. Discussions*

Through analysis, it is found that the rainfall in Beijing in summer has a strong correlation with underlying surface topography, with obvious characteristics of rain in front of mountains, but that rainfall movement is obvious. In Shenzhen, weather conditions that cause rainfall are complex, and the temporal and spatial distribution characteristics of rainfall are also different. The correlations between the temporal and spatial distribution characteristics of rainfall and the underlying terrain conditions are weakly established. The temporal and spatial distribution characteristics of rainfall are closely related to the weather conditions and are more relevant to the source of water vapor.

In this study, machine learning algorithms are used to analyze the pattern of rainfall. Although machine learning technology cannot directly study the physical mechanisms behind rainstorms, it enables the extraction of the spatial–temporal distribution characteristics of regional rainfall through the analysis and mining of historical rainfall data. The temporal and spatial distribution pattern of characteristics has similar temporal and

spatial distribution patterns to typical actual rainstorm events. This comparison shows that temporal and spatial distribution characteristics of rainfall processes are adequately representative of actual rainfall events. These features and laws are consistent with the weather conditions that cause rainfall. The extraction results can be reasonably explained by local climatic characteristics, topographic conditions, and other factors.

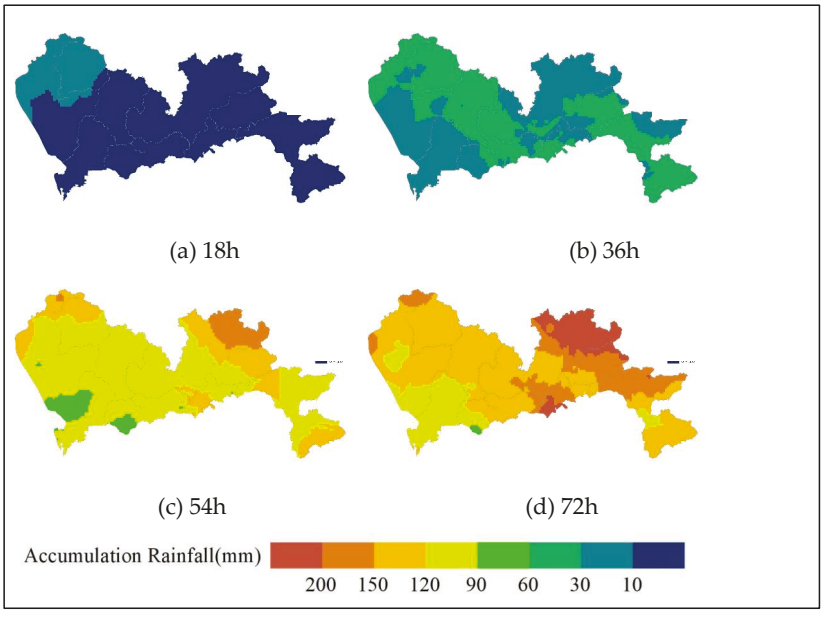



**Figure 6.** Temporal and spatial distribution pattern of the type I characteristics of 72 h rainfall in Shenzhen.

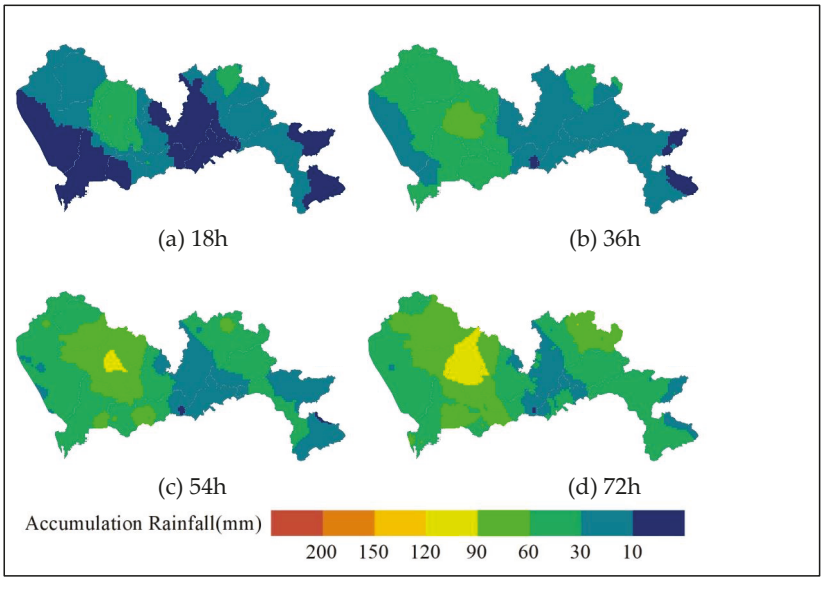



**Figure 7.** Rainfall accumulation of rainstorm of type I occurred on 22 May 2009.

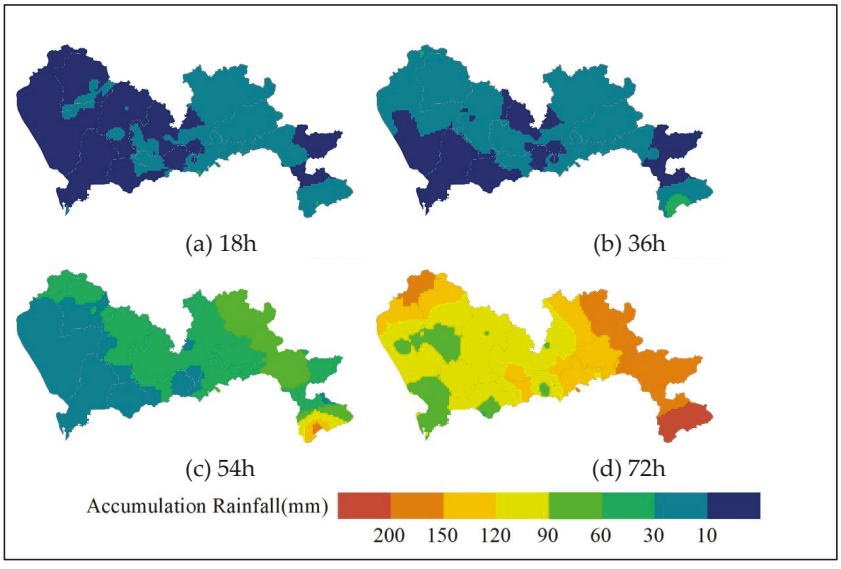



**Figure 8.** Temporal and spatial distribution pattern of the type II characteristics of 72 h rainfall in Shenzhen.

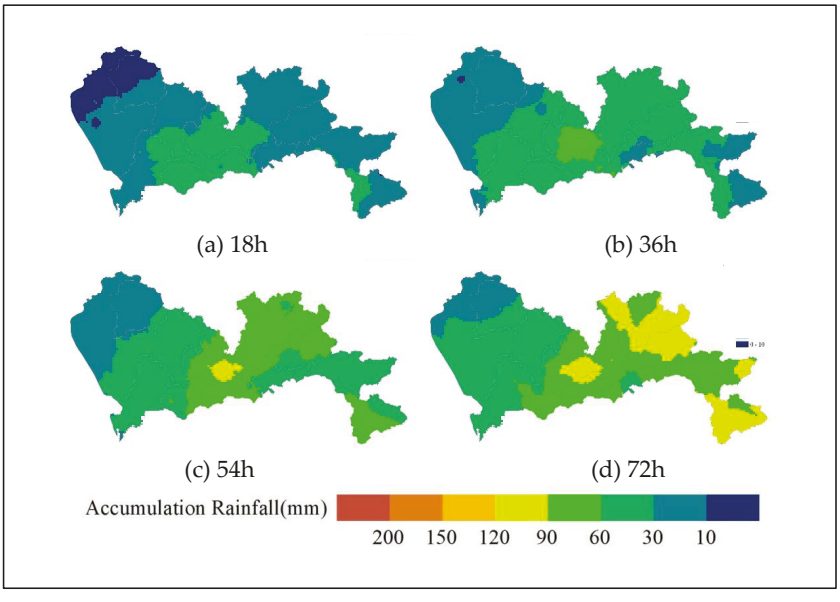



**Figure 9.** Rainfall accumulation of rainstorm of type II occurred on 26 August 2017.

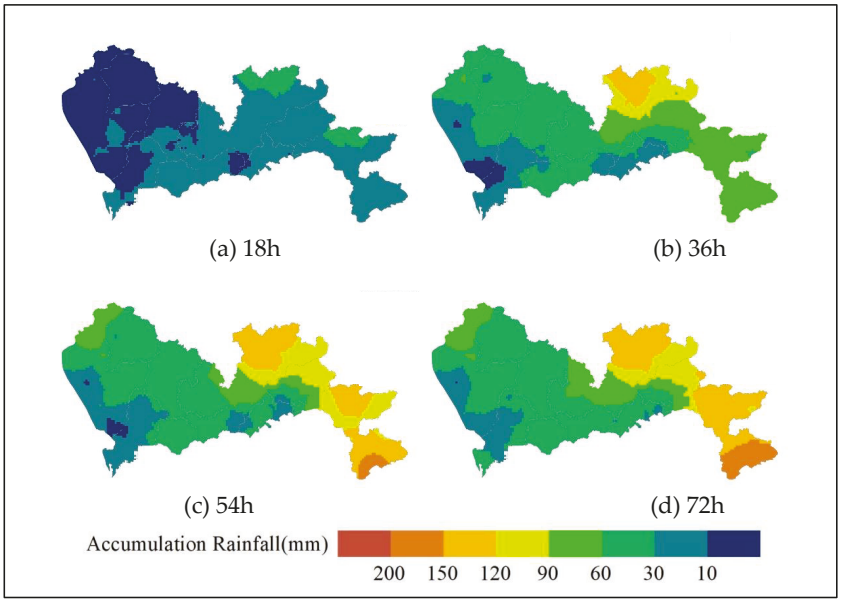



**Figure 10.** Temporal and spatial distribution pattern of the type III characteristics of 72 h rainfall in Shenzhen.

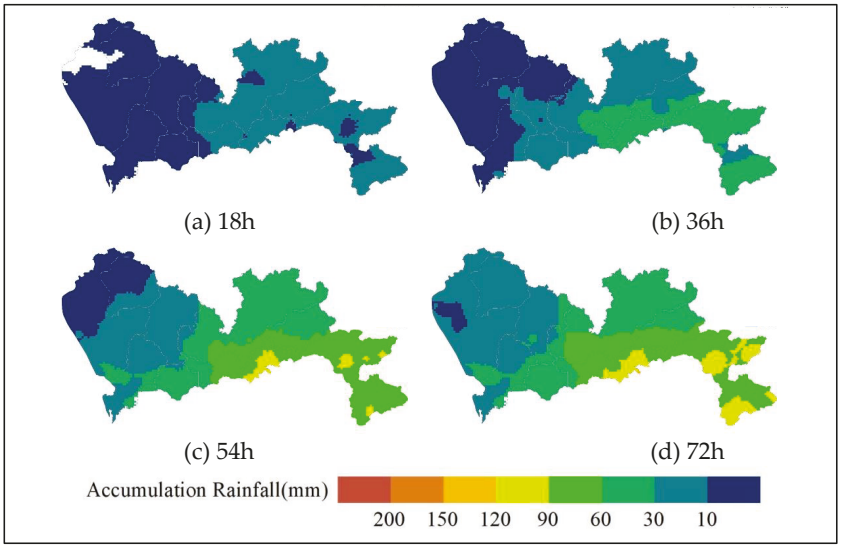



**Figure 11.** Rainfall accumulation of rainstorm of type III occurred on 12 June 2017.

## 4. Conclusions

In this study, we analyzed the 5 min rainfall data from the last 20 years, collected by stations in Beijing and Shenzhen. We described the spatial–temporal characteristics of rainfall patterns in a high-dimensional array. We applied manifold learning algorithms to analyze and extract the spatial–temporal distribution characteristics of rainfall in Beijing and Shenzhen. The results extracted by the machine learning algorithm in this study identified physical mechanisms consistent with the climatic characteristics and topographic conditions of the region.

Our research showed that, although machine learning alone cannot fully explain the physical mechanisms of rainfall, data analysis utilizing machine learning algorithms can identify rainfall patterns and quantitative spatial–temporal characteristics. The method was put forward to analyze spatio-temporal distribution characteristics of rainfall, which can provide a basis for the design of rainfall patterns in different regions.

In the future, with the increase in data timing and the expansion of data range, the algorithm will become more objective and produce reasonable results. In this study, only the ground rainfall observation data were analyzed. The performance of the method can be improved by combining it with the use meteorological data, such as radar echo maps and meteorological cloud maps. The proposed method can be further improved to improve the accuracy of methods to identify early risk of rainstorms.

**Author Contributions:** Y.L. (Yesen Liu), H.R., L.Z., C.W. and Q.G. collected and processed the data; Y.L. (Yuanyuan Liu), Y.L. (Yesen Liu), H.R., S.L. and L.D. proposed the model and analyzed the results; and Y.L. (Yuanyuan Liu), Y.L. (Yesen Liu) and H.R. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by the National Key R&D Program of China (2022YFC3090600), and the Chinese National Natural Science Foundation (No. 51739011).

**Institutional Review Board Statement:** Not applicable for studies not involving humans or animals.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Participants of this study did not agree for their data to be shared publicly, so supporting data are not available.

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

## References

1. Zhang, J.Y.; Wang, Y.T.; He, R.M.; Hu, Q.F.; Song, X.M. Discussion on the urban flood and waterlog ging and causes analysis in China. *Adv. Water Sci.* **2016**, *27*, 485–491. [CrossRef]
2. Chen, X.T. Urban Water Disasters and Strategy of Comprehensive Control of Water Disaster. *J. Catastrophology* **2010**, *25*, 10–15.
3. Zhang, J.Y.; Wang, Y.T.; Hu, Q.F.; He, R.M. Discussion and views on some issues of the sponage city con struction in China. *Adv. Water Sci.* **2016**, *27*, 793–799. [CrossRef]
4. Rafieeinasab, A.; Norouzi, A.; Kim, S.; Habibi, H.; Nazari, B.; Seo, D.-J.; Lee, H.; Cosgrove, B.; Cui, Z. Toward High-Resolution Flash Flood Prediction in Large Urban Areas–Analysis of Sensitivity to Spatiotemporal Resolution of Rainfall Input and Hydrologic Modeling. *J. Hydrol.* **2015**, *531*, 370–388. [CrossRef]
5. Onyutha, C.; Willems, P. Influence of Spatial and Temporal Scales on Statistical Analyses of Rainfall Variability in the River Nile Basin. *Dyn. Atmos. Ocean.* **2017**, *77*, 26–42. [CrossRef]
6. Liu, Y.S.; Liu, Y.Y.; Li, M. Influence of rainfall spatial and temporal heterogeneity on flooding process in urban rivers. *S. N. Water Transf. Water Sci. Technol.* **2022**, *20*, 317–326. [CrossRef]
7. Wang, G.; Sun, G.; Li, J.; Li, J. The Experimental Study of Hydrodynamic Characteristics of the Overland Flow on a Slope with Three-Dimensional Geomat. *J. Hydrodyn.* **2018**, *30*, 153–159. [CrossRef]
8. Li, Z.J.; Jiang, T.T.; Huang, P.N.; Liu, Z.Y.; An, D.; Yao, C.; Ju, X.S. Impact and analysis of watershed precipitation and topography characteristics on model simulation results. *Adv. Water Sci.* **2015**, *26*, 473–480. [CrossRef]
9. Guo, L.; Ding, L.Q.; Sun, D.Y.; Liu, C.J.; He, B.S.; Liu, R.H. Key techniques of flash flood disaster prevention in China. *J. Hydraul. Eng.* **2018**, *49*, 1123–1136. [CrossRef]
10. Emmanuel, I.; Payrastre, O.; Andrieu, H.; Zuber, F. A Method for Assessing the Influence of Rainfall Spatial Variability on Hy drograph Modeling. First Case Study in the Cevennes Region, Southern France. *J. Hydrol.* **2017**, *555*, 314–322. [CrossRef]

11. Fung, K.F.; Chew, K.S.; Huang, Y.F.; Ahmed, A.N.; Teo, F.Y.; Ng, J.L.; Elshafie, A. Evaluation of spatial interpolation methods and spatiotemporal modeling of rainfall distribution in Peninsular Malaysia. *Ain Shams Eng. J.* **2021**, *13*, 101571. [CrossRef]
12. Hitchens, N.M.; Brooks, H.E.; Schumacher, R.S. Spatial and temporal characteristics of heavy hourly rainfall in the United States. *Mon. Weather. Rev.* **2013**, *141*, 4564–4575. [CrossRef]
13. Satya, O.C.; Arsali, M.; Affandi, A.K.; Mandailing, P.M. Spatial distribution and diurnal characteristics of rainfall in South Sumatra and surrounding areas based on Tropical Rainfall Measuring Mission (TRMM) data. *J. Phys. Conf. Ser.* **2020**, *1568*, 012024. [CrossRef]
14. Ndiaye, B.; Moussa, M.A.; Wade, M.; Sy, A.; Diop, A.B.; Diop, A.D. Spatial and Temporal Distribution of Rainfall Breaks in Senegal. *Am. J. Clim. Change* **2021**, *10*, 533–560. [CrossRef]
15. Chaubey, P.K.; Mall, R.K.; Jaiswal, R.; Payra, S. Spatio-Temporal Changes in Extreme Rainfall Events Over Different Indian River Basins. *Earth Space Sci.* **2022**, *9*, e2021EA001930. [CrossRef]
16. Audu, M.O.; Ejembi, E.; Igbawua, T. Assessment of Spatial Distribution and Temporal Trends of Precipitation and Its Extremes over Nigeria. *Am. J. Clim. Change* **2021**, *10*, 22. [CrossRef]
17. Yeung, J.K. Summertime Convective Rainfall in the New York City-New Jersey Metropolitan Region. Ph.D. Thesis, Princeton University, Princeton, NJ, USA, 2012.
18. Viglione, A.; Chirico, G.B.; Woods, R.; Blöschl, G. Generalised synthesis of space-time variability in flood response: An analytical framework. *J. Hydrol.* **2010**, *394*, 198–212. [CrossRef]
19. Zoccatelli, D.; Borga, M.; Viglione, A.; Chirico, G.B.; Blöschl, G. Spatial moments of catchment rainfall: Rainfall spatial organisation, basin morphology and flood response. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 3767–3783. [CrossRef]
20. Wu, D.; Zhao, K.; Yu, H.; Wang, M.J. An analysis of spatial and temporal variations in the axisymmetric precipitation structure associ ated with typhoons making landfall on the southeastern coast of China based on the Doppler radar data. *Acta Meteorol. Sin.* **2010**, *68*, 896–907.
21. Ngongondo, C.; Xu, C.Y.; Gottschalk, L.; Alemaw, B. Evaluation of Spatial and Temporal Characteristics of Rainfall in Malawi: A Case of Data Scarce Region. *Theor. Appl. Climatol.* **2011**, *106*, 79–93. [CrossRef]
22. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
23. Liu, Y.Y.; Liu, H.W.; Huo, F.L. An application of machine learning on examining spatial and temporal distribution of short duration rainstorm. *J. Hydraul. Eng.* **2019**, *50*, 773–779. [CrossRef]
24. Liu, Y.Y.; Li, L.; Zhang, W.H.; Chan, P.W.; Liu, Y.S. Rapid Identification of Rainstorm Disaster Risks Based on an Artificial Intelligence Technology Using the 2DPCA Method. *Atmos. Res.* **2019**, *227*, 157–164. [CrossRef]
25. Sharma, S.K. A novel approach on water resource management with Multi-Criteria Optimization and Intelligent Water De mand Forecasting in Saudi Arabia. *Environ. Res.* **2022**, *208*, 112578. [CrossRef] [PubMed]
26. Salloom, T.; Kaynak, O.; Yu, X.B.; He, W. Proportional integral derivative booster for neural networks-based time-series prediction. *Eng. Appl. Artif. Intell.* **2022**, *108*, 104570. [CrossRef]
27. Sadeghifar, T.; Barati, R. Application of adaptive Neuro-fuzzy inference system to estimate alongshore sediment transport rate (A real case study: Southern shorelines of Caspian Sea). *J. Soft Comput. Civ. Eng.* **2018**, *2*, 72–85.
28. Alizadeh, M.J.; Shahheydari, H.; Kavianpour, M.R.; Shamloo, H.; Barati, R. Prediction of longitudinal dispersion coefficient in natural rivers using a cluster-based Bayesian network. *Environ. Earth Sci.* **2017**, *76*, 86. [CrossRef]
29. Hu, R.; Fang, F.; Pain, C.C.; Navon, I.M. Rapid spatio-temporal flood prediction and uncertainty quantification using a deep learning method. *J. Hydrol.* **2019**, *575*, 911–920. [CrossRef]
30. Liu, Y.Y.; Liu, Y.S.; Zheng, J.W. Intelligent rapid prediction method of urban flooding based on BP neural network and numerical simulation model. *J. Hydraul. Eng.* **2022**, *53*, 284–295.
31. Bao, Z.X.; Zhang, J.Y.; Wang, G.Q.; He, R.M.; Jin, J.L.; Wang, J.; Wu, H.F. Quantitative assessment of the attribution of runoff and sediment changes based on hydrologic model and machine learning: A case study of the Kuye River in the Middle Yellow River basin. *Adv. Water Sci.* **2021**, *32*, 485–496. [CrossRef]
32. Zheng, T.F.; Liu, X.T.; Wang, Q.L.; Yu, X. The characteristics of precipitation in Guangdong Province over the past 50 years: Dependence on rainfall intensity. *J. Trop. Meteorol.* **2017**, *33*, 212–220. [CrossRef]
33. Luo, C.; Liu, Y.S.; Liu, Y.Y.; Mu, J.; Liu, X.P.; Feng, M.M. A smart-framework for urban flood and waterlogging analysis. *China Flood Drought Manag.* **2021**, *31*, 17–22. [CrossRef]
34. Cauteruccio, A.; Colli, M.; Stagnaro, M.; Lanza, L.G.; Vuerich, E. In situ precipitation measurements. In *Springer Handbook of Atmospheric Measurements*; Springer Handbooks; Springer: Cham, Switzerland, 2021; pp. 259–400. ISBN 978-3-030-52170-7. [CrossRef]
35. Cauteruccio, A.; Brambilla, E.; Stagnaro, M.; Lanza, L.G.; Rocchi, D. Experimental evidence of the wind-induced bias of precipitation gauges using Particle Image Velocimetry and particle tracking in the wind tunnel. *J. Hydrol.* **2021**, *600*, 126690. [CrossRef]
36. Belkin, M.; Niyogi, P. Laplacian Eigenmaps for Dimensionality Reduction and Data Representation. *Neural Comput.* **2003**, *15*, 1373–1396. [CrossRef]
37. Roweis, S.T.; Saul, L.K. Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science* **2000**, *290*, 2323–2326. [CrossRef]
38. Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A K-Means Clustering Algorithm. *Appl. Stat.* **1979**, *28*, 100. [CrossRef]

39. Chen, Z.H.; He, G.B.; Cui, C.G. The coupling reaction among convective cloud cluster and moisture frontal zone and low level jet—A possible mechanism of maintenance of continuing heavy rain. *J. Trop. Meteorol.* **2007**, *23*, 246–254.
40. Hu, Y.; Du, Y.; Luo, X. Precipitation Patterns During the "Dragon Boat Water" in South China for the Recent 49 Years. *Meteor. Mon.* **2013**, *39*, 1031–1041.
41. Huang, R.H.; Chen, G.H. Besearch on interannual variations of tracks of tropical cyclones over northwest pacific and their physical mechanism. *Acta Meteorol. Sin.* **2007**, *65*, 683–694.

water

MDPI

*Article*

# Spatiotemporal Modes of Short Time Rainstorms Based on High-Dimensional Data: A Case Study of the Urban Area of Beijing, China

**Wei Liu [1], Sheng Chen [2] and Fuchang Tian [1,*]**

[1] School of Civil Engineering, Tianjin University, Tianjin 300072, China; waterliuwei@tju.edu.cn
[2] China Institute of Water Resources and Hydropower Research, Beijing 100038, China; chensheng@iwhr.com
* Correspondence: tianfuchang@tju.edu.cn; Tel.: +86-22-27401156

**Abstract:** The identification of the characteristics of short time rainstorms in urban areas is a difficult problem. The traditional rainfall definition methods, using rainfall graph or a GIS map, respectively reflect the temporal or spatial variations of a rainfall process, but do not regard a rainfall as one complete process including its temporal and spatial dimension. In this paper, we present an approach to define typical modes of rainfall from the temporal and spatial dimensions. Firstly, independent rainfall processes are divided based on the continuous monitoring data of multiple rainfall stations. Subsequently, algorithms are applied to identify the typical spatiotemporal modes of rainfall and reconstruction of the process of modes, including dimensionality reduction, clustering, and reconstruction. This approach is used to analyze the monitoring data (5 min intervals) from 2004 to 2016 of 14 rainfall stations in Beijing. The results show that there are three modes of rainstorms in the Beijing urban area, which account for 31.8%, 13.7%, and 54.6% of the total processes. Rainstorm of mode 1 moves from the northwest to the center of Beijing, then spreads to the eastern part of the urban area; rainstorm of mode 2 occurs in the southwestern region of the urban area, and gradually northward, but there is no rainfall in the mountainous northwest; rainstorm of mode 3 is concentrated in the central, eastern, and southern regions. The approach and results of this study can be applied to rainstorm forecasting or flood prevention.

**Keywords:** rainstorm mode; high dimension; dimension reduction; cluster

**Citation:** Liu, W.; Chen, S.; Tian, F. Spatiotemporal Modes of Short Time Rainstorms Based on High-Dimensional Data: A Case Study of the Urban Area of Beijing, China. *Water* **2021**, *13*, 3597. https://doi.org/10.3390/w13243597

Academic Editors: Yaohuan Huang, Yesen Liu, Runhe Shi and Hongyan Ren

Received: 9 November 2021
Accepted: 5 December 2021
Published: 14 December 2021



## 1. Introduction

The security of water resources in this changing environment has become a research focus, due to the fact that climate change causes variations in rainfall at a large scale, and human activities influence the spatiotemporal characteristics at the regional scale [1,2]. Big cities are especially concerned as they are the hub of human activities. Human activities lead to the frequent occurrence of extreme rainstorms, through the urban heat island effect and air pollution [3,4]. Statistics show that 60% of the cities in China suffered from waterlogging from 2014 to 2016 [2]. Studies showed that urban waterlogging is directly related to the temporal and spatial distribution of rainstorms [5]. For example, the flood peak of triangle rainfall with a rain-peak in the central or rear is 30% larger than that of even rainfall, irrespective of whether the average rainfall is the same [6]. It is of great significance to study the spatial and temporal modes of rainfall to prevent waterlogging [7].

At present, there are two main approaches to the definition of urban rainfall process. First, different rainfall types are applied to describe various rainfall processes, such as single peak and double peaks, the method based on site monitoring data. Studies of this method focus on monitoring data of a single station or the average of multiple stations. Pilgrim and Cordery [8] put the time of a rain peak at the most likely position, and the proportion of the rain peak in the total rainfall is the average of the proportion of the rainfall peaks in each field. Keifer and Chu [9] designed the Chicago-mode according

to the strength, diachronic, and frequency of rainfall. Huff [10] designed four modes of rainfall, according to the location of the rain peak and duration of rainfall in Illinois, USA. These methods only consider the total rainfall and extreme rainfall of a single station [11]. However, the data of a single station cannot reflect the spatial characteristics of rainfall, especially in metropolitan areas with obvious spatiotemporal variations of the background environment, such as temperature and wind direction [12]. This method is widely used in urban planning or urban construction, but it is increasingly criticized for neglecting the spatial variations of rainfall, especially in large cities. The second method defines rainfall from the spatial perspective based on geographic theory. The general method is to use a spatial interpolation algorithm to interpolate the monitoring data from stations into spatial distribution data, such as the software ANUSPLIN, developed by Hutchinson, of the Australian National University [13]. In recent years, with the advancement of satellite and radar technology, rainfall spatial data can be obtained more directly, with higher accuracy, for example with Global Precipitation Climatology Project (GPCP) and Tropical Rainfall Measuring Mission (TRMM). These spatial distribution data are more suitable for analyzing the distribution of total rainfall or cumulative rainfall in a certain period, but it is difficult to express the correlation between rainfall distributions in different periods. However, in the real rainfall process, the rain belt usually moves rapidly. The spatial distribution of total rainfall or the time history distribution of rainfall intensity at a single point are not enough to accurately describe the dynamic temporal and spatial distribution characteristics of a rainfall, which is very important for the risk emergency management of rainstorm.

For the purpose of reducing the urban storm disaster effectively, it is necessary to express a complete "rainfall process", which includes not only the rainfall graph of stations at different locations, but also the spatial relationship of rainfall of these stations at different period. So, rainfall is a multidimensional data that includes time and spatial characteristics. Some scholars try to integrate the temporal and spatial characteristics of rainfall with multidimensional data. For example, based on the 3-dimensional mosaic reflectivity data from 10 S-band Doppler radars in Guangdong province, an artificial intelligence (AI) algorithm for automatic hail detection and nowcasting is developed in the light of the machine learning (ML) technology [14]. The temporal and spatial distribution characteristics of short duration rainfall in Shenzhen are analyzed and extracted by LLE algorithm [15]. It is necessary to conduct further study on the rainfall process with temporal and spatial dimensions.

In contrast to taking the rainfall data of a single station as the research object, this paper used the ML algorithm to extract the temporal and spatial distribution characteristics of rainfall from the rainfall data of all rainfall stations in the whole research area. The continuous monitoring data of rainfall stations from 2004 to 2016 were divided into different rainfall processes in Beijing. The rainfall processes were taken as the research objects, then, algorithms, such as dimensionality reduction, clustering, and reconstruction were applied to identify the typical spatiotemporal process of rainfall, and then simulate the rainfall process of different modes.

## 2. Study Area and Data Process

According to previous studies, Beijing has become one of the most urbanized cities in China in the past 30 years [16,17]. In 2013, urban population accounted for 86.3% of the total population in Beijing, far higher than the average of 53.73% for China, causing a rapid expansion of built-up areas [17]. In recent years, severe waterlogging has frequently occurred in Beijing [18]. For example, floods and waterlogging caused serious casualties on 21 July 2012 [19].

The Beijing urban area is 396 $km^2$. Fourteen meteorological monitoring stations have continuously recorded the rainfall data of the Beijing urban area in recent years. The continuous monitoring data of 14 rainfall stations were selected from the database, which contains data from 2004 to 2016, at intervals of 5 min. The 14 rainfall stations were evenly distributed in urban areas, as shown in Figure 1.

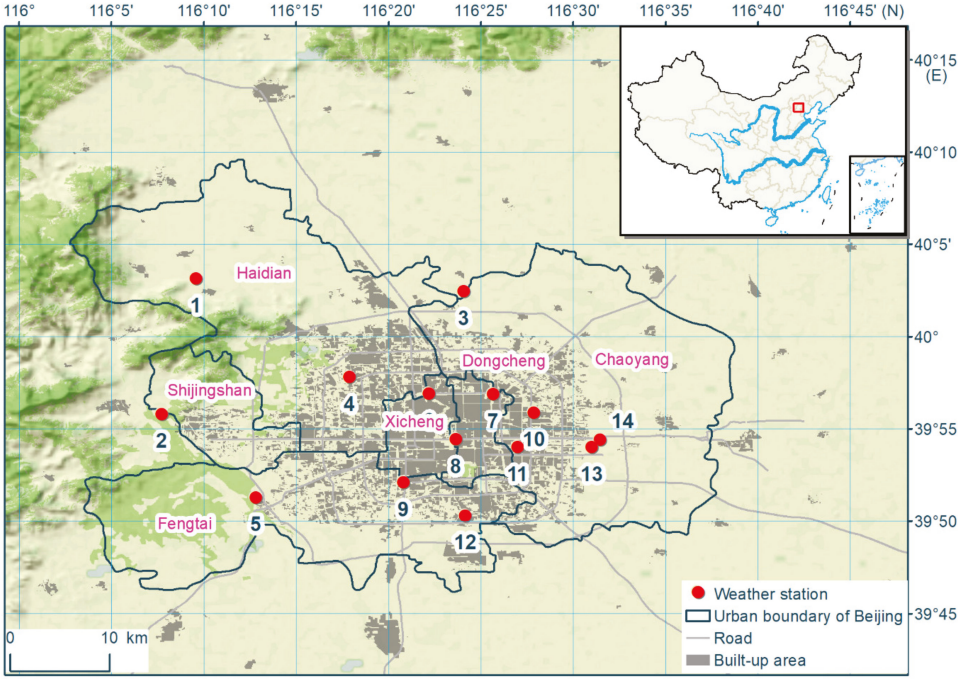



**Figure 1.** Study area.

Statistics show that 89 rainstorms occurred during 2004 and 2016, which is defined as rainfall of 1 hour exceeding 30 mm [20]. For the comparability of rainfall processes, it was necessary to standardize the rainfalls with different duration. According to previous studies, the method of 1-h moving average was used to deal with the rainfall processes, as shown in Figure 2. The data in the red box were selected as the standardized rainfall.

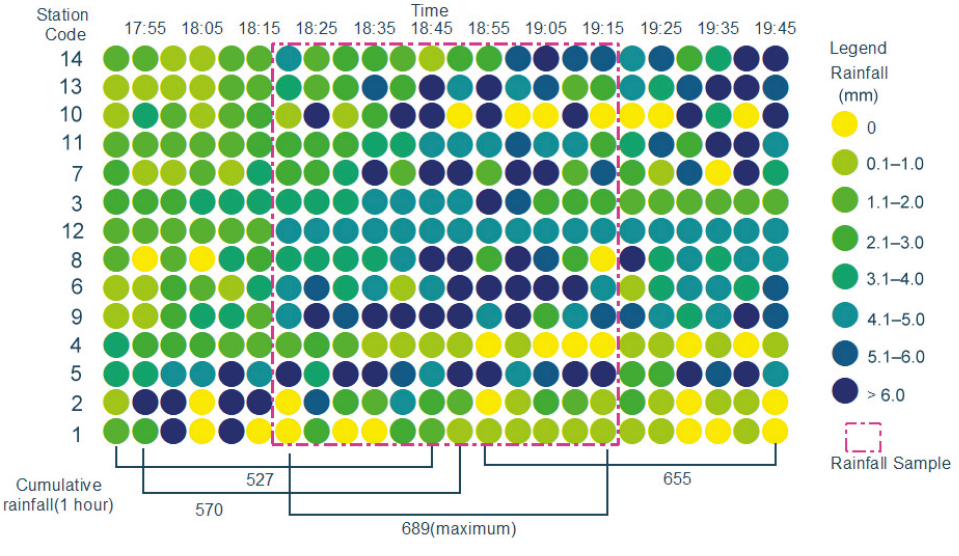



**Figure 2.** Rainfall of 1-h moving average.

Figure 2 shows the data of a rainstorm process at 14 stations. In Figure 2, "cumulative rainfall" represents the sum of rainfall at 14 stations at a certain period. For example, "527" represents the sum of rainfall at all 14 stations from the beginning of rainfall to 1 h. The range surrounded by red dotted/dashed outlines indicates that the "cumulative rainfall" was the largest in this hour compared with any other hour.

According to the early warning standard issued by Beijing Flood Control Office, when the rainfall exceeds 50 mm in one-hour, yellow warning signals will be issued to the public, indicating that serious urban disasters may occur. In 2004–2016, 22 rainfall events had a maximum sliding rainfall of more than 50 mm in one hour, accounting for 24.7% of the total rainfalls. In this paper, the 22 rainstorms were selected as samples, which are called "short duration storm". The rainfall data of these 22 rainstorms are shown in Table 1.

**Table 1.** Data for 22 rainstorms.

| | Instantaneous Rainfall (mm) Every 5 min | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Date** | **5** | **10** | **15** | **20** | **25** | **30** | **35** | **40** | **45** | **50** | **55** | **60** | **Total** |
| 2006-6-27 23:00 | 4.2 | 7.4 | 22.2 | 8.2 | 8.6 | 8.6 | 10.3 | 10.8 | 12.1 | 11.1 | 7.7 | 5 | 116.2 |
| 2006-6-30 22:45 | 3.8 | 3.3 | 3.8 | 4 | 3.4 | 3 | 7 | 19.7 | 18.9 | 18.4 | 5.4 | 8.9 | 99.6 |
| 2006-7-12 4:25 | 3 | 3 | 3 | 3 | 3.5 | 4.1 | 13.6 | 19.3 | 14.2 | 22.2 | 20.6 | 13.1 | 122.6 |
| 2006-7-12 19:05 | 4.5 | 4.5 | 7.3 | 20.3 | 24.2 | 16.3 | 11 | 13.2 | 14.8 | 11.6 | 9.6 | 8.6 | 145.9 |
| 2006-7-13 22:50 | 3.6 | 3.3 | 3.3 | 4.2 | 9.9 | 11.5 | 16 | 11.4 | 7.3 | 7.8 | 6.3 | 4.3 | 88.9 |
| 2006-7-31 8:50 | 10.6 | 7.1 | 3.3 | 9.7 | 12.7 | 16.7 | 20.7 | 28.7 | 18.2 | 17.2 | 15.7 | 11.7 | 172.3 |
| 2007-6-27 13:05 | 8.4 | 12.4 | 16.4 | 15.9 | 23.9 | 28.9 | 24.9 | 20.4 | 18.4 | 13.9 | 11.1 | 9.6 | 204.2 |
| 2008-6-23 14:55 | 4.8 | 3.5 | 5.8 | 4 | 9.2 | 11.7 | 16.3 | 9.2 | 10.3 | 7.7 | 10.7 | 8.5 | 101.7 |
| 2008-7-18 9:10 | 8.4 | 6.8 | 8.8 | 7.8 | 4.8 | 6.8 | 5.8 | 2.8 | 4.8 | 3.8 | 3.8 | 7.1 | 71.5 |
| 2009-7-22 17:15 | 6.5 | 11.5 | 16 | 16.2 | 8.5 | 20.4 | 24.9 | 19.9 | 17.9 | 12.9 | 5 | 6.5 | 166.2 |
| 2009-7-23 15:50 | 2 | 1 | 13.5 | 18.7 | 16.5 | 28 | 23 | 19.4 | 16.6 | 4.6 | 3.6 | 4.6 | 151.5 |
| 2009-8-7 16:20 | 2.1 | 3.4 | 2.7 | 6.6 | 9.3 | 8.6 | 9 | 9.6 | 6 | 3.5 | 5 | 2.2 | 68 |
| 2011-7-1 7:05 | 4 | 5.3 | 4.2 | 8.6 | 4.8 | 12.2 | 16.5 | 19.1 | 19.9 | 19.1 | 19 | 14.5 | 147.2 |
| 2013-7-31 19:20 | 5.1 | 0.6 | 5.1 | 1.1 | 13.6 | 0.1 | 5.6 | 0.6 | 3.6 | 1.1 | 11.1 | 7.8 | 55.4 |
| 2014-6-10 14:05 | 4.9 | 6.9 | 4.6 | 7.5 | 16.5 | 26.9 | 52.5 | 18.8 | 39.3 | 25.2 | 16.5 | 21.2 | 240.8 |
| 2014-6-15 18:55 | 8.5 | 2.3 | 6.2 | 9.6 | 14.2 | 13.1 | 8.7 | 8.4 | 8.6 | 9.1 | 11.2 | 7.3 | 107.2 |
| 2014-7-16 18:55 | 16.3 | 8.8 | 22 | 6.3 | 5.8 | 8.5 | 7.3 | 15.1 | 15.6 | 25 | 26.9 | 20 | 177.6 |
| 2014-8-23 22:30 | 0.3 | 1.8 | 7.9 | 14.3 | 16.8 | 12.7 | 18.3 | 22.8 | 12.3 | 7.8 | 8.4 | 4.8 | 128.2 |
| 2014-8-30 21:50 | 4.9 | 9.4 | 7.9 | 14.6 | 7.7 | 13.4 | 8.4 | 10.9 | 12.8 | 7.7 | 11.2 | 13.9 | 122.8 |
| 2015-8-23 14:10 | 2.5 | 4.5 | 5.7 | 9.5 | 15.6 | 7.4 | 3.7 | 2 | 5.3 | 9.9 | 4.2 | 5.1 | 75.4 |
| 2016-8-6 22:05 | 7.5 | 7 | 11 | 7 | 5.5 | 5.1 | 6 | 7.1 | 5.5 | 5.5 | 6.2 | 17.3 | 90.7 |
| 2016-9-7 18:25 | 5 | 3.7 | 1.3 | 21.3 | 26 | 32.5 | 38.1 | 22.8 | 6.8 | 18.3 | 9.5 | 5.3 | 190.6 |

Data source: Beijing Municipal Land and Water Protection Station.

## 3. Methodology

### *3.1. Flowchart for Extraction of Rainfall Temporal and Spatial Modes*

Rainstorm is described according to the duration, intensity, total amount, and frequency by most researchers. In this paper, a new method is introduced to describe a rainstorm. The spatiotemporal process of a rainstorm was constructed as a high-dimensional array, and then principal component analysis (PCA), dynamic clustering (k-means), and reconstruction were applied to analyze the array [21]. The flowchart is shown in Figure 3.

In Figure 3, "m" is the number of samples, which represents the number of rainfalls in the manuscript. "n" is the dimension of the samples which represents the dimension of the rainfall matrix. "k" represents the dimension of rainfall samples after dimensionality reduction.

(1) The rainstorm events were digitized and structured. High-dimensional arrays were established from temporal and spatial dimension perspectives.
(2) Principal component analysis was used to map high-dimensional array to low latitude array.
(3) Dynamic clustering was used to categorize samples to typical modes for describing the temporal and spatial distribution of rainstorms.

(4) With the inverse calculation of principal component analysis, the low dimensional array was reduced to a high-dimensional array to express the spatiotemporal process of each rainstorm modes.

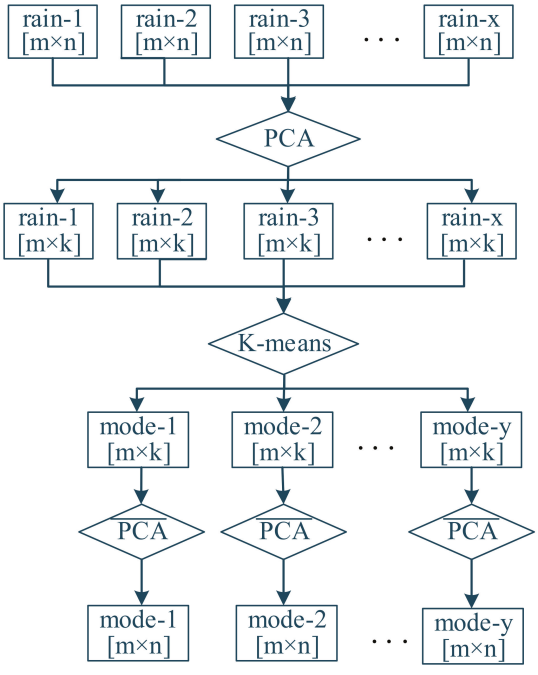



**Figure 3.** Flowchart for extraction of rainfall temporal and spatial modes.

### *3.2. Construction of High-Dimensional Array for Rainstorms*

Rainstorm was characterized as a spatiotemporal process. The continuous monitoring data was divided into independent rainfall periods, where the duration of no rainfall was longer than 120 min [20]. When the maximum rainfall in 1 h was greater than 30 mm, it was called a rainstorm.

We take the monitoring data from multiple stations in a rainstorm as a multi-dimensional array (shown in Figure 2). Based on the processing method introduced by "2. Study Area and Data process" in this paper, the multi-dimensional array was converted to matrix with same rows and columns, the number of which is 14 × 12 in this paper. Fourteen is the number of stations and twelve is the number of periods. The matrix forms one sample in the database of rainstorms (Ω). This is shown in Equation (1).

$$\Omega = \{Q_1, Q_2, \ldots Q_m\},\ Q_i = \begin{bmatrix} r_{1t1} & r_{2t1} \cdots & r_{st1} \\ r_{1t2} & r_{2t2} \cdots & r_{st2} \\ \vdots & \vdots & \vdots \\ r_{1tn} & r_{2tn} \cdots & r_{stn} \end{bmatrix} \quad (1)$$

where $Q_i$ represents rainstorm i, m is the number of rainstorms, $r_{stn}$ is the rainfall at $tn$ period of s rainfall station, $s$ = 1,2,3 … S, $tn$ = 1,2,3 … N, $S$ is the number of stations, and N is the number of periods. The main objective of this paper is to study the "mode" of rainfall, or may be called "structure", so the monitoring rainfall data is standardized by Equation (2).

$$x_{jtn} = \frac{r_{stn}}{\sum_{j-1}^{s} r_{jtn}} \times 100 \quad (2)$$

where $x_{jtn}$ represents the ratio of $j$ station rainfall to all station at $tn$ period, $S$ is the number of stations.

Figure 4 is a sample of visual representation of the standardized rainstorm matrix.

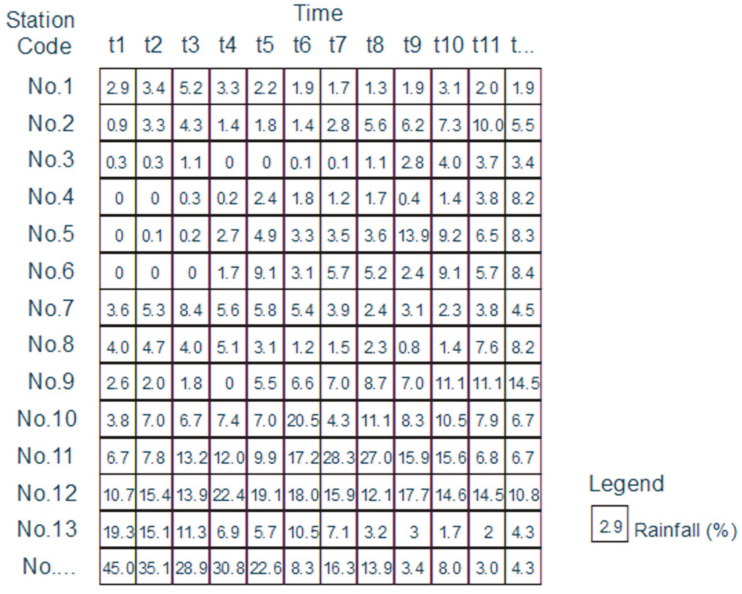



**Figure 4.** Sample of rainstorm matrix.

### *3.3. Dimensionality Reduction of High Dimensional Array*

As shown in Figure 4, the $Q_i$ is a $S \times$ N dimensional array. It was necessary to map the high-dimensional array to low dimensional space, in order to cluster and analyze arrays in Ω [22]. Principal component analysis (PCA) was used to reduce the dimensions of high-dimensional arrays [23].

Taking U represents the arrays in Ω, which is an n×m matrix, where n is the number of rainstorms, and m represents the number of characteristics, which equal S × N in this paper, including the monitoring data of each station at all periods. $Y_{n\times k}$ is the matrix converted from $X_{n\times m}$ by the transformation matrix $V_{m\times k}$, which means that the original data is converted from m dimensions to k dimensions (k ≪ m). The steps are as follows:

(1) The transformation matrix $X_{n\times m}$ was obtained by centralization of the matrix U,
(2) Calculation of covariance matrix $\sigma_{m\times m}$ of $X_{n\times m}$,
(3) Calculating the eigenvalues and eigenvectors of the variance matrix of M, M = $\sigma_{m\times m}$,
(4) The dimension $k$, which can retain more than 90% information of original data, was obtained. K eigenvectors constitute the transformation matrix $V_{m\times k}$ as column vectors,
(5) Descending dimensions by Equation (3).

$$Y_{n\times k} = X_{n\times m} \times V_{m\times k} \quad (3)$$

$Y_{n\times k}$ is the low dimensional matrix after transformation, which $n$ is the number of samples and $k$ is the number of dimensions of new matrix.

### *3.4. Clustering and Feature Selection*

After the dimensionality reduction of high-dimensional samples, the k-means clustering algorithm is used to classify low dimensional samples of $Y_{n\times k}$ [24].

(1) $r$ initial cluster centers are set up: $Z_1(p), Z_2(p), \ldots\ldots Z_r(p)$, where $p$ is the number of iterations.

(2) Calculating the distance from samples $x$ ($x \in X$) to each cluster center, if $D_x(j) = min\{D_x(i)\}$ $i = 1, 2, \cdots r$, then $x \in S_j$, where $S_j$ represents cluster $j$ with the center of $Z_j$.

(3) The new center of each cluster is calculated. The new center of $Z_j$ is calculated by Equation (4).

$$Z_j(p+1) = \frac{1}{N}\sum_{i=1}^{N} x_i \, , j = 1, 2, \cdots r \tag{4}$$

where $N$ is the number of samples contained in the cluster $S_j$, and $x_i$ is the sample in $S_j$. Using $Z_j(p+1)$ as the new cluster center, and the clustering criterion function can be minimized (Equation (5)).

$$J_j = \left[\sum_{x \in S_j(k)} x - z_j(k+1)^2\right]^{\frac{1}{2}} \tag{5}$$

where $j = 1, 2, \cdots K$.

(4) If $Z_j(p+1) \neq Z_j(p)$, $j$ = 1,2, ... $r$, then go to step (2); if $Z_j(p+1) = Z_j(p)$, $j$ = 1,2, ... $r$, the calculation is over.

In this paper, different $r$ values were calculated, and the initial values of different cluster centers were selected for the k-means cluster. Finally, the rainstorms were divided into three modes. The mean value of each rainstorm was taken as a typical mode of the rainstorm.

### *3.5. Reconstruction of Rainstorms*

With the inverse calculation of principal component analysis, the low dimensional array was reduced to a high-dimensional array to express the spatiotemporal process of each rainstorm. The i clustering centers are reconstructed into i×m matrices (Equation (6)).

$$X_{app} = Z_{i \times k} \times V_{m \times k}{}' \tag{6}$$

where $i$ is the number of rainstorm modes and $m$ is the dimension of the original data.

## 4. Results

The method of extracting rainfall temporal and spatial modes was applied to analyze the monitoring data of 22 rainstorms in the urban area of Beijing. It was found that the spatiotemporal distribution can be divided into three modes, and there are obvious differences in the center, spatial distribution, and occurrence time of three modes. The movement process of rainfall centers in different modes at different periods is shown in Figure 5. The centroid coordinates for a certain period was obtained by the method of calculating the geographical center with the weight of rainfall.

As shown in Figure 5, there are obvious differences in the three modes. The geographical centers of mode 1, mode 2, and mode 3 are located in northwest, southwest, and southeast, respectively. Mode 1 moves from the northwest to the urban center, mode 2 mainly spreads from the southwest and south to the north and the urban center, and mode 3 is basically concentrated in the urban center.

The matrixes of 3 modes are shown in Figure 6, which shows the characteristics of rainfall distribution. For example, the rainstorms belonging to mode 2 are more centralized than that belonging to modes and mode 3, which shows rainfall concentrated at several stations of 30748000, 30523900, 30504030, and 30523650. The rainfall spatial distribution at different periods of each mode is shown in Figures 7–9, which represents the percentage of rainfall at all rainfall stations in the period by the depth of the color in the location of the station.

As shown in Figure 6, the proportion of rainfall at each station is significantly different. In mode 1, rainfall is mainly concentrated in six stations, which are stations 3, 6, and 4. The spatial non-uniformity of rainfall in mode 2 is the most obvious, and only four stations

account for a large proportion of rainfall, which are stations 5, 9, 12, and 4. In mode 3, almost all stations have obvious rainfall, except station 1 and station 2.

In mode 1, as shown in Figures 5–7, rainstorm spreads from the northwest mountainous area to the central area of the city and the eastern part of the city. Rainfall began at the beginning of the northwest mountain area, and the rest of the city did not have any rainfall. Rainfall gradually dispersed and rainfall occurred at all stations. There are seven rainstorms of this mode, accounting for 31.8% of all sample. Among them, the single peak and homogeneity stations are 43%, and the double-peak type accounts for 14%.

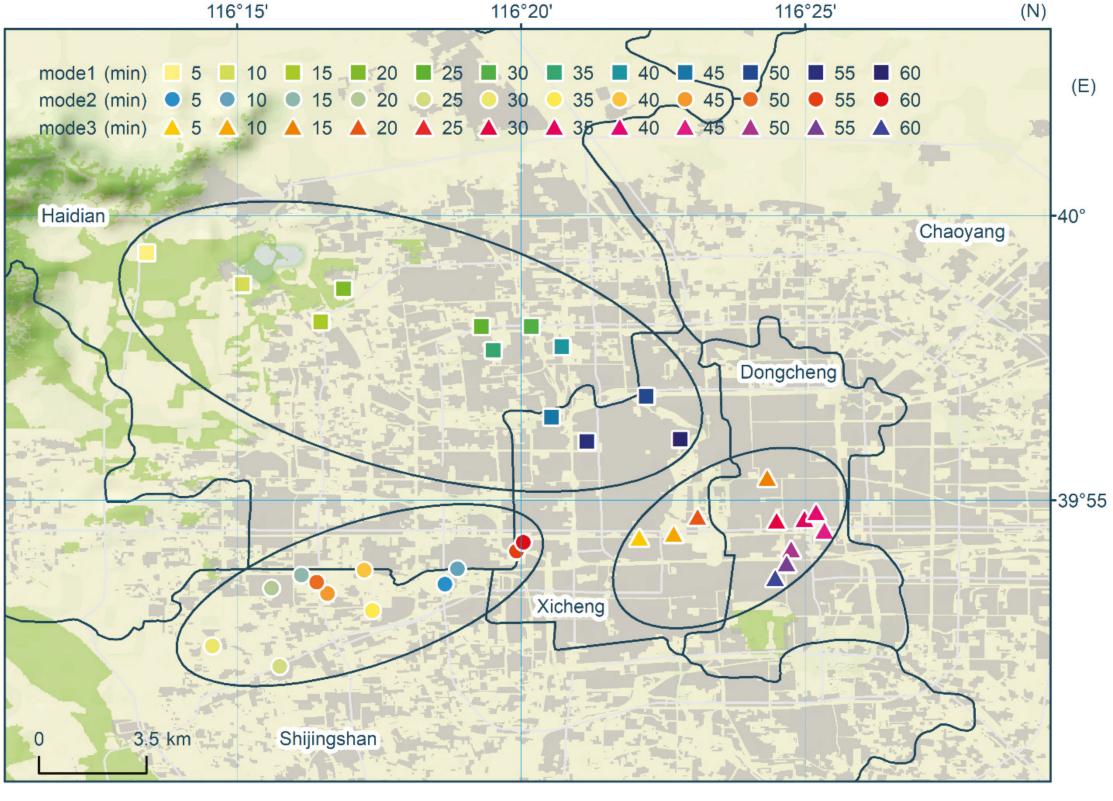



**Figure 5.** Centroids of three modes at different periods.

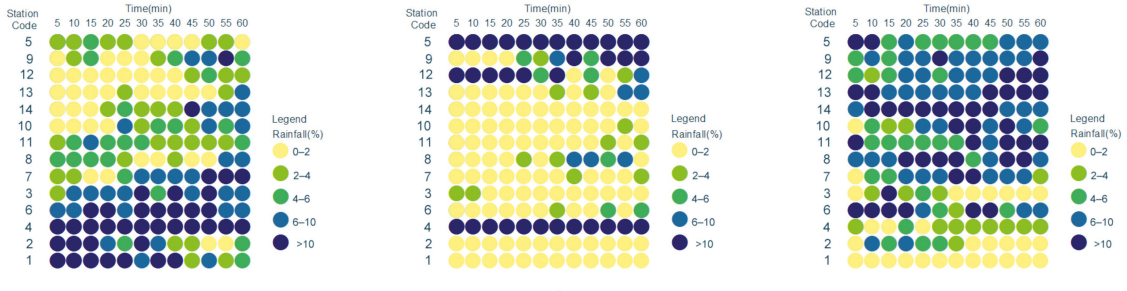



**Figure 6.** Matrixes of 3 modes. (**a**) Mode 1; (**b**) mode 2; (**c**) mode 3.

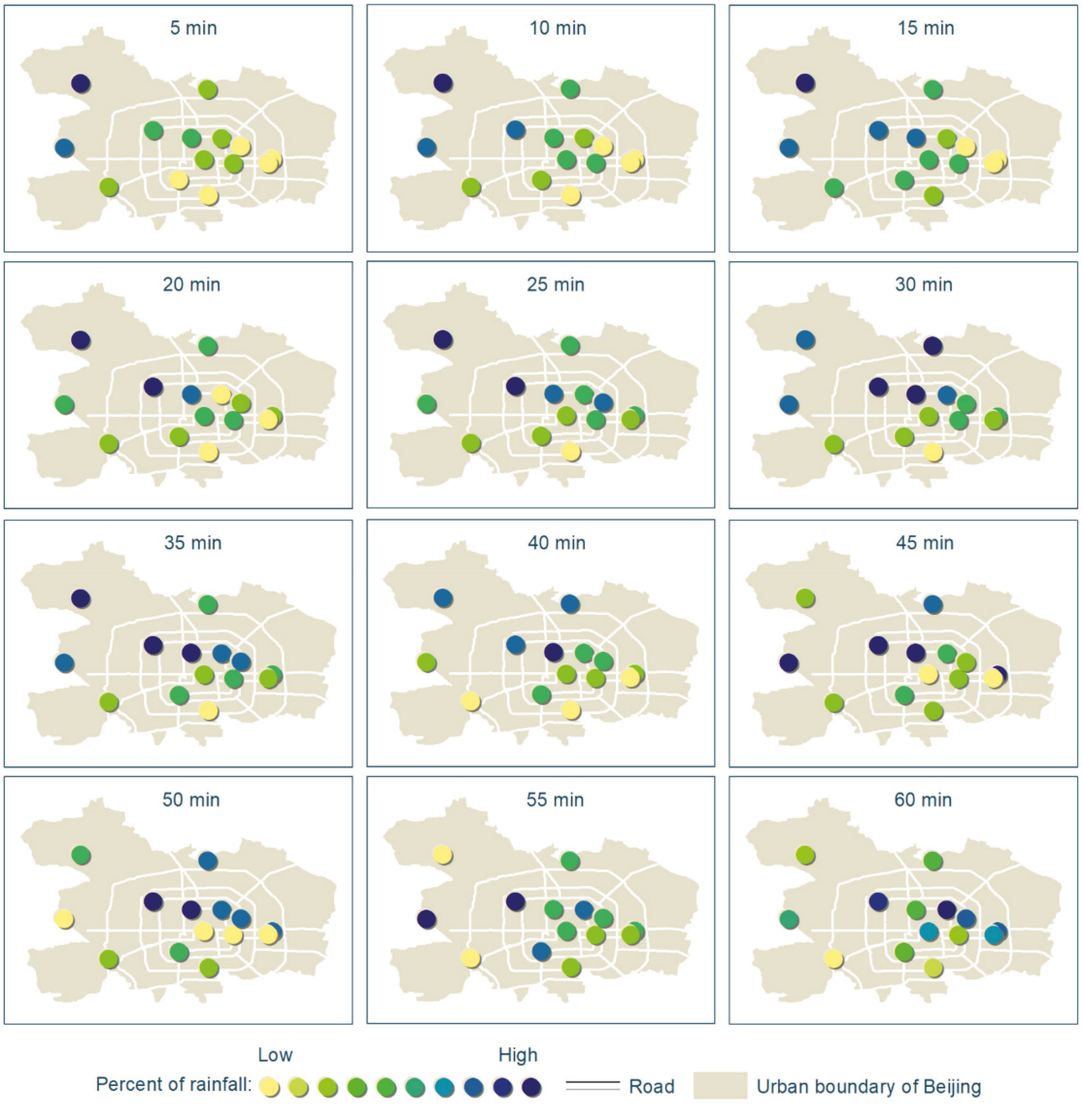



**Figure 7.** Spatiotemporal process of mode 1.

In mode 2, as shown in Figures 5, 6 and 8, the main rainfall was concentrated in the southern and southwestern regions of the city, gradually spreading to the northern and urban central areas, and there was no rainfall in the northwest mountain areas. There are three rainstorms of this mode, accounting for 13.7% of the total sample. Two of them are unimodal and one is homogeneous. In addition, this type of rainfall occurs between 14:00 and 16:00.

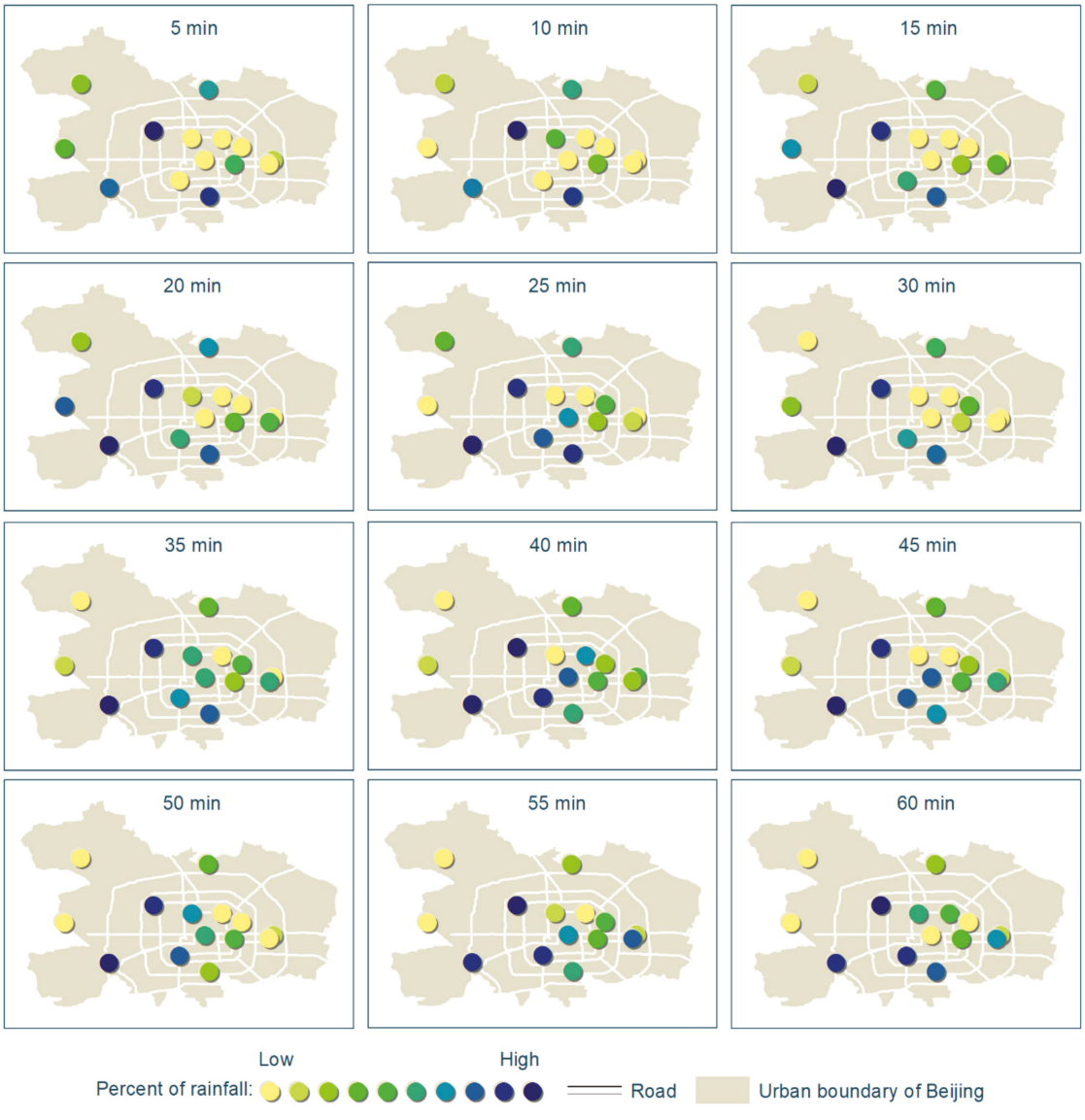



**Figure 8.** Spatiotemporal process of mode 2.

In mode 3, as shown in Figures 5, 6, and 9, the main rainfall was concentrated in the central area of the city and the eastern and southern parts of the city, which basically did not move. There was no rain in the northwest mountain areas. There were 12 rainstorms of this mode, accounting for 54.6% of total sample, among them, 62% were single peak type and 38% were homogeneity type. This mode was the main rainstorm type in summer in Beijing, which mainly occurred from afternoon to evening.

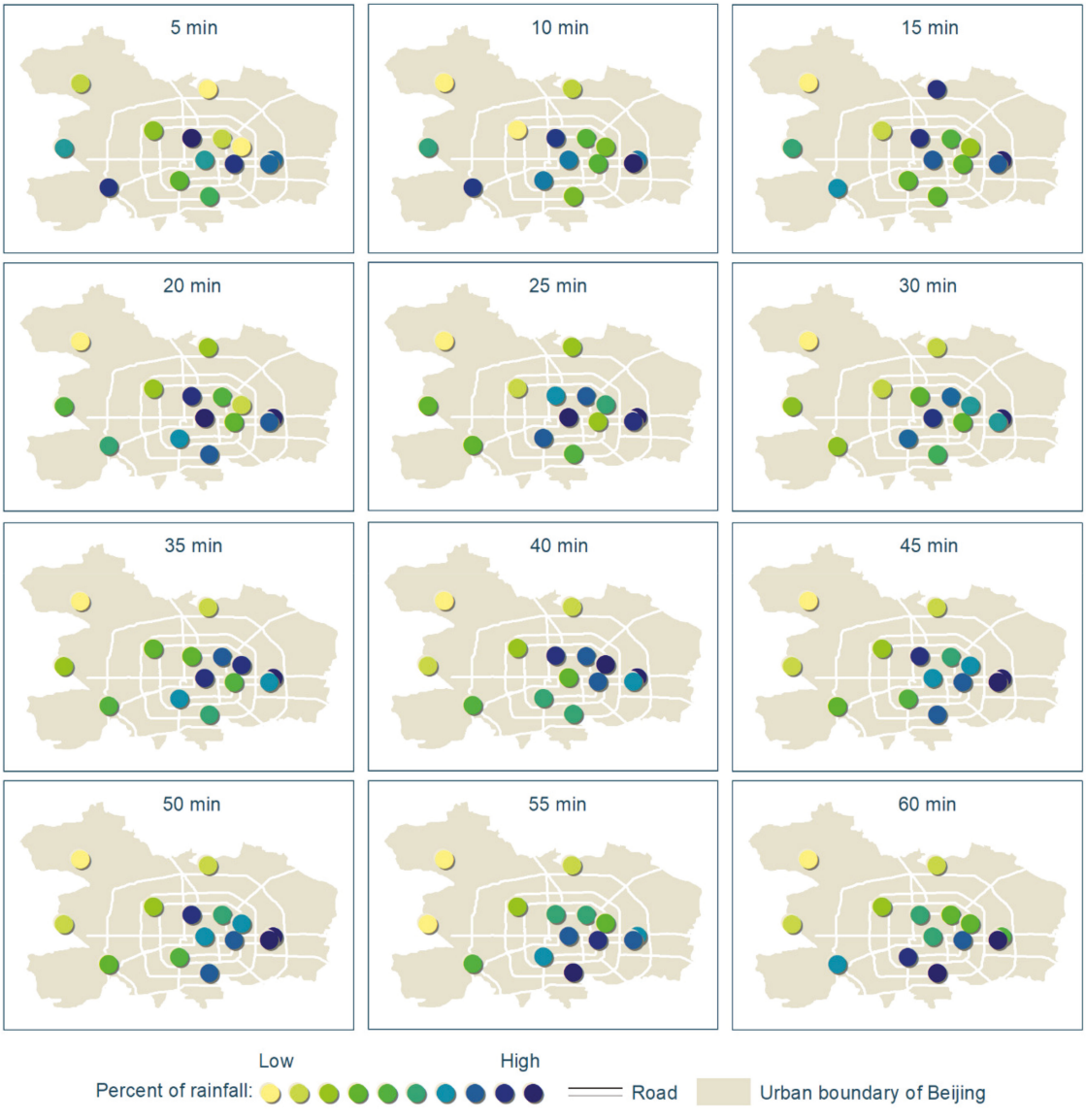



**Figure 9.** Spatiotemporal process of mode 3.

## 5. Discussion

How to identify and extract valuable information from multidimensional and massive rainfall monitoring data is a problem faced by many researchers. In this paper, a new approach for rainfall mode recognition is introduced, and different rainfall modes are identified from the massive monitoring data through the algorithms of dimensionality reduction, clustering, and reconstruction of a high dimensional array. This approach can be applied to both multidimensional data analysis and spatiotemporal data mining.

The results show that there are three modes of rainstorms in the Beijing urban area. Rainstorms of mode 1 moved from the northwest to the center of Beijing, then spread to the eastern part of the urban area; rainstorms of mode 2 occurred in the southwestern region

of the urban area, and gradually northward, but there was no rainfall in the mountainous northwest; rainstorms of mode 3 were concentrated in the central and eastern regions, and basically did not move. The results are consistent with the actual rainstorm process. This approach provides a framework for analysis, and there is uncertainty in some respects. For example, the result of the restructured rainfall has certain randomness and uncertainty in spatiotemporal distribution, which should be paid attention to in future. Because of the abundance of available data, this paper only selected the rainfall data of 14 years in Beijing. It should be noted that urban rainfall is a part of a larger range of rainfall in most cases, the temporal and spatial distribution characteristics of rainfall extracted from the rainfall data of 14 stations may have a certain randomness and uncertainty. More extensive rainfall data should be collected, considering terrain, climate, etc., according to the availability of data. In addition, it is necessary to distinguish the historical evolution of rainfall modes in different periods due to the climate changes in the city.

There are several suggestions for practice, First, the spatial and temporal resolution for rainfall data needs to reflect the temporal and spatial differences of a rainfall. Secondly, the study area should include the urban area and surrounding areas as much as possible, which is mainly to maintain the integrity of rainfall process.

## 6. Conclusions

In this paper, a high dimensional array is introduced for the study of the spatiotemporal distribution of rainfall, which describes rainfall by storing continuous rainfall monitoring data of all rainfall stations.

Through the establishment of high dimensional arrays of each rainstorm and algorithms, such as dimensionality reduction, clustering, feature extraction, and reconstruction, the spatiotemporal distribution of rainstorms in the flood season of Beijing city was analyzed. It was found that there were three spatiotemporal modes of rainstorms in the urban area of Beijing from 2004 to 2016. Rainstorms of mode 1 moved from the northwest mountain area to the central district, and further spread to the eastern part of the area. Rainstorms of mode 2 concentrated in the southwest of the urban area, gradually spreading to the northern and urban central areas; the northwest mountainous area basically had no rain. Rainstorms of mode 3 concentrated in the central area of the urban area and the eastern and southern regions, and basically did not move. The variation of the centroids of different modes shows a significant difference between the modes. The approach and conclusions in this paper can be applied to the study of rainfall modes in other cities or regions at a different scale, so as to provide assistance for rainfall forecasting and flood prevention.

The limitation of current machine algorithms is that it is too dependent on the number and quality of learning samples. If the rainfall stations are dense and the rainfall data quality is accurate, this method can achieve good results. If it is in an area with insufficient data and sparse rainfall stations, results may not be satisfactory. With the increasing density of rainfall stations, the improvement of rainfall data quality, and improvement of machine learning algorithm, the machine learning algorithm will get more reasonable and objective results.

**Author Contributions:** The study was designed by W.L. and F.T. The data were collected and analyzed by W.L. The paper was written by S.C. and F.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

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

## References


1. Willems, P.; Arnbjerg-Nielsen, K.; Olsson, J.; Nguyen, V. Climate change impact assessment on urban rainfall extremes and urban drainage: Methods and shortcomings. *Atmos. Res.* **2012**, *103*, 106–118. [CrossRef]
2. Xia, J.; Wei, S. Perspective on water security issue of changing environment in China. *J. Hydraul. Eng.* **2016**, *47*, 292–301.
3. Kong, F.; Shi, P.J.; Fang, J.; Lu, L.; Fang, J.; Guo, L. Advances and Prospects of Spatiotemporal Pattern Variation of Extreme Precipitation and its Affecting Factors under the Background of Global Climate Change. *J. Catastrophology* **2017**, *32*, 165.
4. Wang, B.; Zhao, L.; Gong, Y. Characteristics of temporal pattern and return period of short-duration rainfall at Beijing Observatory. *Torrential Rain Disasters* **2015**, *35*, 302–308.
5. Rafieeinasab, A.; Norouzi, A.; Kim, S.; Habibi, H.; Nazari, B.; Seo, D.-J.; Lee, H.; Cosgrove, B.; Cui, Z. Toward high-resolution flash flood prediction in large urban areas—Analysis of sensitivity to spatiotemporal resolution of rainfall input and hydrologic modeling. *J. Hydrol.* **2015**, *531*, 370–388. [CrossRef]
6. Zhangchun, W.; Guoping, C.; Zhimin, A. Experimental study on overland flow. *J. Hydraul. Eng.* **1995**, *26*, 84–89.
7. Guan, M.; Sillanpää, N.; Koivusalo, H. Storm runoff response to rainfall pattern, magnitude and urbanization in a developing urban catchment. *Hydrol. Process.* **2016**, *30*, 543–557. [CrossRef]
8. Guoping, C.; Jin, S.; Rongsheng, F. Research on Rainfall Pattern of Urban Design Storm. *Adv. Water Sci.* **1998**, *9*, 41–46.
9. Keifer, C.J.; Chu, H.H. Synthetic Storm Pattern for Drainage Design. *J. Hydraul. Div.* **1957**, *83*, 1332-1–1332-25. [CrossRef]
10. Kibler, D.F. *Urban Stormwater Hydrology*; American Geophysical Union: Washington, DC, USA, 1982.
11. Song, X.; Zhang, J.; Zou, X.; Zhang, C.; AghaKouchak, A.; Kong, F. Changes in precipitation extremes in the Beijing metropolitan area during 1960–2012. *Atmos. Res.* **2019**, *222*, 134–153. [CrossRef]
12. Wei, Z.; Changhe, L.U. Assessing Changes in Rainstorms in Beijing During the Last 50 Years. *J. Resour. Ecol.* **2016**, *7*, 372–377. [CrossRef]
13. Hutchinson, M.F. *ANUSPLIN Version 4.3*; Centre for Resource and Environment Studies, The Australian National University: Canberra, Australia, 2004.
14. Wenhai, Z.; Lei, L. An initial application of artificial intelligence on the detection and nowcasting of hail weather. *ACTA Meteorol. Sin.* **2019**, *77*, 282–291.
15. Liu, Y.Y.; Li, L.; Liu, Y.S.; Chan, P.W.; Zhang, W.H. Dynamic spatial-temporal precipitation distribution models for short-duration rainstorms in Shenzhen, China based on machine learning. *Atmos. Res.* **2020**, *237*, 104861. [CrossRef]
16. Zhang, Y.; Li, Q.; Huang, H.; Wu, W.; Du, X.; Wang, H. The Combined Use of Remote Sensing and Social Sensing Data in Fine-Grained Urban Land Use Mapping: A Case Study in Beijing, China. *Remote. Sens.* **2017**, *9*, 865. [CrossRef]
17. Liu, X.; Tian, G.; Feng, J.; Ma, B.; Wang, J.; Kong, L. Modeling the Warming Impact of Urban Land Expansion on Hot Weather Using the Weather Research and Forecasting Model: A Case Study of Beijing, China. *Adv. Atmos. Sci.* **2018**, *35*, 723–736. [CrossRef]
18. Zhicong, Y.; Guo, W.; Li, N.; Xie, Y. Numerical Simulation of Urban Ponding in Beijing. *Meteorol. Mon.* **2015**, *41*, 1111–1118.
19. Gu, X.; Li, N.; Zhou, Y.; Wu, J. Thinking on urban waterlogging disaster defense initiated by "7 • 21" extraordinary rainstorm in Beijing. *J. Nat. Disasters* **2013**, *22*, 1–6.
20. Weidong, L.; You, H.; Ren, G.; Yang, P.; Zhang, B. Subtle Precipitation Characteristics in Beijing Area. *Clim. Environ. Res.* **2014**, *19*, 61–68.
21. Zhang, T.; Detang, L.; Daolun, L.; Yi, D. Research on Statistical Information Reconstruction of Images Based on Multiple-Point Geostatistics Integrating Soft Data with Hard Data. *J. Comput. Res. Dev.* **2010**, *47*, 573–578.
22. Jolliffe, I.T. *Principal Components Analysis*; Wiley Online Library: New York, NY, USA, 2002.
23. Mitchell, T. *Machine Learning*; China Machine Press: Beijing, China, 2003; pp. 417–433.
24. Ng, R.T.; Han, J. Efficient and Effective Clustering Methods for Spatial Data Mining. In Proceedings of the VLDB Conference, Santiago, Chile, 12–15 September 1994.

water

MDPI

*Article*

# The Impact of Rainfall Movement Direction on Urban Runoff Cannot Be Ignored in Urban Hydrologic Management

**Yesen Liu [1], Yaohuan Huang [2,*], Yuanyuan Liu [1], Kuang Li [1] and Min Li [1]**

[1] State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing 100038, China; liuys@iwhr.com (Y.L.); lyyiwhr@gmail.com (Y.L.); likuang0000@163.com (K.L.); limin@iwhr.com (M.L.)

[2] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

* Correspondence: huangyh@igsnrr.ac.cn; Tel.: +86-22-2740-1156

**Abstract:** Urban floods have been exacerbated globally, associated with increasing spatial-temporal variations in rainfall. However, compared with rainfall variabilities of intensity and duration, the effect of rainfall movement direction is always ignored. Based on 1313 rainfall scenarios with different combinations of rainfall intensity and rainfall movement direction in the typically rainy city of Shenzhen in China, we find that the effect of rainfall movement direction on the peak runoff may reach up to 20%, which will decrease to less than 5% under heavy rainfall intensity conditions. In addition, our results show that the impact of rainfall movement direction is almost symmetrical and is associated with the direction of the river. The closer rainfall movement direction is to the Linear Directional Mean of rivers, the larger is the peak runoff of section. Our results reveal that rainfall movement direction is significant to urban peak runoff in the downstream reaches, which should be considered in urban hydrological analysis.

**Keywords:** urban floods; rainfall movement direction (RMD); rainfall intensity (RI); peak runoff; Linear Directional Mean (LDM); Shenzhen

**Citation:** Liu, Y.; Huang, Y.; Liu, Y.; Li, K.; Li, M. The Impact of Rainfall Movement Direction on Urban Runoff Cannot Be Ignored in Urban Hydrologic Management. *Water* **2021**, *13*, 2923. https://doi.org/10.3390/w13202923

Academic Editor: Renato Morbidelli

Received: 12 September 2021
Accepted: 15 October 2021
Published: 17 October 2021



## 1. Introduction

Extreme rainstorms have been increasing in urban areas due to continued global warming and rapid urbanization [1–4].This increase is expected to lead to an increment in urban runoff generation and, consequently, to the intensification of urban flood risks [5,6]. During 2010–2018, more than 160 cities in China suffered from floods each year [7]. Hydrological processes in urban areas are sensitive to small-scale temporal and spatial variations in rainfall [8]. Furthermore, the probability of extreme rainfall in urban areas is very likely to increase, with high spatial and temporal variabilities [9,10], yielding further uncertainties in urban runoff estimations and flood-related damages [11,12]. Understanding the impacts of spatial and temporal rainfall variabilities on runoff is significant for urban hydrologic management (e.g., urban drainage design and construction, forecasting and prevention of flood risk) [13–15] and the development of an Early Warning System (EWS) [16,17].

However, the interactions among extreme rainfall variability, river features and runoff responses remain poorly understood, especially in urban areas [11,18]. Such attributions require sufficient information about the spatial distribution of short-term rainstorms and runoff responses, which is lacking in measurements for the sudden rainfall process and the complex inhomogeneity of urban areas introduced by the building envelope. With the development of new instruments, techniques and methods for capturing rainfall and hydrological processes at high resolution, urban hydrological models, such as the Storm Water Management Model [19], have been proposed and applied in urban hydrologic management [20,21]. Although various components of rainfall variability such as the RI, rainfall duration spatial and temporal resolution, and degree of imperviousness are

involved in previous studies, these components are far from complete. For example, similar extreme rainfall patterns of RI and rainfall duration in the same urban area may induce different flood damages. As a result, there remain critical errors and uncertainties in the impacts of rainfall variabilities on hydrological process such as runoff and floods [22].

Rainfall movement direction (RMD) is a significant component of rainfall variability [23–29] that is always ignored. While some studies have found no impact of RMD on hydrological responses [8,30], relatively few studies reported in the literature have remained inconclusive with respect to the impact of RMD on urban runoff. This contrast may be explained by (1) the limited RMDs observed from rainstorm events, which is almost fixed rather than 360 degree and lacks information in an urban area; (2) isolated analyses of the impacts of RMD were conducted in these studies by neglecting the interactions between RMD and the directions of urban river segments.

In this study, we focus on verifying whether RMDs play a significant role in generating peak runoff in urban areas. This paper is organized as follows: Section 2 introduces the study area and the experimental design. Results and discussion are given in Section 3, which also summarizes the main findings of the impact of RMD on runoff.

## 2. Materials and Methods

### *2.1. Study Area*

The typically rainy city of Shenzhen in China, with four urban rivers, is selected as the study area [31], this city has complete hydrological infrastructure and observation information systems. Shenzhen, located in the southeast coast of China, has the most rapid urbanization of any city in China with high urban flood risk caused by extreme rainfall. Shenzhen city ranks fifth among the 136 coastal cities in the world in terms of future flood loss risk [32]. Shenzhen city is a typical case study area, in which the spatial and temporal variations in rainfall is wide with rapid movement and obvious "squall line" [33]. Additionally, urban hydrological datasets are sufficient in Shenzhen, which is the benefit from a 30 million RMB project of building an urban flood model, which began in 2018. These datasets include underground pipe network data (nearly 3000 km), 1:1000 terrain data, 13 years of rainfall monitoring data, historical water level data, reservoir operation data, and more.

### *2.2. Methods*

#### 2.2.1. Construction of Rainfall Schemes

We designed the idealized experimental conditions of rainfall intensity (RI) and RMD to construct comprehensive rainfall schemes. According to the RI value of 132.7 mm/h recorded once every 1000 years in Shenzhen, we constructed 13 RIs with 10 mm intervals from 10 mm/h to 130 mm/h, and chose the commonly used Chicago rain pattern (Table 1) to represent short-term rainfall. In addition, 100 moving directions with equal intervals were designed to reflect the continuous RMD. We collected rainfall monitoring data from 63 meteorological stations in Shenzhen from 2008 to 2018, with a time resolution of 5 min. The rainfall events were extracted from the data and the rainfall center of every 5 min is calculated, then the moving speed of each rainfall events is calculated through the movement of the center. The average speed of all rainfall events is about 10 km/h.The asynchronous rainfall process of the whole basin is set to 3 h based on the average moving speed of the rainfall center (10 km/h) and the diameter of the circumscribed circle of the basin (30 km). By combining RI and RMD, 1300 rainfall processes (13 RIs $\times$ 100 RMDs = 1300) were constructed in Figure 1. The two rainfall distribution maps in Figure 1 show distributions when the RI is 80 mm and the RMDs are 15 and 80 respectively.

**Table 1.** Rainfall process of 13 rainfall intensity schemes.

| RI Scheme | Total Rainfall | Rainfall of 5 min Intervals | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 |
| RI1: | 10 | 0.4 | 0.6 | 0.8 | 1.4 | 2.0 | 1.2 | 0.9 | 0.7 | 0.6 | 0.5 | 0.5 | 0.4 |
| RI2: | 20 | 0.9 | 1.1 | 1.6 | 2.8 | 4.0 | 2.4 | 1.8 | 1.4 | 1.2 | 1.0 | 0.9 | 0.8 |
| RI3: | 30 | 1.3 | 1.7 | 2.4 | 4.3 | 5.9 | 3.7 | 2.7 | 2.1 | 1.8 | 1.6 | 1.4 | 1.2 |
| RI4: | 40 | 1.8 | 2.3 | 3.2 | 5.7 | 7.9 | 4.9 | 3.6 | 2.9 | 2.4 | 2.1 | 1.8 | 1.7 |
| RI5: | 50 | 2.2 | 2.8 | 4.0 | 7.1 | 9.9 | 6.1 | 4.5 | 3.6 | 3.0 | 2.6 | 2.3 | 2.1 |
| RI6: | 60 | 2.7 | 3.4 | 4.7 | 8.5 | 11.9 | 7.3 | 5.4 | 4.3 | 3.6 | 3.1 | 2.8 | 2.5 |
| RI7: | 70 | 3.2 | 4.0 | 5.5 | 9.6 | 13.2 | 9.3 | 6.2 | 5 | 4.2 | 3.7 | 3.3 | 3.0 |
| RI8: | 80 | 3.9 | 4.7 | 6.3 | 10.7 | 14.6 | 10.3 | 7.1 | 5.8 | 5.0 | 4.4 | 4.0 | 3.6 |
| RI9: | 90 | 4.5 | 5.4 | 7.1 | 11.8 | 16.1 | 11.3 | 7.9 | 6.5 | 5.7 | 5.1 | 4.6 | 4.3 |
| RI10: | 100 | 5.1 | 6.1 | 7.9 | 12.9 | 17.6 | 12.3 | 8.7 | 7.3 | 6.4 | 5.7 | 5.2 | 4.8 |
| RI11: | 110 | 5.8 | 6.7 | 8.6 | 13.9 | 18.8 | 14.1 | 9.5 | 8.0 | 7.0 | 6.3 | 5.9 | 5.5 |
| RI12: | 120 | 6.3 | 7.4 | 9.4 | 15.2 | 20.7 | 15.3 | 10.3 | 8.7 | 7.6 | 7.0 | 6.4 | 6.0 |
| RI13: | 130 | 6.8 | 8.0 | 10.2 | 16.4 | 22.4 | 16.4 | 11.2 | 9.4 | 8.4 | 7.6 | 6.9 | 6.5 |

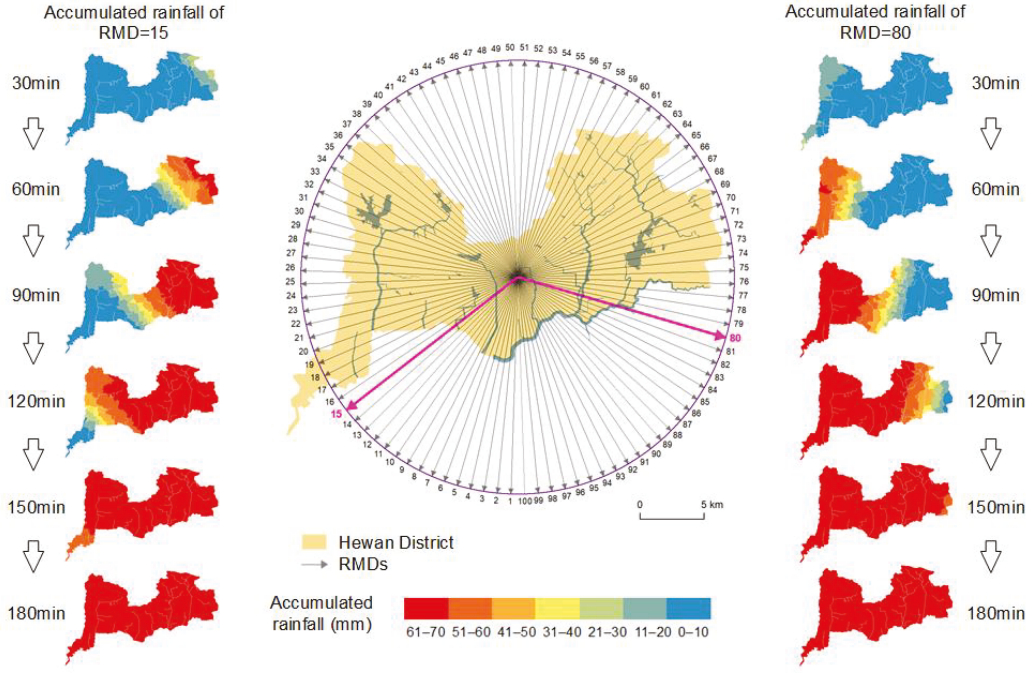



**Figure 1.** Schematic diagram of rainfall schemes.

2.2.2. Runoff Numerical Simulation

We applied the IFMS/urban (Integrated Urban Flood Modeling System) platform, an urban flood simulation platform developed by China Institute of Water Resources and Hydropower Research, to simulate the runoff processes of the studied rivers (Figure 2). Before the simulations, the model was calibrated by the rainfall–runoff relationship measured from 2018 to 2020. Based on 1313 rainfall processes, we constructed a set of input conditions for the model, including the initial water level of each section, the previous rainfall, and the infiltration process; these conditions were kept consistent for comparison. We calculated the runoff processes of 414 sections in each scheme by using a parallel computing program and output the result to structured files in CSV format.


**Figure 2.** Location of the four rivers.

IFMS/urban model is a flood analysis model developed in China, which has been successfully applied in Shenzhen, Beijing, Chengdu, Shanghai, Foshan and other cities [34,35]. Its main modules include: (1) Flood simulation and analysis: River, lake, nearshore flow, flood simulation analysis and calculation of flood protection area and flood storage area; urban rainstorm and waterlogging, overflow (break) flood and dam break flood simulation Analysis; storm surge simulation and analysis; (2) Engineering scheduling simulation: reservoir and lake gate dam, weir, culvert, box culvert and pumping station and other water conservancy engineering facilities scheduling simulation; pipeline one-way valve control; lake reservoir, underground storage, tank storage, and so on. (3) Data management and pre-processing functions: 2D and 3D data and result display platform; data pre-processing (grid generation, pipeline processing, data management).

#### 2.2.3. Index of Peak Runoff Deviation

We designed a concise index to describe the peak runoff deviation (Equation (1)):

$$I_{fp} = \frac{Q_{fp}}{Q_p} \tag{1}$$

where $I_{fp}$ is the peak runoff deviation of a section caused by the rainfall scheme with RI = $p$ and RMD = $f$, $Q_p$ is the peak runoff ($m^3/s$) of the section under the condition of a synchronous rainfall scheme with RI = $p$, and$Q_{fp}$ is the peak runoff ($m^3/s$) with RI = $p$ and RMD = $f$, where $f \in 1, 2, \ldots, 100$.

In this paper, the $I_{fp}$ values of 414 sections were extracted from all 1300 rainfall processes.

#### 2.2.4. Rainfall Movement Direction and Flow Concentration Direction

The trend in a set of line features is measured by calculating the average angle of the lines, which called the "Linear Directional Mean"(LDM) in GIS [36,37]. The LDM often

represents the paths of objects that move, such as rainfall. We use the LDM to represent the confluence direction above a certain river section (Equation (2)):

$$\mathrm{LDM} = \arctan\frac{\sum_{i=1}^{n}\sin\theta_i}{\sum_{i=1}^{n}\cos\theta_i} \tag{2}$$

where $\theta_i$ is the angle between section $i$-1 and section $i$ in the confluence area.

The $LDM$ is adjusted according to the angle quadrant, and the adjusted $LDM$ is between 0° and 360°:

$$\mathrm{LDM} = \begin{cases} LDM, \sum_{i=1}^{n}\sin\theta_i \geq 0 \; and \; \sum_{i=1}^{n}\cos\theta_i > 0 \\ 180 - LDM, \sum_{i=1}^{n}\sin\theta_i \geq 0 \; and \; \sum_{i=1}^{n}\cos\theta_i < 0 \\ 360 - LDM, \sum_{i=1}^{n}\sin\theta_i < 0 \; and \; \sum_{i=1}^{n}\cos\theta_i > 0 \\ 180 + LDM, \sum_{i=1}^{n}\sin\theta_i < 0 \; and \; \sum_{i=1}^{n}\cos\theta_i < 0 \end{cases}$$

The angle between the $LDM$ and RMD is calculated as Equation (3):

$$\Delta_{LDM} = \begin{cases} \left|D_{rainfall} - LDM_{river}\right|, \left|D_{rainfall} - LDM_{river}\right| \leq 180 \\ 360 - \left|D_{rainfall} - LDM_{river}\right|, \left|D_{rainfall} - LDM_{river}\right| > 180 \end{cases} \tag{3}$$

where $D_{rainfall}$ represents RMD, $LDM_{river}$ represents the LDM, and $\Delta_{LDM}$ is between 0 and 180°.

2.2.5. Dynamic Clustering of Sections

We use the dynamic clustering machine learning method to classify the $I_{fp}$ values of all sections. The $I_{fp}$ value of each section in different RMD with same RI is used as the attribute to construct the sample set. The calculation is as follows:

$$S_i = \left\{ I_{f(1)p}, I_{f(2)p}, \ldots I_{f(100)p} \right\} \tag{4}$$

$$\Omega = \{S_1, S_2, \ldots S_{429}\} \tag{5}$$

where $S_i$ is section i and $I_{f(1)p}$ is the simulated result with RMD = $f(1)$ and RI = $p$.

The dynamic cluster analysis method is used to classify the sections in Ω; then, the features of each subset, that is, the typical features influencing the peak runoff of the section, are extracted.

The k-means clustering algorithm is used to classify low dimensional with 4 steps:

(1) A total of r initial cluster centers is set up: $Z_1(p), Z_2(p), \ldots\ldots Z_r(p)$, where $p$ is the number of iterations.

(2) The distance from samples x ($x \in X$) to each cluster center is calculated,

$$\text{if } D_x(j) = min\{D_x(i)\}\ i = 1, 2, \cdots \text{r, then } x \in S_j$$

where $S_j$ represents cluster j with the center of $Z_j$.

(3) The new center of each cluster is calculated. The new center of $Z_j$ is calculated as follows.

$$Z_j(p+1) = \frac{1}{N}\sum_{i=1}^{N} x_i\ , j = 1, 2, \cdots r$$

where $N$ is the number of samples contained in cluster $S_j$ and $x_i$ is sample $i$. using $Z_j(p+1)$ as the new cluster center, the clustering criterion function can be minimized:

$$J_j = \left[ \sum_{x \in S_j(k)} \| x - z_j(k+1) \|^2 \right]^{\frac{1}{2}}$$

where $j = 1, 2, \cdots, K$.

(4) If $Z_j(p+1) \neq Z_j(p)$, $j = 1, 2, \ldots,$ r, then go to step (2); if $Z_j(p+1) = Z_j(p)$, $j = 1, 2, \ldots,$ r, then cease the iteration.

In this paper, different r values are calculated, and the initial values of different cluster centers are selected for the k-means cluster analysis. Then, the rainstorms are divided into 3 modes. The mean value of each rainstorm is taken as a typical rainstorm mode.

In addition, we reconstructed the typical features of the river sections. With the inverse calculation of the principal component analysis, the low-dimensional array is reduced to a high-dimensional array to express the spatiotemporal process of each rainstorm. The $i$ clustering centers are reconstructed into $i \times m$ matrices (Equation (6)):

$$X_{app} = \mathrm{Z}_{i \times p} \times \mathrm{V}_{m \times p}{}' \quad (6)$$

where $i$ is the number of rainstorm modes and $m$ is the dimension of the original data.

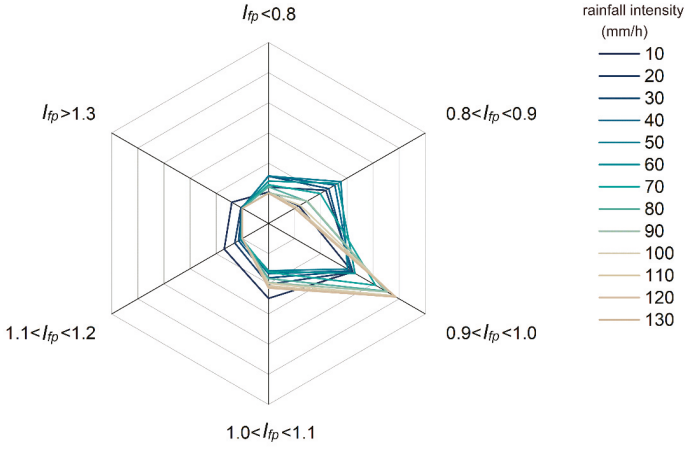
## 3. Results and Discussion

### *3.1. Influence of Variation in RI and RMD Combinations on the Peak Runoff*

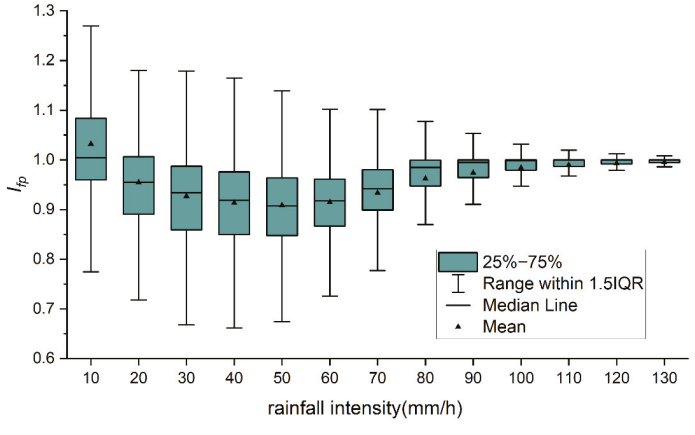
We found that the influence of variation in RI and RMD combinations on the peak runoff can reach 30% based on the *index of peak runoff deviation ($I_{fp}$)* in the study area. The statistics of 538,200 $I_{fp}$s show that the variations in the RI and RMD are more likely to cause negative effects on peak runoff, with 77.1% of $I_{fp}$s < 1.0 and 22.9% of $I_{fp}$s > 1.0. However, this influence is always ignored, possibly due to most $I_{fp}$s (83.3%) ranging from 0.9 to 1.1, indicating that the effect of these variations is smaller than 10% in most instances. Additionally, a heavy RI will reduce the effect of RMD on the peak runoff, with more concentrated of $I_{fp}$s value of approximately 1.0 correspondingly with the more enhanced RI. We found that 96% of $I_{fp}$s values ranged from 0.9 to 1.1 when RI > 70mm/h. Furthermore, the range of $I_{fp}$s decreased to 0.95–1.05 when the RI reach an extreme value larger than the once-in-one-hundred-years value (100 mm/h). We also found that the duration of the peak runoff was lengthened with extremely high RI values (Figure 3), which is associated with the processes of drainage networks operating under maximum waterlogging-elimination capacities. Figure 3A shows that the heavier the RI is, the more concentrated the $I_{fp}$ is between 0.9 and 1.1, and the number of $I_{fp}$ values less than 1.0 is obviously greater than that of $I_{fp}$ values more than 1 for all RIs. Figure 3B shows that with an increase in the RI, the fluctuation range of $I_{fp}$ decreases obviously and tends toward 1.0.

### *3.2. Influence of Variation in RMD on the Peak Runoff*

To estimate the impact of RMD, we isolate the effect of RMD from the RI by analyzing the spatial pattern of the RMD impact under similar RI. Based on the dynamic clustering of 414 sections with 100 RMDs each, we do find three typical patterns with proportions of 42%, 21% and 37% of total number of sections, including (1) model 1 shows a nondirectional effect of RMD by reducing the peak runoff with most $I_{fp}$s < 1.0; (2) model 2 shows an obvious directional effect of RMD with a symmetrical distribution of $\mathrm{I_{fp}}$s, in which the minimum effect of direction is opposite to direction of the maximum; (3) model 3 shows completely no obvious effect of RMD on peak runoff neither in direction nor in magnitude with $I_{fp}$s approximately equal to 1.0 (Figure 4).




(**A**)




(**B**)

**Figure 3.** Influence of the rainfall intensity (RI) and rainfall movement direction (RMD) on the peak runoff. (**A**) The distribution of $I_{fp}$ values under 13 different RIs. (**B**) The distribution of $I_{fp}$ under different RIs.

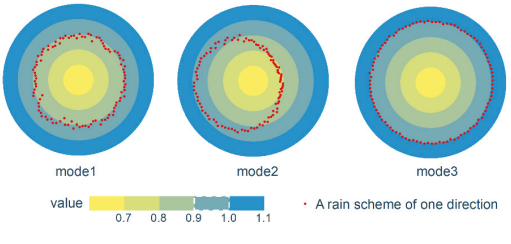



**Figure 4.** Distribution of $I_{fp}$s for typical modes.

Spatially, these three models are primarily related to the location of sections in rivers (Figure 5).

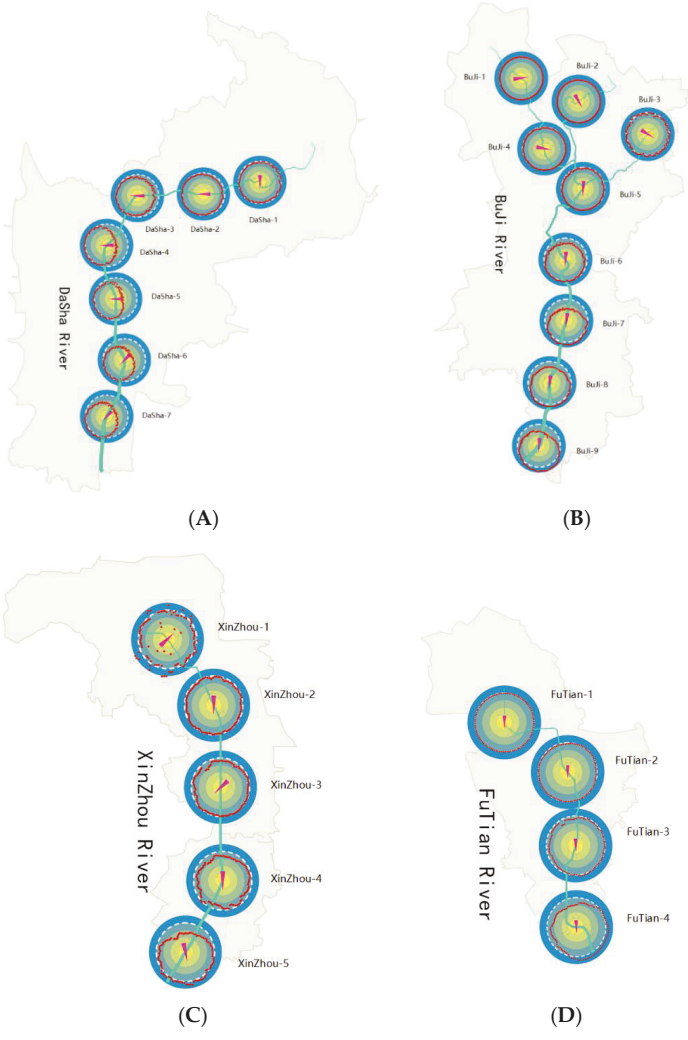



**Figure 5.** Influence of the peak runoff on the sections of four rivers. (**A**) Dasha River. (**B**) Buji River. (**C**) Xinzhou River. (**D**) Futian River. Only certain parts of the 414 sections are drawn.

Figure 5 shows that the in lower reaches of the rivers, the larger the variation range of $I_{fp}$ is, the more symmetrical is the influence. In model 1 and model 3, the influence of RMD on the peak runoff is negligible in the upper reaches of the river. However, model 2 shows symmetrical variations in the lower reaches of the river. We also find that the river length potentially influences the RMD effect. Taking four rivers in Shenzhen as an example, the Dasha River and Buji River show the effect of model 2 from middle reaches to lower reaches. While, this effect occurs at the ends of the lower reaches in the other two rivers, Xinzhou River and Futian River, the lengths of which are just half of the two aforementioned rivers. The impact of RMD on the peak runoff is primarily present in the lower reaches, possibly due to the longer duration and larger area of the flow concentration that gradually increased the asynchrony of rainfall with the peak runoff.

Furthermore, we find a slightly jagged shape in models of $I_{fp}$s compared with the natural watershed, which may be explained by the uneven distribution of underground pipes used to drain flows into rivers in urban areas and their corresponding covering areas. In rapidly urbanizing area, the lack of pipe network datasets will increase the complexity of urban flood analyses.

### *3.3. How RMD Affect the Peak Runoff across Rivers*

To further understand how RMD affect the peak runoff, we calculated three angles of sections: $\theta_{maxPR\text{-}FC}$ (the angle between RMD with maximum peak runoff and flow concentration direction), $\theta_{min\ PR\text{-}FC}$ (the angle between RMD with minimum peak runoff and flow concentration direction) and $\theta_{max\ PR\text{-}FD}$ (the angle between RMD with maximum peak runoff and flow direction). In this paper, we used "Linear Directional Mean" (LDM) to represent the flow concentration direction, which is the geometric mean of all the reaches upstream. In contrast with the flow direction, we do find a significant relationship between RMD and the flow concentration direction during maximum peak runoff (Figure 6), with decreasing trends of $\theta_{maxPR\text{-}FC}$ from upstream to downstream. For example, the angles of the second half of downstream reaches of the Buji River are almost smaller than 20°, and some sections even reach 0°. In contrast, $\theta_{minPR\text{-}FC}$ is gradually increased with the opposite RMD to $\theta_{maxPR\text{-}FC}$, which is consistent with the result of model 2 shown in Figure 4. Additionally, the $\theta_{max\ PR\text{-}FC}$ values are larger in more meandering rivers (e.g., Dasha River) than in rivers with straight channels (e.g., Buji River). We can therefore conclude that RMD is a significant factor to peak runoff downstream in urban areas, and this influence is not isolated but needs to be combined with the spatial feature of rivers such as direction and bending.

The impact of RMD on the peak runoff indicates that river flood risks and discharge capacities should be evaluated from more rainfall variations including RI, rainfall duration, etc. Commonly used methods that do not consider the impact of RMD, such as constructing rainfall processes by independent zoning [34,38,39], may underestimate or overestimate the peak runoff magnitude in rivers. Figure 7 compares the difference of peak runoff between considering and not considering the impact of RMD, in four urban rivers of Shenzhen with similar RIs. We found that the uncertainties yielded from RMD may reach −40% and 50% (Figure 7). In most cases, these uncertainties range from −20% to 20%. These analyses confirm that the impact of RMD on runoff cannot be ignored in urban hydrologic management.

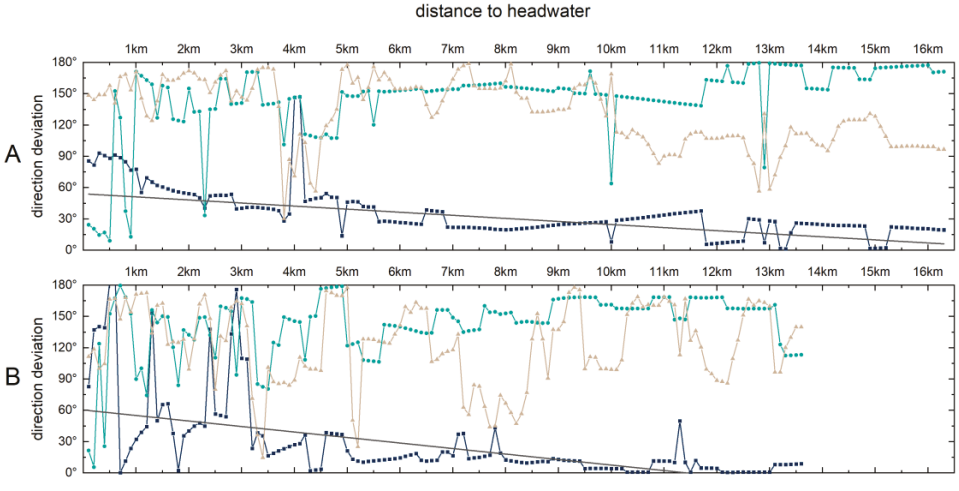



**Figure 6.** *Cont.*

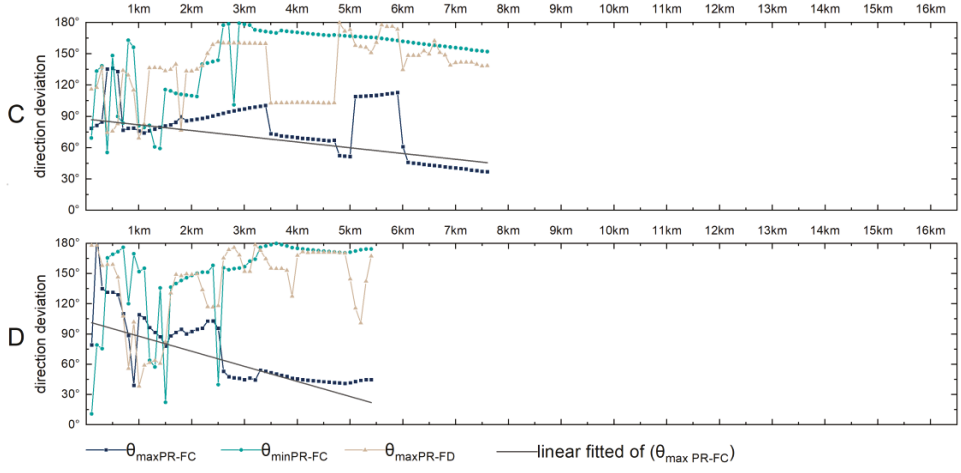



**Figure 6.** Influences of rainfall movement direction (RMD) and the direction of river confluence on the peak runoff. (**A**) Dasha River. (**B**) Buji River. (**C**) Xinzhou River. (**D**) Futian River.

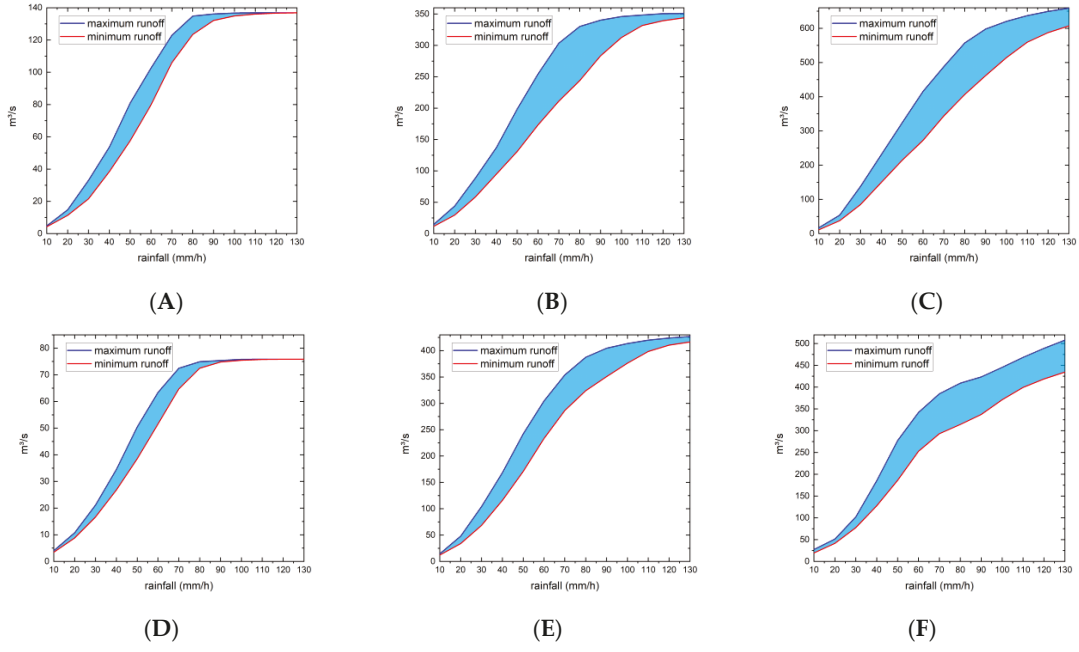



**Figure 7.** The peak runoff ranges with the same rainfall intensity (RI) value but different RMDs for 6 river sections. (**A**), section 50 of the Dasha River. (**B**), section 100 of the Dasha River. (**C**), section 150 of the Dasha River. (**D**), section 30 of the Buji River. (**E**), section 80 of the Buji River. (**F**), section 130 of the Buji River.

## 4. Conclusions

Our results show that rainfall movement direction (RMD) are very likely responsible for variations in flood risks in different sections of urban rivers, which should not be ignored in urban hydrologic management. Although the impact of RMD on peak runoff in rivers is always covered up by heavy rainfall intensity (RI), this impact is significant in the downstream reaches of urban river when combined with spatial features of rivers, such

as the river direction and bending. We propose that RMD should be involved in urban hydrologic models and predication, as it may yield substantial uncertainties in peak runoff in rivers. We provide an empirical evaluation to quantify the contribution of RMD to peak runoff in urban rivers, indicating that the river flood risk and discharge capacity should be evaluated based on more variations in rainfall such as RMD. Given the importance of the prevention and treatment of urban waterlogging with the accelerating process of urbanization, this empirical evaluation of peak runoff variations due to changing RMDs represents a contribution for the development of an Early Warning System (EWS) for the study area and provides critical information to inform policy and decision making.

**Author Contributions:** Y.L. (Yesen Liu) and M.L. collected and processed the data, Y.L. (Yesen Liu), K.L. and Y.L. (Yuanyuan Liu) proposed the model and analyzed the results, Y.H. and Y.L. (Yesen Liu) wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by National Natural Science Foundation of China (Grant NO. 41822104), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA19040402), the National Key Research and Development Program of China (Grant No. 2017YFB0503005), National Science and Technology Major Project of China's High Resolution Earth Observation system (21-Y20B01-9001-19/22), the Fundamental Research Funds of IWHR (WH0145B032024), the Fundamental Research Funds of IWHR(WH0145B032024).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

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

## References

1. Lai, Y.; Li, J.; Gu, X.; Chen, Y.D.; Kong, D.; Gan, T.Y.; Liu, M.; Li, Q.; Wu, G. Greater flood risks in response to slowdown of tropical cyclones over the coast of China. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 14751–14755. [CrossRef] [PubMed]
2. DeGaetano, A.T. Time-Dependent Changes in Extreme-Precipitation Return-Period Amounts in the Continental United States. *J. Appl. Meteorol. Clim.* **2009**, *48*, 2086–2099. [CrossRef]
3. Papalexiou, S.M.; Montanari, A. Global and Regional Increase of Precipitation Extremes under Global Warming. *Water Resour. Res.* **2019**. [CrossRef]
4. Arnell, N.W.; Gosling, S.N. The impacts of climate change on river flood risk at the global scale. *Clim. Chang.* **2016**, *134*, 387–401. [CrossRef]
5. Cristiano, E.; Deidda, R.; Viola, F. EHSMu: A New Ecohydrological Streamflow Model to Estimate Runoff in Urban Areas. *Water Resour. Manag.* **2020**, *34*, 4865–4879. [CrossRef]
6. Davenport, F.; Burke, M.; Diffenbaugh, N. Contribution of historical precipitation change to US flood damages. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2017524118. [CrossRef]
7. Xu, Z.; Chen, H.; Ren, M.; Cheng, T. Progress on disaster mechanism and risk assessment of urban flood/waterlogging disasters in China. *Shuikexue Jinzhan/Adv. Water Sci.* **2020**, *31*, 713–724. [CrossRef]
8. Veldhuis MC, T.; Zhou, Z.; Yang, L.; Liu, S.; Smith, J. The role of storm scale, position and movement in controlling urban flood response. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 417–436. [CrossRef]
9. Yan, M.; Chan, J.; Zhao, K. Impacts of Urbanization on the Precipitation Characteristics in Guangdong Province, China. *Adv. Atmos. Sci.* **2020**, *37*, 696–706. [CrossRef]
10. Marelle, L.; Myhre, G.; Steensen, B.; Hodnebrog, Ø.; Alterskjær, K.; Sillmann, J. Urbanization in megacities increases the frequency of extreme precipitation events far more than their intensity. *Environ. Res. Lett.* **2020**. [CrossRef]
11. Cristiano, E.; ten Veldhuis, M.C.; van de Giesen, N. Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas—A review. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 3859–3878. [CrossRef]
12. Hailegeorgis, T.T.; Alfredsen, K. Analyses of extreme precipitation and runoff events including uncertainties and reliability in design and management of urban water infrastructure. *J. Hydrol.* **2017**, *544*, 290–305. [CrossRef]
13. Hénonin, J.; Russo, B.; Roqueta, D.S.; Sanchezdiezma, R.; Domingo, N.S.; Domingo, S.; Thomsen, F.; Mark, O. *Urban Flood Real-time Forecasting and Modelling: A State-of-the-Art Review*; Publishr: Copenhagen, Denmark, 2010.
14. Xu, H.; Ma, C.; Xu, K.; Lian, J.; Long, Y. Staged optimization of urban drainage systems considering climate change and hydrological model uncertainty. *J. Hydrol.* **2020**, *587*, 124959. [CrossRef]
15. Pang, B.; Shi, S.; Zhao, G.; Shi, R.; Zhu, Z. Uncertainty Assessment of Urban Hydrological Modelling from a Multiple Objective Perspective. *Water* **2020**, *12*, 1393. [CrossRef]

16. Luca, D.L.D.; Versace, P. Diversity of Rainfall Thresholds for early warning of hydro-geological disasters. *Adv. Geosci.* **2017**, *44*, 53–60. [CrossRef]
17. Singh, A.; Zommers, Z. *Reducing Disaster: Early Warning Systems for Climate Change*; Springer: Washington, DC, USA, 2014; pp. 89–108. [CrossRef]
18. Cristiano, E.; Ten Veldhuis, M.-C.; Wright, D.; Smith, J.; van de Giesen, N. The Influence of Rainfall and Catchment Critical Scales on Urban Hydrological Response Sensitivity. *Water Resour. Res.* **2019**, *55*, 3375–3390. [CrossRef]
19. Rossman, L. Modeling Low Impact Development Alternatives with SWMM. *J. Water Manag. Modeling* **2010**. [CrossRef]
20. Lyu, H.; Ni, G.; Cao, X.; Ma, Y.; Tian, F. Effect of Temporal Resolution of Rainfall on Simulation of Urban Flood Processes. *Water* **2018**, *10*, 880. [CrossRef]
21. Malik, S.; Pal, S.; Sattar, A.; Singh, S.; Das, B.; Chakrabortty, R.; Mohammad, P. Trend of extreme rainfall events using suitable Global Circulation Model to combat the water logging condition in Kolkata Metropolitan Area. *Urban Clim.* **2020**, *32*. [CrossRef]
22. Mu, D.; Luo, P.; Lyu, J.; Zhou, M.; Huo, A.; Duan, W.; Nover, D.; He, B.; Zhao, X. Impact of temporal rainfall patterns on flash floods in Hue City, Vietnam. *J. Flood Risk Manag.* **2020**, *14*, e12668. [CrossRef]
23. Zoccatelli, D.; Borga, M.; Viglione, A.; Chirico, G.B.; Blöschl, G. Spatial moments of catchment rainfall: Rainfall spatial organisation, basin morphology, and flood response. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 3767–3783. [CrossRef]
24. Smith, J.A.; Bradley, A.A.; Baeck, M.L. The SpaceTime Structure of Extreme Storm Rainfall in the Southern Plains. *J. Appl. Meteorol.* **1994**, *33*, 1402–1417. [CrossRef]
25. Ogden, F.L.; Richardson, J.R.; Julien, P.Y. Similarity in Catchment Response: 2. Moving Rainstorms. *Water Resour. Res.* **1995**, *31*, 1543–1547. [CrossRef]
26. Singh, V.P. Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. *Hydrol. Process.* **1997**, *11*, 1649–1669. [CrossRef]
27. Emmanuel, I.; Andrieu, H.; Leblois, E.; Flahaut, B. Temporal and spatial variability of rainfall at the urban hydrological scale. *J. Hydrol.* **2012**, *430–431*, 162–172. [CrossRef]
28. Nikolopoulos, E.I.; Borga, M.; Zoccatelli, D.; Anagnostou, E.N. Catchment-scale storm velocity: Quantification, scale dependence and effect on flood response. *Int. Assoc. Sci. Hydrol. Bull.* **2014**, *59*, 1363–1376. [CrossRef]
29. Emmanuel, I.; Andrieu, H.; Leblois, E.; Janey, N.; Payrastre, O. Influence of rainfall spatial variability on rainfall-runoff modelling: Benefit of a simulation approach? *J. Hydrol.* **2015**, *531*, 337–348. [CrossRef]
30. Zoccatelli, D.; Borga, M.; Zanon, F.; Antonescu, B.A.; Stancalie, G. Which rainfall spatial information for flash flood response modelling? A numerical investigation based on data from the Carpathian range, Romania. *J. Hydrol.* **2010**, *394*, 148–161
31. Liu, Y.Y.; Li, L.; Liu, Y.S.; Chan, P.W.; Zhang, W.H. Dynamic spatial-temporal precipitation distribution models for short-duration rainstorms in Shenzhen, China based on machine learning. *Atmos. Res.* **2020**, *237*, 104861. [CrossRef]
32. Hallegatte, S.; Green, C.; Nicholls, R.; Corfee-Morlot, J. Future flood losses in major coastal cities. *Nat. Clim. Chang.* **2013**, *3*, 802–806. [CrossRef]
33. Zhou, L.; Wu, X.; Ji, Z.; Gao, G. Characteristic analysis of rainstorm-induced catastrophe and the countermeasures of flood hazard mitigation about Shenzhen city. *Geomatics* **2017**, *8*, 1886–1897. [CrossRef]
34. Zang, W.; Shu, L.; Huang, S.; Li, J.; Fu, Y.; Sun, Y.; Zheng, J. Impact of urbanization on hydrological processes under different precipitation scenarios. *Nat. Hazards* **2019**, *99*, 1233–1257. [CrossRef]
35. Haijun, Y.U.; Jianming, M.A. Application of IFMS Urban software in urban flood risk mapping. *China Flood Drought Manag.* **2018**, *28*, 13–17.
36. Kim, I.H.; Feng, C.C.; Wang, Y.C. A simplified linear feature matching method using decision tree analysis, weighted linear directional mean, and topological relationships. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 1042–1060. [CrossRef]
37. Mitchell, A. *The ESRI Guide to GIS Analysis*; ESRI Press: Redlands, CA, USA, 1999.
38. Wheater, H.S.; Chandler, R.E.; Onof, C.J.; Isham, V.S.; Bellone, E.; Yang, C.; Lekkas, D.; Lourmas, G.; Segond, M.L. Spatial-temporal rainfall modelling for flood risk estimation. *Stoch. Environ. Res. Risk Assess.* **2005**, *19*, 403–416. [CrossRef]
39. Aronica, G.; Cannarozzo, M. Studying the hydrological response of urban catchments using a semi-distributed linear non-linear model. *J. Hydrol.* **2000**, *238*, 35–43. [CrossRef]

water

MDPI

*Article*

# Improvement in Ridge Coefficient Optimization Criterion for Ridge Estimation-Based Dynamic System Response Curve Method in Flood Forecasting

**Kexin Liu [1,2,*], Weimin Bao [1], Yufeng Hu [2], Yiqun Sun [1], Dongjing Li [2], Kuang Li [2] and Lili Liang [3]**

[1] Department of Hydrology and Water Resources, University of Hohai, Nanjing 210000, China; wmbao163@163.com (W.B.); yiqun.sun@Hotmail.com (Y.S.)
[2] China Institute of Water Resources and Hydropower Research, Beijing 100038, China; hyf6969@163.com (Y.H.); tkin.l@outlook.com (D.L.); likuang0000@iwhr.com (K.L.)
[3] Institute of Science and Technology, China Three Gorges Corporation, Beijing 100038, China; liangli0921@163.com
* Correspondence: liukexin0412@163.com

**Abstract:** The ridge estimation-based dynamic system response curve (DSRC-R) method, which is an improvement of the dynamic system response curve (DSRC) method via the ridge estimation method, has illustrated its good robustness. However, the optimization criterion for the ridge coefficient in the DSRC-R method still needs further study. In view of this, a new optimization criterion called the balance and random degree criterion considering the sum of squares of flow errors (BSR) is proposed in this paper according to the properties of model-simulated residuals. In this criterion, two indexes, namely, the random degree of simulated residuals and the balance degree of simulated residuals, are introduced to describe the independence and the zero mean property of simulated residuals, respectively. Therefore, the BSR criterion is constructed by combining the sum of squares of flow errors with the two indexes. The BSR criterion, L-curve criterion and the minimum sum of squares of flow errors (MSSFE) criterion are tested on both synthetic cases and real-data cases. The results show that the BSR criterion is better than the L-curve criterion in minimizing the sum of squares of flow residuals and increasing the ridge coefficient optimization speed. Moreover, the BSR criterion has an advantage over the MSSFE criterion in making the estimated rainfall error more stable.

**Keywords:** flood forecasting; error correction; residual property; ridge coefficient criterion

**Citation:** Liu, K.; Bao, W.; Hu, Y.; Sun, Y.; Li, D.; Li, K.; Liang, L. Improvement in Ridge Coefficient Optimization Criterion for Ridge Estimation-Based Dynamic System Response Curve Method in Flood Forecasting. *Water* **2021**, *13*, 3483. https://doi.org/10.3390/w13243483

Academic Editor: Renato Morbidelli

Received: 5 November 2021
Accepted: 1 December 2021
Published: 7 December 2021



## 1. Introduction

Flood forecasting, an important non-structural measure, plays an important role in regional flood control, flood warning, risk decision making, etc. [1–3]. The hydrological model simplifies and conceptualizes the flood process with a set of equations aiming at obtaining the outlet flow. However, not all problems can be solved with such a model as the flood forecasting accuracy is often hampered and influenced by many error factors existing in the hydrologic system, including the errors in the model inputs, the errors in the model initial condition, the errors in the model simplification and the errors in the model parameters. Therefore, many scholars have devoted themselves to the research of error correction methods. For example, the autoregressive (AR) model estimates the flow error existing in a certain forecasting period by using the correlation of error series, and it was later developed into improved methods such as the recursive autoregressive model and the forgetting factor recursive autoregressive model [4–6]; Kalman filtering (KF) technology is widely used to update hydrological element time series in flood forecasting, and many improved types have been gradually formed, including the extended Kalman filter (EKF) [7] and the ensemble Kalman filter (EnKF) [8]. Data assimilation technology has also shown satisfying results in improving the prediction accuracy of models, including dynamic identifiability analysis (DYNIA) [9], and the Bayesian recursive estimation technique

(BaRE) [10]. Robust theory [11] and the comprehensive correction method [12] are also used in flood forecasting error correction. However, error correction technology still needs to be improved; for example, the autoregressive model assumes that the error series has a linear correlation, but its performance near the flood peak is often not satisfying [12]. Additionally, some studies have also shown that the EnKF does not perform very well when the error structure is far away from a Gaussian distribution [13]. Scholars are still trying to solve the above problems.

A new error correction method called the dynamic system response curve (DSRC) method has been proposed by Bao et al. [14]. This method constructs a feedback model which is conceptualized on updating the hydrologic element series by tracing back to the source of the error. With the help of the first-order Taylor linearization to approximate the hydrologic model, error correction is achieved by solving the corresponding equations using the least square method. The DSRC method was initially applied to correct single hydrological elements, including runoff [15], rainfall [16] and model state variables [17]. Then, it was used to correct several hydrological elements comprehensively [18]. However, some studies [17–20] found that the correction results are not always stable, reflected in the excessive correction of hydrological element series and none-smooth simulated flow hydrographs. To solve the above problems, the DSRC-R method was developed by Si et al. [19] from the point of the regularization known as ridge estimation, and this method has improved the stability of correction results to some degree. Nevertheless, the selection criterion of the ridge coefficient still needs further study and improvement. Previous studies [17–19] often chose the ridge coefficient based on the L-curve criterion [21]. However, it has been found in practice that two aspects still need attention. One is that the L-curve seems insufficient to reflect the properties of model-simulated residual errors including the independence and the zero mean property of simulated residuals, which hinders the performance of DSRC-R in some cases; the other is that the application of the L-curve criterion takes a long time, which is not conducive to the real-time performance of flood forecasting. The L-curve criterion involves derivative calculation, and the difference method makes the operation efficiency lower.

Therefore, in this paper, we analyze the L-curve criterion and introduce the concepts of the random degree of simulated residuals and the balance degree of simulated residuals to describe the properties of the model-simulated residuals, and then a new criterion called the balance and random degree criterion considering the sum of squares of flow errors (BSR) is proposed. The new criterion takes the independence and the zero mean property of simulated residuals into account, which is conducive to obtaining a ridge coefficient in line with the statistical characteristics of residuals. The new criterion does not involve derivative calculation; thus, it can greatly shorten the search time of the ridge coefficient in optimization, improve the operational efficiency and enhance the real-time flood forecasting performance.

## 2. Methodology

### *2.1. DSRC Method*

The main idea behind the DSRC method is that it firstly retrieves the rainfall errors from the outlet flow errors, then updates the rainfall series and finally reruns the model with the updated rainfall series. In this method, given a hydrological model $Q = Q(P)$ that generates outlet flow $Q$ as a function of rainfall $P$, the variation process of the outlet flow change by the unit perturbation in rainfall is called the system response curve. Based on this, multi-time system response curves form the system response matrix $S$, and this matrix sets up the relation between rainfall errors $\Delta P$ and flow errors $\Delta Q$; then, the estimation of rainfall errors $\hat{\Delta P}_{LS}$ can be computed via the least square method. In this round, the input rainfall series is updated with $\hat{\Delta P}_{LS}$, and then the model is rerun with the updated rainfall series to correct the forecasting results. In this study, the DSRC method was combined with the Xinanjiang (XAJ) model, a hydrological model constructed by Professor Zhao Renjun of Hohai University which is widely used in flood forecasting in humid areas of

China [22]. Additionally, here, we mainly talk about the calculation process of the DSRC method shown in Figure 1 rather than the XAJ model, as this paper will introduce it in Section 3. Additionally, the theoretical derivation of the DSRC method can be gained from [16].

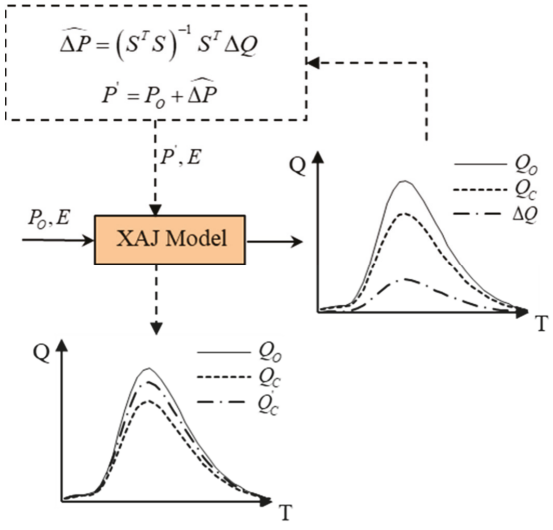



**Figure 1.** Schematic diagram of the DSRC method. $Q_C$, $Q_O$ and $Q_C'$ are model-simulated flow, observed flow and updated model-simulated flow, respectively; $\Delta Q$ is the model-simulated deviation series of the outlet flow; $\hat{\Delta P}$ is the estimated rainfall error series; $P_O$ is the initial rainfall series; $P'$ is the updated rainfall series; and $E$ is pan evaporation.

According to [16], the rainfall error estimation $\hat{\Delta P}_{LS}$ is expressed as the following Equation (1):

$$\hat{\Delta P}_{LS} = \left(S^T S\right)^{-1} S^T \Delta Q \tag{1}$$

where $\Delta Q$ is the model-simulated deviation series of the outlet flow, $\Delta Q = Q_O - Q_C$; $Q_O = [Q_{O,1}, Q_{O,2}, \cdots Q_{O,M}]^T$ is the observed flow series; and $Q_C = [Q_{C,1}, Q_{C,2}, \cdots Q_{C,M}]^T$ is the simulated flow series computed from the observed rainfall series.

In Equation (1), $S$ is the system response matrix defined as

$$S = \begin{bmatrix} \frac{\partial Q_1(P)}{\partial p_1} & \frac{\partial Q_1(P)}{\partial p_2} & \cdots & \frac{\partial Q_1(P)}{\partial p_N} \\ \frac{\partial Q_2(P)}{\partial p_1} & \frac{\partial Q_2(P)}{\partial p_2} & \cdots & \frac{\partial Q_2(P)}{\partial p_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial Q_M(P)}{\partial p_1} & \frac{\partial Q_M(P)}{\partial p_2} & \cdots & \frac{\partial Q_M(P)}{\partial p_N} \end{bmatrix} \tag{2}$$

where $p_1, p_2 \cdots p_n$ are the initial rainfall values; and the indices $M$ and $N$ represent the lengths of observed flow and rainfall, respectively ($M \geq N$). $\frac{\partial Q_i(P)}{\partial p_j}$ represents the influence of the $j$-th rainfall on the $i$-th outlet flow. When $i < j$, it is obvious that $\frac{\partial Q_i(P)}{\partial p_j} = 0$ because rainfall does not affect the outlet flow that occurs before it. $\frac{\partial Q_i(P)}{\partial p_j}$ is generally obtained by the difference, which is $\frac{\partial Q_i(P)}{\partial p_j} = \frac{Q(p_1,\ldots,p_j+\Delta p,\ldots,p_N) - Q(p_1,\ldots,p_j,\ldots,p_N)}{\Delta p}$.

Accordingly, the updated rainfall series $P'$ can be expressed as Equation (3):

$$P' = P_O + \hat{\Delta P} \tag{3}$$

where $P'$ is the updated rainfall series; $P_O$ is the initial rainfall series; and $\hat{\Delta P}$ is the rainfall error estimation series. In the DSRC method, $\hat{\Delta P}$ is replaced by $\hat{\Delta P}_{LS}$ mentioned above.

In order to improve the forecasting accuracy, the updated rainfall series $P'$ is introduced into the hydrological model for recalculation, and then the updated model-simulated flow series $Q'_C$ is obtained by Equation (4):

$$Q'_C = Q(P') \tag{4}$$

where $Q'_C = [Q'_{C,1}, Q'_{C,2}, \cdots, Q'_{C,N}]^T$ is the updated model-simulated flow series.

### 2.2. DSRC-R Method

Correction results from the DSRC method are sometimes unstable. Relevant studies [19,20] have pointed out that the DSRC method is prone to be ill-conditioned, which generates unstable results when this method is applied to small basins, or when the length of flow information is short. Therefore, Si et al. [19] combined the DSRC method with the ridge estimation method and proposed the DSRC-R method, which is more robust than the DSRC method. In previous studies, the ridge coefficient $\beta$ was often selected via the L-curve criterion [17–19]; however, the correction results were not always stable. Thus, the criterion for obtaining the appropriate ridge coefficient $\beta$ in the DSRC-R method needs further study. Here, we directly provide the formula of rainfall error estimation $\hat{\Delta P}_{RE}$ as the following Equation (5). For more details about the derivation process of DSRC-R, please refer to [19].

$$\hat{\Delta P}_{RE} = \left(S^T S + \beta I\right)^{-1} S^T \Delta Q \tag{5}$$

where $\beta$ is the ridge coefficient; $I$ is the identity matrix; and $\hat{\Delta P}_{RE}$ is the rainfall error estimation series of the DSRC-R method.

$\hat{\Delta P}_{RE}$ can be introduced into Equation (3) to update the rainfall series, and then the model can be rerun with $P'$ to correct the forecasted flow. The flow chart of the DSRC-R method is shown in Figure 2.

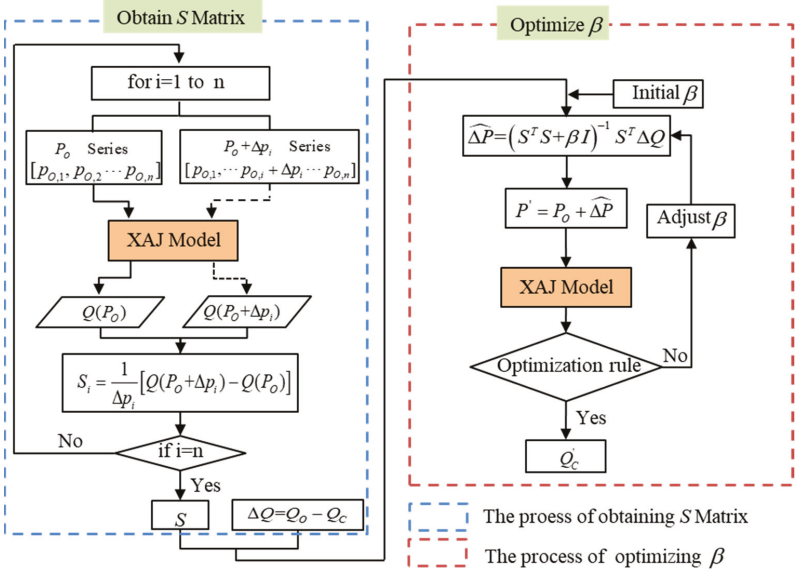



**Figure 2.** Flow chart of the DSRC-R method. $S_i$ is the system response curve of the $i$-th rainfall, that is, the $i$-th column of matrix S.

### 2.3. L-Curve Criterion

The ridge coefficient $\beta$ is significantly important in the DSRC-R method. In some previous studies [17–19], the L-curve criterion was used to determine the value of $\beta$. According to [23], the main idea behind the L-curve criterion in selecting the appropriate ridge estimation coefficient $\beta$ in the DSRC-R method can be summarized as: the balance of $\lg\|Q(P_O + \hat{\Delta P}_{RE}) - Q_O\|$ and $\log\|\hat{\Delta P}_{RE}\|$, and the appropriate ridge coefficient lies at the corner of the curve, usually with the largest curvature. That is, the technique to find the appropriate ridge coefficient can be expressed as Equation (6).

$$\max_{\beta}\left(\frac{|f'g'' - f''g'|}{\left[(f')^2 + (g')^2\right]^{3/2}}\right) \tag{6}$$

where $f(\beta) = \log\|\hat{\Delta P}_{RE}\|_2^2$; $g(\beta) = \lg\|Q(P_O + \hat{\Delta P}_{RE}) - Q_O\|_2^2$; and $\|\bullet\|_2^2$ is the modular square.

It can be proved (see Appendices A and B) that $f(\beta)$ and $g(\beta)$ can be expressed as Equations (7) and (8).

$$f(\beta) = \log[\sum_{i=1}^{n}(\frac{k_i}{\lambda_i + \beta})^2] \tag{7}$$

$$g(\beta) = \lg\|Q(P_O + \sum_{i=1}^{n}\frac{k_i}{\lambda_i + \beta}v_i) - Q_O\|_2^2 \tag{8}$$

where $\lambda_i (i = 1, \cdots, N)$ is the eigenvalue of the matrix $S^TS$; $v_i (i = 1, \cdots, N)$, orthogonal to each other, is the unit eigenvector corresponding to $\lambda_i (i = 1, \cdots, N)$; and $k_i (i = 1, \cdots, n)$ is a group of coefficients that enable $S^T\Delta Q$ to be linearly expressed by $v_i (i = 1, \cdots, N)$, that is, $S^T\Delta Q = \sum_{i=1}^{n} k_i v_i$.

The L-curve criterion has a good effect on selecting the ridge coefficient in the DSRC-R method, but there are still some problems that are worthy of attention. First, the result is sometimes unsatisfactory. The reason is that the L-curve criterion seems to insufficiently reflect the properties of the model-simulated residuals, although it pays attention to the balance of $\lg\|Q(P_O + \hat{\Delta P}_{RE}) - Q_O\|$ and $\log\|\hat{\Delta P}_{RE}\|$. Second, the optimization of the ridge coefficient consumes too much time. The L-curve criterion involves first-order and second-order derivatives, as shown in Equation (6). Additionally, the explicit expression of $g(\beta)$ cannot be obtained at present, meaning its derivative can only be obtained by a difference method; thus, it will take a long time and is not conducive to the real-time performance of flood forecasting.

### 2.4. New Optimization Criterion (BSR)

In previous studies [17–19], the L-curve criterion was generally adopted to find the suitable ridge estimation coefficient $\beta$ in the DSRC-R method. Nevertheless, the L-curve criterion has some shortcomings such as insufficient consideration of model-simulated residuals, imperfect utilization of information and huge consumption of time. Therefore, this study takes the properties of the model-simulated residuals into consideration and then explores a new optimization criterion which is more suitable for the DSRC-R method.

For any model, it is always expected that the simulated residual series satisfies the zero mean property and non-correlative statistical property; in other words, let the mean of the residual series and correlation coefficient be as small as possible. This shows us that the criterion for determining the ridge coefficient should consider Equations (9) and (10). We use Equation (9) to express the zero mean property of the residuals, and this indicator is called the balance degree of simulated residuals (BDSR). We use Equation (10) to express the correlation of the residual series, and its reciprocal is called the random degree of simulated residuals (RDSR), which is shown in Equation (11). Additionally, RDSR indicates

the independence of the residuals, and we generally expect a smaller value of the BDSR indicator and a bigger value of the RDSR indicator.

$$BDSR = abs\left[\sum_{i=1}^{m}\left(Q_{O,i} - Q'_{C,i}\right)\right] \tag{9}$$

$$r_e = abs[\frac{\sum_{i=1}^{m-1}(\Delta Q'_i - \overline{\Delta Q_X})(\Delta Q'_{i+1} - \overline{\Delta Q_Y})}{\sqrt{\sum_{i=1}^{m-1}(\Delta Q'_i - \overline{\Delta Q_X})^2 \sum_{i=1}^{m-1}(\Delta Q'_{i+1} - \overline{\Delta Q_Y})^2}}] \tag{10}$$

$$RDSR = \frac{1}{r_e} \tag{11}$$

where *abs* represents the absolute value sign; $r_e$ is the absolute value of the correlation coefficient of adjacent residuals;$\Delta Q'_i = Q_{O,i} - Q'_{C,i}$; $\overline{\Delta Q_X} = \frac{1}{m-1}\sum_{i=1}^{m-1}\Delta Q'_i$; and $\overline{\Delta Q_Y} = \frac{1}{m-1}\sum_{i=1}^{m-1}\Delta Q'_{i+1}$.

Traditional methods often take the minimum sum of squares of flow errors (MSSFE) as the objective function for calibration parameters, but the effective information contained in this method is not sufficient to obtain the ridge coefficient. The reason is that the derivation process of the DSRC-R method utilizes the least square method, and when the system is linear, the least sum of squares of flow errors equals the least square method; thus, the value of $\beta$ should be zero in this circumstance. Although the DSRC-R method belongs to non-linear system inversion methods, the value of $\beta$ still has a decreasing tendency, and this is not conducive to the stability of the method. Therefore, we need to further excavate more useful information in the simulated errors.

In this paper, we take the independence and the zero mean property of simulated residuals into account, combine these two points with the sum of squares of flow errors (SSFE) and lastly explore a new criterion called the balance and random degree criterion considering the sum of squares of flow errors (BSR criterion). Generally, we hope to find a large value of RDSR which is more consistent with the property of the residuals and is conducive to avoiding system errors; moreover, we hope to find a small value of BDSR which can satisfy the zero mean property of residuals and can decrease the flood volume errors. Based on this, considering the SSFE indicator, we propose the BSR criterion, the mathematical form of which is provided in Equation (12). Overall, the new BSR criterion considers the traditional sum of squares of residuals; furthermore, it is possible to find the ridge coefficient which satisfies the properties of flow residuals.

$$\min_{\beta}\frac{(BDSR+1)SSFE}{RDSR} \tag{12}$$

where "+1" is used to avoid the value of BDSR being zero; and $SSFE = \sum_{i=1}^{m}\left(Q_{O,i} - Q'_{C,i}\right)^2$.

Since the BSR criterion pays more attention to the independence and the zero mean property of simulated residuals than the L-curve criterion, it is more likely to select a ridge coefficient $\beta$ which satisfies the properties of residuals, and thus a better performance can be achieved with the DSRC-R method; moreover, the BSR criterion does not involve derivation calculation, meaning it can improve operational efficiency and save much more time.

In real-time flood forecasting, in order to quickly obtain an appropriate $\beta$, we need to utilize an automatic optimization method. Additionally, in this paper, we adopt the particle swarm optimization algorithm, which was first proposed by Kennedy and Eberhart and constructed on the concept of mimicking the social behavior of birds [24–27]. This

algorithm has been widely used in many types of optimization problems. For more details about particle swarm algorithms, please refer to [24].

### *2.5. The Entire Research Process*

In order to make the whole research process more clear, we created a flow chart, as shown in Figure 3. The figure shows the entire research process including the proposal of the BSR criterion, the research of a synthetic case and a real case and the comparison of three criteria (BSR, L-curve and MSSFE).

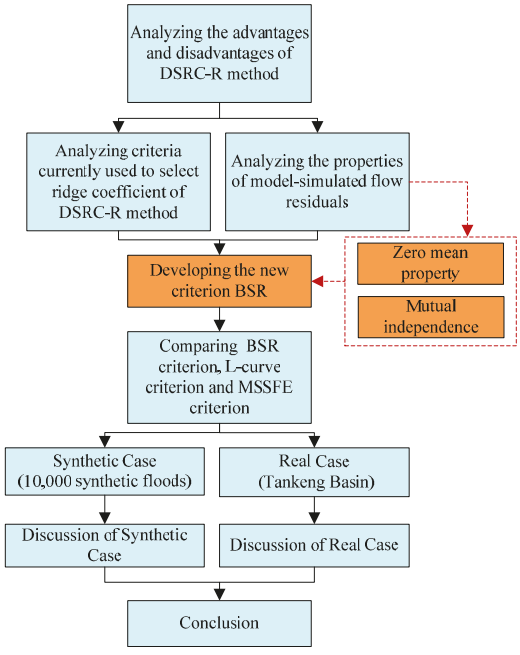



**Figure 3.** The research flow chart showing the entire research process.

## 3. Case Study

This research projecting synthetic and real-data studies aimed at comparing the performance of the DSRC-R method under three criteria which include the BSR criterion, the L-curve criterion and the MSSFE criterion.

### *3.1. Model Description*

The selected hydrological model in this research is the Xinanjiang (XAJ) model by Professor Zhao Renjun of Hohai University [22], which is one of the most widely used conceptual hydrological models in China. The XAJ model, including the inputs of observed precipitation as well as pan evaporation, and the outputs of forecasted flow as well as evaporation, can be used in different spatial and temporal scales and be divided into four layers: the first layer utilizing the three-layer evapotranspiration (TLE) model to realize basin evaporation; the second layer utilizing the saturated runoff production (SRP) model to realize the basin runoff production; the third layer utilizing the free water storage model to realize runoff separation; and the fourth layer utilizing the linear reservoir method and the Muskingum method to realize the basin flow concentration. When applying the XAJ model, firstly, divide the basin into several sub-basins and then compute the runoff and outlet flow in every sub-basin; lastly, gather the flow of each sub-basin at the outlet of the basin. For

more details about the XAJ model, please refer to [22]. The structure of the XAJ model is shown in Figure 4, and the meaning of parameters in each layer is shown in Table 1.

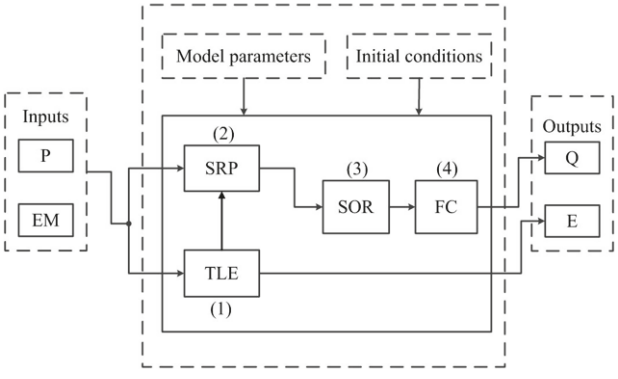



**Figure 4.** Schematic diagram of the Xinanjiang (XAJ) model. TLE represents the three-layer evapotranspiration model; SRP represents the saturated runoff production layer; SOR represents the runoff separation layer; and FC represents the flow concentration layer.

**Table 1.** The parameters of the XAJ model.

| Layer | Function | Parameter | Meaning |
|---|---|---|---|
| First layer | Evaporation | K | Ratio of potential evapotranspiration to pan evaporation |
| | | WUM | Areal mean tension water capacity of the upper layer |
| | | WLM | Areal mean tension water capacity of the lower layer |
| | | WDM | Areal mean tension water capacity of the deeper layer |
| | | C | Coefficient of deep evapotranspiration |
| Second layer | Runoff production | IM | Ratio of impervious area |
| | | WM | Areal mean tension water capacity |
| | | B | Exponent of the tension water capacity distribution curve |
| Third layer | Runoff separation | SM | Areal mean free water capacity of the surface soil layer |
| | | EX | Exponent of the free water capacity curve |
| | | KI | Outflow coefficients of the free water storage to interflow |
| | | KG | Outflow coefficients of the free water storage to groundwater |
| Fourth layer | Flow Concentration | CS | Recession constant of the surface water storage |
| | | CI | Recession constant of the interflow storage |
| | | CG | Recession constant of the groundwater storage |
| | | KE | Storage time constant |
| | | XE | Weight factor |

### *3.2. Synthetic Case*

The synthetic case was to design a typical artificial basin whose catchment area, station distribution, model parameters, hydrological features of the basin condition, "observed" precipitation, evaporation, "observed" outlet flow, error factors and any other information about the basin are all known in order to compare different schemes expediently. The synthetic case included 10,000 synthetic precipitations and corresponding floods in order to compare the performance of the DSRC-R method under three criteria (the BSR criterion, the L-curve criterion and the MSSFE criterion).

A major point of the synthetic case was to obtain the "observed" flow. Here, we referred to [16] and utilized Equation (13) to obtain it. By using different $P_O$ and $\Delta P$, we can obtain a different $Q_O$.

$$Q_O = Q(P_O + \Delta P) + e \tag{13}$$

where $Q_O$ is the "observed" flow series; $P_O$ is the initial precipitation series; $\Delta P$ is the given error series, and each value in the series $\Delta P$ does not exceed 30% of the corresponding value in $P_O$; and $e$ is Gaussian white noise which cannot exceed 5% of the initial value in this study.

This synthetic case assumed that the basin area is 1000 km$^2$ and there are 8 precipitation stations in the basin. The value of parameters from each layer is shown in the following Table 2.

**Table 2.** The parameters of the XAJ model in the synthetic basin.

| **Parameter** | **K** | **WM** | **WUM** | **WLM** | **WDM** | **IM** | **B** | **C** | **SM** |
|---|---|---|---|---|---|---|---|---|---|
| Value | 1.1 | 150 | 20 | 80 | 50 | 0.01 | 0.3 | 0.16 | 10 |
| **Parameter** | **EX** | **KI** | **KG** | **CS** | **CI** | **CG** | **KE** | **XE** | |
| Value | 1.5 | 0.35 | 0.35 | 0.78 | 0.865 | 0.995 | 1.50 | 0.380 | |

#### 3.2.1. Data

A major point of the synthetic case was to generate different initial series of precipitations $P_O$. In order to increase the diversity of $P_O$, we applied the following method. Firstly, we chose 55 typical areal precipitation processes from a real basin, then transformed the position of the rainfall peak in each precipitation and eventually formed 500 synthetic typical precipitation processes. When generating synthetic rainfall, we selected a synthetic typical precipitation, randomly adjusted each rainfall period ranging less than 30% of the typical rainfall and then obtained the proportion of each time interval of the rainfall series; then, we randomly generated the total rainfall and allocated it to each time interval according to the above proportion, thus forming one synthetic precipitation, that is, one initial series of precipitation $P_O$. Then, we introduced $P_O$ and the given $\Delta P$ into Equation (13) to obtain the "observed" flow series. In this case, we constructed 10,000 synthetic precipitations and corresponding floods, and the total rainfall of each flood ranged from 10 to 200 mm.

#### 3.2.2. Statistical Indicators

In the synthetic case, what we consider most is the performance of the DSRC-R method under different criteria rather than the contrast of the results between the DSRC and DSRC-R methods, as was accomplished in [19]. The criteria include the BSR criterion, the L-curve criterion and the MSSFE criterion. The relevant statistical indicators include the relative error of flood peak (RPF), relative error of runoff depth (RRD), Nash–Sutcliffe effiency coefficent (NSE), time needed to update a flood (TU) and root mean square error (RMSE). RMSE can be utilized to evaluate the robustness of the DSRC-R method under different criteria, and this index was applied in [17]. The smaller the value of the RMSE indicator, the more robust the DSRC-R method will be.

The statistical indicators between synthetic cases and real-data cases are different. In synthetic cases, RMSE is one of the indicators; however, it is not covered in real-data cases

as there is no way to obtain the "true" precipitation. The mathematical definitions of each statistical indicator are expressed as follows:

Relative error of flood peak (RPF):

$$\mathrm{RPF} = (Q_{OP} - Q_{CP})/Q_{OP} \times 100\% \tag{14}$$

Relative error of runoff depth (RRD):

$$\mathrm{RRD} = (R_O - R_C)/R_O \times 100\% \tag{15}$$

Nash–Sutcliffe effiency coefficent (NSE):

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N} (Q_{C,i} - Q_{O,i})^2}{\sum_{i=1}^{N} (Q_{O,i} - \overline{Q}_O)^2} \tag{16}$$

Root mean square error (RMSE):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (p_{T,i} - p_i')^2} \tag{17}$$

where $Q_{CP}$ is the forecasted value of the flood peak; $Q_{OP}$ is the observed value of the flood peak; $R_C$ is the forecasted depth of runoff; $R_O$ is the observed depth of runoff; $\overline{Q}_O$ is the average value of flow; $Q_{C,i}$ and $Q_{O,i}$ are the forecasted flow and observed flow in the $i$-th time interval; and $p_{T,i}$ and $p_i'$ are the original precipitation and the updated precipitation in the $i$-th time interval. In the synthetic case, $p_{T,i} = p_{O,i} + \Delta p_i$, and $p_i' = p_{O,i} + \hat{\Delta p}_i$. The larger the NSE (NSE $\leq 1$), the higher the forecasting accuracy, the smaller the RMSE and the more robust the DSRC-R method will be.

3.2.3. Computational Process of DSRC Method and DSRC-R Method

The mechanism of the DSRC method is that it firstly utilizes the error information of the outlet flow to invert and estimate the rainfall error, then updates the original rainfall series and lastly reruns the model with the updated rainfall series to correct the forecasting result. The specific steps of the DSRC method are as follows:

1. Add the additional precipitation in the $i$-th time interval $\Delta p_i$ to the precipitation in the $i$-th time interval $p_{O,i}$ while keeping the precipitation in the $j$-th time interval $(j \neq i) p_{O,j}$ unchanged; then, obtain the new precipitation series $P_O + \Delta p_i$.
2. Introduce the original precipitation series $P_O$ and new precipitation series $P_O + \Delta p_i$ into the model and obtain the series $Q(P_O)$ and $Q(P_O + \Delta p_i)$, respectively. Then, $S_i$ is obtained by the equation $S_i = [Q(P_O + \Delta p_i) - Q(P_O)]/\Delta p_i$, where $S_i$ is the dynamic system response curve of the $i$-th rainfall, that is, the $i$-th column of matrix $S$.
3. Cycle Steps 1 and 2 $n$ times and obtain the precipitation dynamic system response matrix $S$.
4. Add the estimated precipitation error series $\hat{\Delta P}_{LS}$ to the original precipitation series $P_O$ and obtain the updated precipitation series $P'$.
5. Introduce the updated precipitation series $P'$ into the model in order to obtain the updated forecasted flow $Q_C'$.

According to relevant research [19,20], when the flow data are insufficient, the DSRC method will tend to be unstable, characterized by wide fluctuations in the precipitation error estimated series and potentially oscillation, which will influence the flow correction effect. Therefore, a more robust method, DSRC-R, is proposed through combination with the ridge estimation method. Although the DSRC-R method ensures the stability of the error estimate, the ridge coefficient varies with floods. Thus, the ridge coefficient selection

should be in accordance with the instant data. The steps of ridge coefficient optimization are as follows:

1. Initialize the ridge coefficient $\beta$;
2. Obtain the precipitation error estimation series $\hat{\Delta P}_{RE} = \left(S^T S + \beta I\right)^{-1} S^T \Delta Q$. Add $\hat{\Delta P}_{RE}$ to $P_O$ in order to obtain $P'$;
3. Rerun the model with $P'$ and obtain the updated flow process $Q'_C$;
4. Judge whether the results meet the criteria (the criteria adopted in this essay include the BSR criterion, the L-curve criterion and the MSSFE criterion). If yes, turn to Step 6; if no, go back to Step 5;
5. Adjust the ridge coefficient $\beta$ according to the optimization algorithm (this essay applied the particle swarm optimization algorithm) and then turn to Step 2;
6. Finish the optimization process and acquire the optimal ridge coefficient $\beta$.

In order to evaluate the correction effect of the three criteria, the synthetic case not only includes some flow indicators (RPF, RRD, NSE) but also includes some precipitation indicators (RMSE). Additionally, RMSE is proposed to quantitatively describe the robustness of the method and the oscillation phenomena. The smaller the RMSE, the more robust the precipitation error estimation will be. RMSE has been applied in previous studies and has been successful. However, in the actual case, due to the inability to acquire the "true" value of rainfall, the performance of the three criteria cannot be evaluated by the RMSE indicator.

#### 3.2.4. Results and Discussion

The performance of the DSRC-R method on 10,000 synthetic floods under three criteria is shown in Table 3, and one typical synthetic flood is shown in Figure 5. As can be seen in Table 3, although the three criteria (BSR, L-curve and MSSFE) have a certain effect, some differences still exist. In terms of the indicators RPF and RRD, the BSR criterion has the best performance (1.90% and 1.01%). The MSSFE criterion takes second place in this regard (1.97% and 1.27%), and the L-curve criterion has the worst performance (2.19% and 1.30%). In terms of the indicator NSE, the BSR criterion and MSSFE criterion have the same result, with a value of 0.999, which outnumbers that of the L-curve criterion, with 0.001. In terms of operational efficiency, the L-curve criterion consumes much more time, where the average TU of a flood is 12.21 s. The value of the ridge coefficient $\beta$ under the three criteria has a big difference, and $\beta$ tends to be smaller (average value of 64.26) when it applies the MSSFE criterion. In terms of the indicator RMSE, the value under the BSR criterion is 0.759, which is significantly less than the value (1.040) under the MSSFE criterion, indicating that the BSR criterion is more conducive to improving the robustness of the DSRC-R method than the MSSFE criterion.

**Table 3.** The results of the synthetic case.

| Items [1] | RPF | RRD | NSE | RDSR | BDSR | TU | $\beta$ | RMSE |
|---|---|---|---|---|---|---|---|---|
| Before correction | 9.38 | 5.12 | 0.985 | —— | —— | —— | —— | —— |
| L-curve | 2.19 | 1.30 | 0.998 | 5.57 | 80.8 | 12.21 | 985.11 | 0.941 |
| MSSFE | 1.97 | 1.27 | 0.999 | 9.31 | 74.8 | 3.99 | 64.26 | 1.040 |
| BSR | 1.90 | 1.01 | 0.999 | 12.88 | 66.9 | 4.11 | 821.35 | 0.759 |

[1] The values of indicators (RPF, RRD, NSE, RDSR, BDSR, TU, $\beta$, RMSE) in the table are the average values of 10,000 synthetic floods.

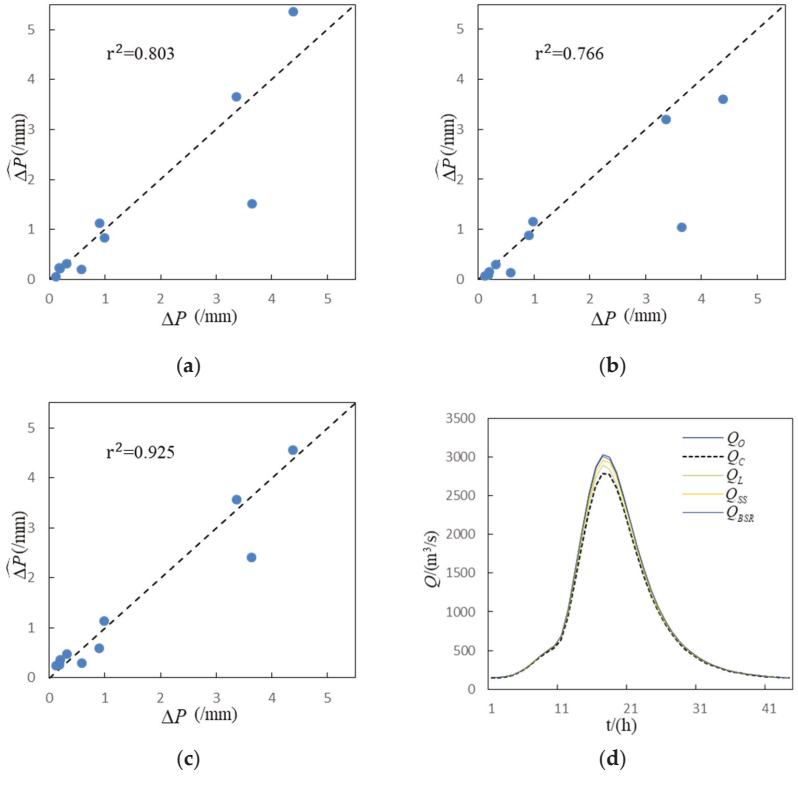



**Figure 5.** The performance of the DSRC-R method in a typical flood under three criteria. (**a**) The contrast between the real value of the rainfall error and the estimation value under the MSSFE criterion; (**b**) the contrast between the real value of the rainfall error and the estimation value under the L-curve criterion; (**c**) the contrast between the real value of the rainfall error and the estimation value under the BSR criterion; (**d**) the contrast of a typical flood forecasted flow, where $Q_O$ is the "observed" flow, $Q_C$ is the forecasted flow, $Q_L$ is the updated forecasted flow under the L-curve criterion, $Q_{SS}$ is the updated forecasted flow under the MSSFE criterion and $Q_{BSR}$ is the updated forecasted flow under the BSR criterion.

As it is depicted in Figure 5a–c, the estimated value of the rainfall error under the BSR criterion is the closest to the "true" value, the total difference is 0.8 mm (1.2 mm errors under the L-curve criterion and 0.9 mm errors under the MSSFE criterion), the correlation coefficient reaches 0.962 (0.875 under the L-curve criterion and 0.896 under the MSSFE criterion) and the spots are evenly distributed on both sides of the 1:1 line. All of the above contribute to the DSRC-R method achieving the best performance under the BSR criterion. As it is shown in Figure 5d, the indicators RPT and RRD have values of 0.9% and 0.6%, respectively, under the BSR criterion. RPT has a value of 2.7% and RRD a value of 1.4% under the MSSFE criterion. Lastly, RPT has a value of 4.5% and RRD a value of 3.8% under the L-curve criterion. In a typical flood, the optimal value of $\beta$ under the MSSFE criterion is 2.47, which is significantly less than 651.33 under the BSR criterion. This result shows that the instability of the method is not sufficiently alleviated. Therefore, the points in Figure 5a are scattered. The L-curve criterion does not fully consider the properties of the simulated residuals, meaning the flow correction result in Figure 5d is not satisfactory.

Compared with the L-curve criterion, the BSR criterion improves the performance of the DSRC-R method. This is because the BSR criterion takes more account of the properties

of model-simulated residuals, including the mutual independence between residuals and the zero mean property of residuals. As is shown in Table 3, the average BDSR under the BSR criterion is 66.9 $m^3/s$, which is significantly less than 80.8 $m^3/s$ under the L-curve criterion, indicating that the former can better reflect the zero mean property of the simulated residuals; moreover, the average RDSR under the BSR criterion is 12.88, which is significantly greater than 5.57 under the L-curve criterion, indicating that the former is more conducive to the mutual independence of simulated residuals. The above analysis shows that the BSR criterion is more conducive to the optimization of the ridge coefficient, in line with the properties of simulated residuals, than the L-curve criterion. Therefore, the BSR criterion improves the performance of the DSRC-R method. In terms of operation efficiency, the BSR criterion takes less time than the L-curve criterion because the former does not involve derivative calculation, while the latter involves derivative difference calculation.

Compared with the BSR criterion, the MSSFE criterion tends to make the value of $\beta$ smaller (average value of 64.26), which is not conducive to improving the robustness of the DSRC-R method. The average RSME under the MSSFE criterion is 1.040, which is significantly greater than 0.759 under the BSR criterion, indicating that the BSR criterion makes the DSRC-R method more stable. This is because the MSSFE method is equivalent to the least square method in linear systems, meaning the value of $\beta$ should be 0 if the DSRC-R method is applied in a linear system. Although the XAJ model is a non-linear system, the value of $\beta$ will still tend to be small, which is not conducive to improving ill-conditioned problems. However, the BSR criterion introduces BDSR and RDSR, and this makes more use of the effective information contained in the simulated errors, which is conducive to avoiding a value of $\beta$ that is too small, making the DSRC-R method more stable.

To sum up, the BSR criterion has greater advantages among the three criteria. It pays more attention to extracting effective information from simulated errors of the outlet flow (in fact, any type of error will eventually be reflected here). It takes more account of the mutual independence and zero mean property of residuals, which is conducive to selecting a more reasonable value of $\beta$ and improving the performance of the DSRC-R method.

### *3.3. Real Case*

The research basin in this study is Tankeng basin, with a total area of 3330 $km^2$ and 15 precipitation stations, which is located at the tributary of the Ou River in Zhejiang Province. Tankeng basin is in the subtropical monsoon climate zone, and it enjoys a temperate climate with well-marked seasons and plenty of rainfall and sunshine. Runoff from Tankeng basin is mainly supplied by precipitation. The annual average precipitation of Tankeng basin is between 1500 and 2100 mm, and multi-annual average evaporation is 969.9 mm. The precipitation spatial distribution is uneven, as is the annual precipitation temporal distribution; based on this, the entire year can be divided into three parts with the first part "spring rain" ranging from March to April, the second part "plum rain" ranging from May to June and the third part "thunderstorm" ranging from July to September. In Tankeng basin, multi-annual average flow is 120 $m^3/s$, and the maximal runoff happens in June, with the proportion of the entire annual runoff reaching up to 19.7%. More details about Tankeng basin and its rainfall station distribution are illustrated in Figure 6.

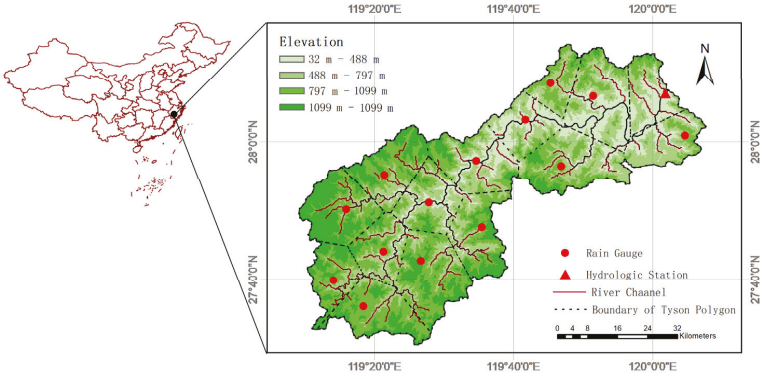



**Figure 6.** Map showing the location of Tankeng basin and depicting the stations. The map describes the location, longitude and latitude range and shape of Tankeng basin and shows the location of the rainfall stations and the flow station.

### 3.3.1. Data

There are two time scales for observed data including the day scale and the hour scale. Evaporation data are day-scale information, while precipitation and flow data are both day-scale information and hour-scale information. This research collected observed data such as evaporation, observed flow and precipitation from 1980 to 2005. The rainfall data were from 15 rainfall stations, evaporation data came from one evaporation station and observed flow data came from the outlet flow of Tankeng basin.

Tankeng basin has been used in previous studies using the XAJ model [20,28], and thus no calibration was required here. The parameters of the XAJ model at Tankeng basin are shown in Table 4.

**Table 4.** Parameters of the XAJ model for Tankeng basin.

| **Parameter** | **K** | **WM** | **WUM** | **WLM** | **WDM** | **IM** | **B** | **C** | **SM** | **EX** |
|---|---|---|---|---|---|---|---|---|---|---|
| Value | 1.296 | 150 | 20 | 80 | 50 | 0.01 | 0.3 | 0.16 | 10 | 1.5 |
| **Parameter** | **KI** | **KG** | **CS** | **CI** | **CG** | **KE** | **XE** | | | |
| Value | 0.35 | 0.35 | 0.65 | 0.865 | 0.95 | 1.466 | 0.380 | | | |

### 3.3.2. Results and Discussion

The statistical indicators of the real case are different from those of the synthetic case. In the real case, the "true" value of rainfall cannot be obtained, meaning RMSE cannot be applied. Here, the indicators include RPF, RRD, NSE and TU.

In this study, 31 historical floods in the basin were selected to compare the results of three criteria including the L-curve criterion, MSSFE criterion and BSR criterion. In order to show the performance of the DSRC-R method under each criterion, the statistical indicators of each flood are listed in Table A1 (see Appendix C). In order to more clearly compare the results of the BSR criterion and the other two criteria, scatter diagrams of each indicator of 31 floods were constructed, as shown in Figure 7. As it is illustrated in Table A1, in terms of the flood peak, the average RPF is 5.42% under the L-curve criterion, 3.95% under the MSSFE criterion and 3.49% under the BSR criterion; therefore, it is apparent that the BSR criterion has a better performance. In terms of the runoff depth, the average RRD is 2.77% under the BSR criterion, which is 3.91% and 1.31% lower, respectively, than that under the L-curve and MSSFE criteria. In terms of NSE, the BSR criterion has the maximal NSE of 0.940, while the NSE under the L-curve criterion is 0.933 and 0.938 under the MSSFE criterion. In terms of operational efficiency, the average time consumed under the BSR

criterion is 12.7 s, which is 25.7 s lower than that under the L-curve criterion, and this illustrates that the operation efficiency is dramatically improved when the BSR criterion is adopted.

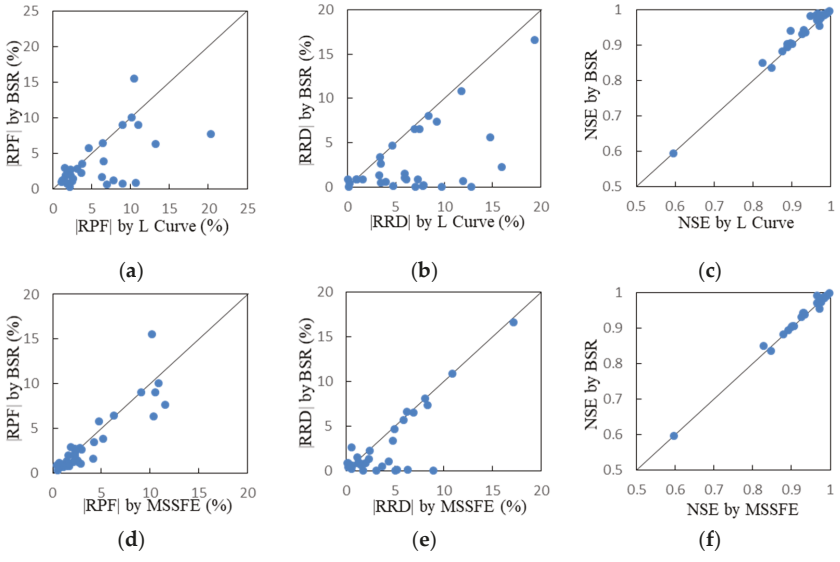



**Figure 7.** The contrast of application effects of 31 floods under three criteria. (**a**) The contrast of RPF of each flood between the BSR criterion and L-curve criterion; (**b**) the contrast of RRD of each flood between the BSR criterion and L-curve criterion; (**c**) the contrast of NSE of each flood between the BSR criterion and L-curve criterion; (**d**) the contrast of RPF of each flood between the BSR criterion and MSSFE criterion; (**e**) the contrast of RRD of each flood between the BSR criterion and MSSFE criterion; (**f**) the contrast of NSE of each flood between the BSR criterion and MSSFE criterion.

Figure 7a–c show the comparison of the statistical indicators of 31 floods between the BSR criterion and L-curve criterion. Figure 7a,b are inclined to the lower side of the 1:1 line, and Figure 7c is inclined to the upper side of the 1:1 line, indicating that the BSR criterion is more effective than the L-curve criterion for most floods. In terms of the flood peak and runoff depth, the BSR criterion is more effective in 24 floods and 29 floods, respectively; additionally, the NSE of 23 floods is greater under the BSR criterion. Figure 7d–f show the comparison of flood indicators between the BSR criterion and MSSFE criterion. Similar to Figure 7a–c, under the BSR criterion, the forecasting accuracy of 21 flood peaks and 25 flood runoff depths is higher, and NSE of 18 floods is larger, indicating that the BSR criterion has more advantages over the MSSFE criterion for most floods.

Compared with the L-curve criterion, the BSR criterion improves the performance of the DSRC-R method. This is because the BSR criterion introduces BDSR and RDSR, thus more fully considering the properties of the model-simulated residuals including the zero mean property and the mutual independence. The results in Table A1 show that the average BDSR corresponding to the BSR criterion is 648.1 $m^3/s$, which is far less than 1269.9 $m^3/s$ of the L-curve criterion, indicating that the BSR criterion is more conducive to the zero mean property of the simulated residuals. At the same time, the average RDSR corresponding to the BSR criterion is 1.69, which is greater than 1.28 of the L-curve criterion, indicating that the BSR criterion tends to meet the mutual independence of the simulated residuals. Therefore, under the BSR criterion, the DSRC-R method has a better performance. In the real case, the mean value of $\beta$ under the MSSFE criterion is 327.67, which is far less than 2048.97 under the BSR criterion. This is because the MSSFE criterion only considers the sum of squares of errors, which tends to make $\beta$ small, while the BSR criterion focuses

on the nature of the simulated residuals and extracts more effective information from the outlet flow error information, which is conducive to avoiding a small value of $\beta$. Therefore, the application of the BSR criterion in the real case is better than that of the MSSFE criterion.

## 4. Conclusions and Prospect

A previous study [19] proposed the DSRC-R method and verified that it has stronger robustness. However, the selection criterion of the ridge coefficient, usually the L-curve criterion [17–19], has only received limited attention. This essay constructed the BSR criterion based on the properties of model-simulated residuals, utilizing the indicator RDSR to quantitatively describe the independence of the residual series while utilizing the indicator BDSR to quantitatively describe the zero mean property of residuals. Additionally, we then contrasted the performance of the DSRC-R method under three different criteria (BSR criterion, L-curve criterion and MSSFE criterion) through synthetic and real-data studies.

From the results, we found that among the three criteria, the BSR criterion is more suitable for the DSRC-R method. Compared with the L-curve criterion, the BSR criterion improves the performance of the DSRC-R method. This is because the BSR criterion introduces RDSR and BDSR, which quantitatively describe the mutual independence of model-simulated residuals and zero mean property of model-simulated residuals, respectively. Moreover, the BSR criterion saves more time than the L-curve criterion because the BSR criterion does not involve derivative calculation. In addition, compared with the traditional MSSFE criterion, the BSR criterion is more conducive to enhancing the robustness of the DSRC-R method. The MSSFE criterion tends to make the ridge coefficient $\beta$ smaller, which is unfavorable to the performance of the DSRC-R method. Meanwhile, the BSR criterion is conducive to avoiding a small ridge coefficient by extracting more effective information contained in the simulated errors, and this makes the DSRC-R method more robust and improves its performance.

Further research is needed. The BSR criterion proposed in this paper improves the performance of the DSRC-R method, and this seems to benefit from the rational use of outlet flow information. In recent years, data assimilation technologies combined with radar information and remote sensing information have been continuously emerging and have received a significant amount of attention. However, making full use of outlet flow information to update hydrological elements deserves more attention. This is because hydrological elements such as rainfall, evaporation and soil moisture will eventually be reflected in the outlet flow. Therefore, much more effort needs to be made in this regard.

**Author Contributions:** Conceptualization, W.B. and K.L. (Kexin Liu); methodology, W.B. and K.L. (Kexin Liu); software, Y.H. and Y.S.; formal analysis, K.L. (Kexin Liu); data curation, L.L.; writing—original draft preparation, K.L. (Kuang Li); writing—review and editing, D.L.; visualization, Y.S.; supervision, Y.H.; project administration, Y.H.; funding acquisition, L.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Fundamental Research Funds of IWHR (grant No. WH0145B032024), the Special Project of National Science and Technology Basic Resources Investigation (grant No. 2021xjkk04005) and the National Key R&D Program of China (grant No. 2018YFC0406400).

**Data Availability Statement:** Data used for this study can be made available upon request.

**Acknowledgments:** We gratefully acknowledge financial support from the Fundamental Research Funds of IWHR and the National Key R&D Program of China.

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

## Appendix A. Proof of $\hat{\Delta P}_{RE} = \sum_{i=1}^{n} \frac{k_i}{\lambda_i+\beta} v_i$ in the DSRC-R Method

Let $\lambda_i (i = 1, \cdots, n)$ be the eigenvalue of matrix $U^T U$.

Next, $\lambda_i + \beta (i = 1, \cdots, n)$ is the eigenvalue of matrix $U^T U + \beta I$. Observe that matrix $S^T S$ and matrix $U^T U + \beta I$ have the same eigenvectors.

We already have $S^T S$ as a symmetric matrix; therefore, the N mutually orthogonal unit eigenvector $v_i (i = 1, \cdots, n)$ exists. Hence

$$S^T S v_i = \lambda_i v_i \therefore (S^T S + \beta I) v_i = (\lambda_i + \beta) v_i$$

Let $A = (v_1, \cdots, v_n)$ be the orthogonal matrix.

Next, $A$ is the full rank; thus, $Ak = S^T \Delta Q$ must have a unique solution, ... Then, $S^T \Delta Q$ has a linear expression with $v_i (i = 1, \cdots, n)$ and can be solved with only one group $k_i (i = 1, \cdots, n)$. Hence

$$\begin{aligned}
(S^T S + \beta I) \Delta \hat{P}_{RE} &= S^T \Delta Q \\
&= \sum_{i=1}^{n} k_i v_i \\
&= \sum_{i=1}^{n} \tfrac{k_i}{\lambda_i + \beta} [(\lambda_i + \beta) v_i] \\
&= \sum_{i=1}^{n} \tfrac{k_i}{\lambda_i + \beta} [(S^T S + \beta I) v_i] \\
&= (S^T S + \beta I) \sum_{i=1}^{n} \tfrac{k_i}{\lambda_i + \beta} v_i \\
&\therefore (S^T S + \beta I) \Delta \hat{P}_{RE} = (S^T S + \beta I) \sum_{i=1}^{n} \tfrac{k_i}{\lambda_i + \beta} v_i
\end{aligned}$$

Notice that $S^T S + \beta I$ is reversible. Hence, $\Delta \hat{P}_{RE} = \sum_{i=1}^{n} \frac{k_i}{\lambda_i + \beta} v_i$.

## Appendix B. Proof of $f(\beta) = \log[\sum_{i=1}^{n} (\frac{k_i}{\lambda_i + \beta})^2]$

Notice that if we want to prove the equation $f(\beta) = \log[\sum_{i=1}^{n} (\frac{k_i}{\lambda_i + \beta})^2]$, firstly, we need to prove $\|\Delta \hat{P}_{RE}\|_2^2 = \sum_{i=1}^{n} (\frac{k_i}{\lambda_i + \beta})^2$.

Furthermore,

$$\|\Delta \hat{P}_{RE}\|_2^2 = (\Delta \hat{P}_{RE}, \Delta \hat{P}_{RE}) = (\sum_{i=1}^{n} \frac{k_i}{\lambda_i + \beta} v_i, \sum_{i=1}^{n} \frac{k_i}{\lambda_i + \beta} v_i) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{k_i}{\lambda_i + \beta} \frac{k_j}{\lambda_j + \beta} (v_i, v_j)$$

where $(v_i, v_j)$ is the inner product of $v_i, v_j$.

Notice that $v_i$ ($i$ = 1, ..., n) is the mutually orthogonal unit vector.

$$\begin{aligned}
&\therefore (v_i, v_i) = 0 \text{ when } i \neq j \text{ and } (v_i, v_i) = 1 \ (i = 1, 2 \cdots n) \\
&\therefore \|\Delta \hat{P}_{RE}\|_2^2 = \sum_{i=1}^{n} (\tfrac{k_i}{\lambda_i + \beta})^2 \\
&\therefore f(\beta) = \log[\sum_{i=1}^{n} (\tfrac{k_i}{\lambda_i + \beta})^2].
\end{aligned}$$

## Appendix C

**Table A1.** The performance of the DSRC-R method on each flood in Tankeng basin.

| Flood Number | Before Updating | | | L-Curve | | | | | | MSSFE | | | | | | | BSR | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RPF | RRD | NSE | RPF | RRD | NSE | BDSR | RDSR | TU | RPF | RRD | NSE | TU | BDSR | RDSR | RPF | RRD | NSE | TU | BDSR | RDSR |
| 31850823 | 10.1 | −14.41 | 0.934 | 6.42 | −9.26 | 0.974 | 468.0 | 1.32 | 25.2 | 6.25 | −8.31 | 0.975 | 8.1 | 959.6 | 1.30 | 6.41 | −7.38 | 0.985 | 15.3 | 6.2 | 1.33 |
| 31860520 | 2.77 | −11.18 | 0.743 | 2.56 | 9.71 | 0.972 | 2958.1 | 1.59 | 42.5 | 2.52 | 8.89 | 0.973 | 21.1 | 2718.9 | 1.83 | 1.38 | −0.02 | 0.954 | 18.7 | 2461.2 | 3.25 |
| 31870910 | −11.7 | −7.38 | 0.952 | 8.93 | 3.38 | 0.963 | 1256.9 | 1.06 | 29.4 | −1.72 | −4.77 | 0.986 | 6.4 | 1418.6 | 1.09 | 0.74 | −3.35 | 0.984 | 9.9 | 40.6 | 2.35 |
| 31880520 | −16.44 | −10.91 | 0.762 | 6.93 | 7.49 | 0.888 | 2070.8 | 1.10 | 43.3 | 0.95 | −6.17 | 0.903 | 11.1 | 236.0 | 1.31 | 0.65 | −6.57 | 0.904 | 10.0 | 32.1 | 1.48 |
| 31880922 | −5.48 | −13.33 | 0.755 | −1.19 | 6.99 | 0.976 | 2314.3 | 1.13 | 28.6 | −0.66 | 5.06 | 0.979 | 30.6 | 2530.1 | 1.16 | 1.15 | 0.03 | 0.979 | 35.3 | 149.5 | 1.56 |
| 31890527 | −17.28 | −3.54 | 0.805 | 3.65 | 3.97 | 0.971 | 673.3 | 1.05 | 46.6 | 2.29 | 0.56 | 0.971 | 9.2 | 11.1 | 1.02 | 2.25 | 0.54 | 0.975 | 14.7 | 5.5 | 1.33 |
| 31890618 | −15.11 | −14.14 | 0.753 | 6.26 | 12.78 | 0.966 | 1210.2 | 1.08 | 24.9 | −4.18 | −3.05 | 0.977 | 7.8 | 2479.9 | 1.05 | 1.63 | −0.02 | 0.973 | 8.8 | 2.78 | 1.13 |
| 31890721 | −13.78 | −17.3 | 0.866 | 1.87 | 7.78 | 0.982 | 1387.1 | 1.24 | 32.8 | 1.46 | 6.31 | 0.984 | 10.0 | 1623.5 | 1.26 | 0.76 | −0.12 | 0.986 | 13.3 | 49.0 | 1.35 |
| 31920830 | −13.34 | 9.85 | 0.959 | −1.57 | 6.98 | 0.992 | 1847.3 | 1.40 | 39.3 | −1.82 | 6.88 | 0.992 | 15.5 | 1304.8 | 1.61 | −2.94 | 6.5 | 0.992 | 17.0 | 1161.8 | 1.73 |
| 31940821 | 13.96 | −15.18 | 0.935 | 4.64 | 1.61 | 0.991 | 2249.6 | 1.29 | 17.6 | 4.76 | 1.31 | 0.991 | 9.5 | 1934.0 | 1.33 | 5.76 | −0.82 | 0.99 | 8.5 | 12.5 | 1.37 |
| 31960801 | −17.42 | −1.81 | 0.937 | −2.26 | 0.22 | 0.971 | 4241.8 | 1.99 | 24.3 | −2.31 | 0.2 | 0.971 | 5.4 | 4231.6 | 1.98 | −1.9 | 0.4 | 0.971 | 9.7 | 4344.9 | 2.06 |
| 31970703 | −4.69 | −0.37 | 0.781 | 1.32 | 14.73 | 0.983 | 3240.7 | 1.15 | 29.4 | 2.12 | 5.89 | 0.983 | 8.0 | 3179.3 | 1.16 | 1.2 | 5.64 | 0.983 | 13.5 | 3068.3 | 1.17 |
| 31980618 | −5.91 | 4.28 | 0.96 | 1.43 | 15.92 | 0.997 | 1120.8 | 2.18 | 26.4 | 1.43 | 2.41 | 0.997 | 10.9 | 1127.6 | 2.22 | 1.32 | 2.23 | 0.997 | 7.9 | 1012.9 | 3.78 |
| 31990525 | −18.28 | −8.19 | 0.917 | 10.67 | 7.84 | 0.947 | 400.8 | 1.06 | 31.7 | −0.45 | 0.48 | 0.983 | 8.3 | 135.7 | 1.08 | −0.84 | 0.22 | 0.983 | 7.6 | 2.2 | 1.09 |
| 31990711 | −2.94 | −8.89 | 0.867 | 2.23 | 3.46 | 0.967 | 1430.2 | 1.13 | 24.6 | 2.27 | −0.54 | 0.967 | 9.9 | 1133.0 | 1.14 | 2.71 | −2.64 | 0.99 | 6.4 | 8.3 | 1.14 |
| 31000609 | −20.05 | −7.3 | 0.946 | −8.96 | 0.14 | 0.964 | 671.1 | 1.78 | 32.3 | −10.56 | −1.68 | 0.968 | 12.1 | 663.2 | 1.80 | −9.03 | 0.03 | 0.971 | 10.3 | 623.6 | 1.80 |
| 31000823 | 8.75 | 68.53 | 0.907 | 3.14 | 7.29 | 0.991 | 1621.2 | 1.06 | 26.2 | 2.81 | −1.37 | 0.992 | 15.6 | 990.1 | 1.63 | 2.81 | −0.87 | 0.991 | 8.6 | 967.0 | 1.66 |
| 31030624 | −11.61 | 58.79 | 0.866 | −1.83 | 6.08 | 0.981 | 659.4 | 1.03 | 24.9 | −1.07 | −1.97 | 0.982 | 8.0 | 522.1 | 1.02 | 0.7 | −0.81 | 0.986 | 5.2 | 32.6 | 1.89 |
| 31040812 | 19.29 | 36.59 | 0.84 | 20.29 | 5.93 | 0.896 | 1214.4 | 1.52 | 15.2 | 11.58 | 1.11 | 0.93 | 4.0 | 1203.5 | 1.49 | 7.68 | −1.46 | 0.941 | 6.5 | 1205.4 | 1.53 |
| 31850604 | −13.11 | −17.32 | 0.754 | 6.49 | 3.3 | 0.934 | 573.5 | 1.14 | 11.7 | 5.14 | 2.3 | 0.935 | 5.3 | 443.1 | 1.15 | 3.87 | 1.31 | 0.935 | 6.8 | 4.4 | 1.28 |
| 31850626 | −10.91 | −3.17 | 0.821 | −2.03 | 11.97 | 0.896 | 59.6 | 2.01 | 16.3 | −2.93 | −1.49 | 0.906 | 10.5 | 388.3 | 1.08 | −2.61 | −0.7 | 0.905 | 9.7 | 22.9 | 2.12 |
| 31860330 | −27.8 | −18.3 | 0.68 | 3.71 | 0.96 | 0.968 | 28.7 | 1.15 | 38.7 | 4.23 | 1.23 | 0.969 | 36.9 | 112.8 | 1.14 | 3.47 | 0.86 | 0.968 | 33.0 | 8.7 | 1.15 |
| 31860428 | 41.36 | 19.54 | 0.02 | 10.92 | 11.8 | 0.825 | 1488.3 | 1.09 | 19.7 | 9.04 | 10.89 | 0.828 | 6.8 | 517.8 | 1.32 | 9.02 | 10.87 | 0.849 | 5.9 | 515.1 | 1.32 |
| 31870411 | −2.08 | −6.7 | 0.724 | 13.18 | 5.89 | 0.875 | 2131.9 | 1.29 | 14.5 | 10.36 | 4.31 | 0.879 | 5.4 | 1975.7 | 1.22 | 6.34 | 0.99 | 0.879 | 14.1 | 1806.3 | 1.37 |
| 31870527 | −23.87 | −9.43 | 0.738 | −10.41 | 3.49 | 0.847 | 298.6 | 1.19 | 16.3 | −10.2 | 3.62 | 0.847 | 9.7 | 225.3 | 1.20 | −15.5 | 0.5 | 0.836 | 16.8 | 225.1 | 1.20 |
| 31880327 | −19.21 | −6.61 | 0.869 | −10.11 | −4.66 | 0.901 | 404.8 | 1.02 | 24.8 | −10.87 | −4.94 | 0.901 | 7.6 | 168.0 | 1.06 | −10.08 | −4.65 | 0.903 | 8.5 | 8.36 | 1.06 |
| 31880524 | 8.66 | −2.92 | 0.429 | 1.59 | 0.09 | 0.926 | 50.4 | 1.03 | 19.9 | 1.61 | 0.1 | 0.926 | 10.3 | 23.3 | 1.05 | −1.94 | −0.83 | 0.931 | 6.4 | 4.1 | 1.21 |
| 31880613 | −29.34 | −10.44 | 0.193 | 1.07 | 0.91 | 0.595 | 353.5 | 1.23 | 16.1 | −0.91 | −0.15 | 0.597 | 9.0 | 431.7 | 1.20 | 0.91 | 0.82 | 0.595 | 5.3 | 4.8 | 1.24 |
| 31880627 | 0.97 | 47.39 | 0.412 | 2.51 | 4.76 | 0.967 | 876.5 | 1.17 | 18.4 | 2.84 | 5.15 | 0.967 | 5.3 | 411.9 | 2.04 | −1.09 | −0.07 | 0.971 | 17.0 | 480.5 | 2.70 |
| 31890413 | 14.65 | 10.1 | 0.125 | 2.18 | 8.35 | 0.931 | 617.2 | 1.03 | 20.8 | 0.54 | 8.11 | 0.932 | 3.8 | 516.5 | 1.07 | 0.27 | 8.07 | 0.932 | 8.6 | 492.5 | 2.16 |
| 31980513 | 28.8 | 57.29 | 0.098 | 7.8 | 19.32 | 0.889 | 1438.5 | 1.03 | 13.8 | 2.63 | 17.14 | 0.892 | 8.1 | 1312.3 | 1.65 | 1.25 | 16.6 | 0.893 | 16.2 | 1332.4 | 2.58 |
| Average [1] | 14.51 | 16.81 | 0.721 | 5.42 | 6.68 | 0.933 | 1269.6 | 1.28 | 25.7 | 3.95 | 4.08 | 0.938 | 10.7 | 1126.7 | 1.34 | 3.49 | 2.77 | 0.940 | 12.1 | 648.1 | 1.69 |

[1] The average values of RPF and RRD actually represent the average of the absolute values of RPF and RRD, respectively.

## References


1. Wang, Y.; Liu, R.; Guo, L.; Tian, J.; Zhang, X.; Ding, L.; Wang, C.; Shang, Y. Forecasting and providing warnings of flash floods for ungauged mountainous areas based on a distributed hydrological model. *Water* **2017**, *9*, 776. [CrossRef]
2. Cheng, W.; Huang, C.; Hsu, N.; Wei, C. Risk Analysis of Reservoir Operations Considering Short-Term Flood Control and Long-Term Water Supply: A Case Study for the Da-Han Creek Basin in Taiwan. *Water* **2017**, *9*, 424. [CrossRef]
3. Bao, W.; Si, W.; Shen, G.; Zhang, X.; Li, Q. Runoff error updating based on unit hydrograph inversion. *Adv. Water Sci.* **2012**, *23*, 315–322.
4. Abrahart, R.J.; See, L. Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. *Hydrol. Process.* **2000**, *14*, 2157–2172. [CrossRef]
5. Broersen, P.M.; Weerts, A.H. Automatic Error Correction of Rainfall-Runoff models in Flood Forecasting Systems. *IMTC IEEE* **2005**, *6*, 963–968.
6. Bogner, K.; Pappenberger, F. Multiscale error analysis, correction, and predictive uncertainty estimation in a flood forecasting system. *Water Resour. Res.* **2011**, *47*, 1772–1780. [CrossRef]
7. Sun, L.; Nistor, I.; Seidou, O. Streamflow data assimilation in SWAT model using Extended Kalman Filter. *J. Hydrol.* **2015**, *531*, 671–684. [CrossRef]
8. Moradkhani, H.; Gupta, H.; Sorooshian, S.V.; Houser, P. Combined parameter and state estimation of hydrological models using ensemble Kalman filter. *Adv. Water Resour.* **2005**, *28*, 135–147. [CrossRef]
9. Wagener, T.; Mcintyre, N.; Lees, M.J.; Wheater, H.S.; Gupta, H.V. Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis. *Hydrol. Process.* **2003**, *17*, 455–476. [CrossRef]
10. Misirli, F.; Gupta, H.V.; Sorooshian, S.; Thiemann, M. *Bayesian Recursive Estimation of Parameter and Output Uncertainty for Watershed Models*; American Geophysical Union (AGU): Washington, DC, USA, 2013.
11. Bao, W.; Ji, H.; Hu, Q.; Qu, S.; Zhao, C. Robust estimation theory and its application to hydrology. *Adv. Water Sci.* **2003**, *14*, 528–532.
12. Qu, S.; Bao, W. Comprehensive correction of real-time flood forecasting. *Adv. Water Sci.* **2003**, *14*, 167–171.
13. Dechant, C.M.; Moradkhani, H. Examining the effectiveness and robustness of sequential data assimilation methods for quantification of uncertainty in hydrologic forecasting. *Water Resour. Res.* **2012**, *48*, W04518. [CrossRef]
14. Bao, W.; Si, W.; Qu, S. Flow Updating in Real-Time Flood Forecasting Based on Runoff Correction by a Dynamic System Response Curve. *J. Hydrol. Eng.* **2014**, *19*, 747–756.
15. Si, W.; Bao, W.; Qu, S. Runoff error correction in real-time flood forecasting based on dynamic system response curve. *Adv. Water Sci.* **2013**, *24*, 497–503.
16. Si, W.; Bao, W.; Gupta, H.V. Updating real-time flood forecasts via the dynamic system response curve method. *Water Resour. Res.* **2015**, *51*, 5128–5144. [CrossRef]
17. Sun, Y.; Bao, W.; Jiang, P.; Si, W.; Zhou, J.; Zhang, Q. Development of a Regularized Dynamic System Response Curve for Real-Time Flood Forecasting Correction. *Water* **2018**, *10*, 450. [CrossRef]
18. Sun, Y.; Bao, W.; Jiang, P.; Ji, X.; Gao, S.; Xu, Y.; Si, W. Development of Multivariable Dynamic System Response Curve Method for Real-Time Flood Forecasting Correction. *Water Resour. Res.* **2018**, *54*, 4730–4749. [CrossRef]
19. Si, W.; Gupta, H.V.; Bao, W.; Jiang, P.; Wang, W. Improved Dynamic System Response Curve Method for Real-Time Flood Forecast Updating. *Water Resour. Res.* **2019**, *55*, 7493–7519. [CrossRef]
20. Liu, K.; Zhang, X.; Bao, W.; Zhao, L.; Shu, H.; Li, J. A system response correction method with runoff error smooth matrix. *J. Hydraul. Eng.* **2015**, *46*, 960–966.
21. Hansen, P.C.; Oleary, D.P. The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM J. Sci. Comput.* **1993**, *14*, 1487–1503. [CrossRef]
22. Zhao, R.J. The Xinanjiang model applied in China. *J. Hydrol.* **1992**, *135*, 371–381.
23. Hansen, P.C. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev.* **1992**, *34*, 561–580. [CrossRef]
24. Kennedy, J.; Eberhart, R.C. Particle swarm optimization. In Proceedings of the International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995.
25. Sierra, M.R.; Coello, C.A. Improving PSO-Based Multi-objective Optimization Using Crowding, Mutation and E-Dominance. *Lect. Notes Comput. Sc.* **2005**, *3410*, 505–519.
26. Mesloub, S.; Mansour, A. Hybrid PSO and GA for global maximization. *Int. J. Comput. Math* **2009**, *2*, 597–608.
27. Chen, Y.; Li, J.; Xu, H. Improving flood forecasting capability of physically based distributed, hydrological models by parameter optimization. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 375–392. [CrossRef]
28. Zhang, X.; Liu, K.; Bao, W.; Li, J.; Lai, S. Runoff error proportionality coefficient correction method based on system response. *Adv. Water Sci.* **2014**, *25*, 789–796.

water

MDPI

*Article*

# Analysis of Water Environment Quality Changes and Influencing Factors during the "Thirteenth Five-Year Plan" Period in Heilongjiang Province

**Wei Chen [1,†], Yu Bai [2,†], Bo Li [1,*], Chengcheng Feng [1,*] and Mi Zhou [3]**

[1] Heilongjiang Ecological Environment Monitoring Center, Harbin 150056, China; mxy0925@126.com
[2] China Association of Environmental Protection Industry, Beijing 100037, China; baiyu@cnemc.cn
[3] China National Environmental Monitoring Centre, Beijing 100012, China; zhoumi@cnemc.cn
* Correspondence: hlj_emcs@126.com (B.L.); zoefeng1987@foxmail.com (C.F.)
† These authors contributed equally to this work and should be regarded as co-first authors.

**Abstract:** Heilongjiang Province is located in the northeastern part of China and is the province with the highest latitude in China. As Heilongjiang Province is the most important grain production base in China, the Chinese government attaches great importance to the quality of the ecological environment in Heilongjiang Province, especially the analysis of changes in the quality of the water environment and their driving factors. We studied the changes in the environmental quality of surface water in Heilongjiang Province during the "13th Five-Year Plan" period (2016–2020), and analyzed the surface water for four major pollutants including the permanganate index, chemical oxygen demand, ammonia nitrogen and total phosphorus, and the change trends of the proportion of the water quality of class I–III and the proportion of the water quality of inferior class V. The results show that the environmental quality of surface water in Heilongjiang Province has improved significantly during the "13th Five-Year Plan". The analysis of the driving factors of the change of surface water environment quality shows that the population, the primary industry, the tertiary industry and forestry are the main factors affecting the change of water environment quality in Heilongjiang Province.

**Keywords:** "Thirteenth Five-Year Plan" period; water environment quality; Heilongjiang Province; correlation analysis; surface water

**Citation:** Chen, W.; Bai, Y.; Li, B.; Feng, C.; Zhou, M. Analysis of Water Environment Quality Changes and Influencing Factors during the "Thirteenth Five-Year Plan" Period in Heilongjiang Province. *Water* **2022**, *14*, 2367. https://doi.org/10.3390/w14152367

Academic Editors: Yaohuan Huang, Yesen Liu, Runhe Shi and Hongyan Ren

Received: 24 April 2022
Accepted: 16 July 2022
Published: 31 July 2022



## 1. Introduction

Located in northeastern China, Heilongjiang Province is the northernmost, easternmost and highest latitude province in China. Its north and east are separated from Russia across Heilongjiang River, its west is adjacent to the Inner Mongolia Autonomous Region, and its south is bordered by Jilin Province. With a total land area of 473,000 $Km^2$, Heilongjiang Province is the sixth largest province in China. The geographical location and water system of Heilongjiang Province are shown in Figure 1.

The period from 2016 to 2020 was the time when China implemented the Outline of the Thirteenth Five-Year Plan for National Economic and Social Development of the People's Republic of China, hereinafter referred to as the "Thirteenth Five-Year Plan" period. During this period, the government of Heilongjiang Province focused on the improvement of the Songhua River Basin. All 44 black and odorous water bodies were treated, 72 industrial parks realized centralized sewage treatment, and 43 water source protection areas were all rectified for environmental problems. These results demonstrate the determination of Heilongjiang Province to improve the surface water environment.

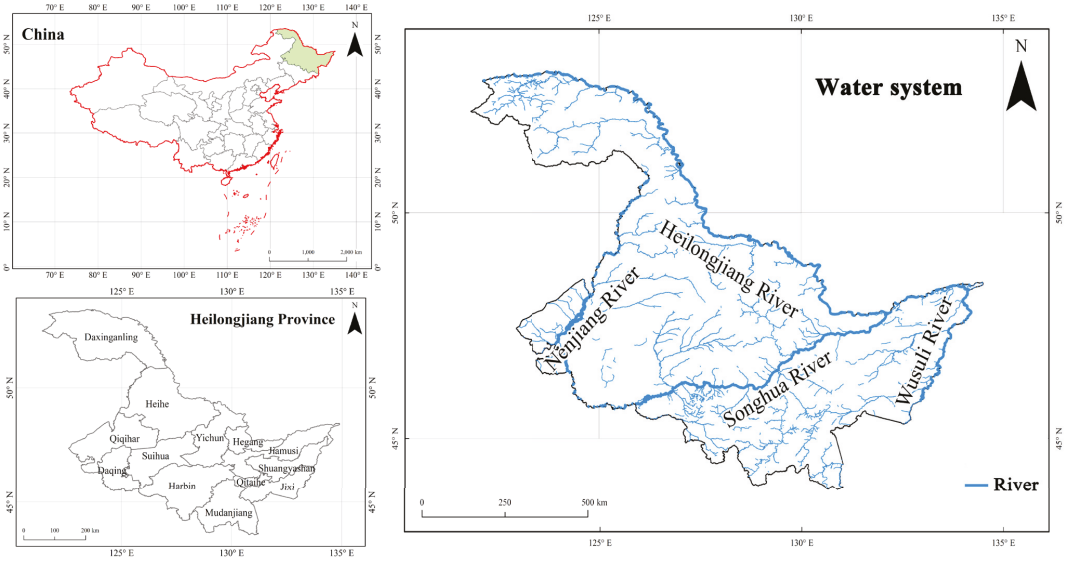



**Figure 1.** Geographical location and water system of Heilongjiang Province.

At present, scholars' research on the surface water environment mainly focuses on the investigation and monitoring of new pollutants. For example, the research of Meng Qiao et al. (2022) showed that OPAHs exist preferentially in the water environment and pose a non-negligible ecological risk to the surface water ecosystem [1]. Hai-Yan Zou et al. (2021) studied the characteristics of antibiotic resistance genes (ARGs) in surface water affected by mining, and the results show that heavy metals from mining activities have significant effects on ARGs in surface water to varying degrees [2]. Nina Henning et al. (2021) detected GBP-Lactam, NA-GBP and CCHA at levels up to 260 ng/L in the Rhine River and its tributaries, suggesting monitoring of these compounds in drinking water [3]. Silvia Galafassi et al. (2019) reported on microplastic emissions, listing all identified sources of microplastic waste to date and a quantitative assessment of environmental inputs to surface water [4]. G. Sammut et al. (2017) conducted an extensive survey of perfluoroalkyl substances (PFAS) in surface waters of the Maltese Islands and the results show that all surface water samples are contaminated with at least one PFAS, with PFOS and PFOA detected in surface water at 100% and 95%, respectively [5].

In addition to increased monitoring of new pollutants, researchers have proposed a number of methods in recent years that may improve surface water monitoring. For example, Koyel Sur et al. (2021) conducted a pilot study in northwestern India, and he used green and shortwave-infrared (SWIR) bands to modify the Modified Normalized Difference Water Index (MNDWI) method, using this method to monitor the environmental quality of surface water [6]. The monitoring data can be processed and displayed using the Google Earth Engine (GEE) platform. Sama Azadi et al. (2021) conducted continuous monitoring of surface water around a new highway in southern Norway; he used the Gamma Test theory (GTT) method to optimize the water quality monitoring network (WQMN) of the road so that WQMN can be suitable for projects with limited design and construction time and budget or projects lacking sufficient data [7]. With the development of science and technology, there are more and more types of new pollutants, and their impact on the water environment is becoming more and more complex. Monitoring and research on new pollutants is often necessary. The development of science and technology will also lead to innovations in monitoring technology, and research on the testing and application scope of new technologies is also necessary to improve the level of environmental monitoring.

However, for a region as large as Heilongjiang Province, the introduction of new pollutants or new monitoring techniques into surface water monitoring should be carefully considered. This is because we are still at the stage of "little knowledge" about the sources of new pollutants and their impact on the ecological environment. At the same time, compared with the existing four major pollutants (permanganate index, chemical oxygen demand, ammonia nitrogen and total phosphorus), the representation of new pollutants in water environmental quality monitoring is still low. In addition, the application of new measurement technology requires a lot of preliminary evaluation work to ensure the stability, practicability and data accuracy of the technology. The new approach needs also to be approved by other provinces in China. Therefore, it is still applicable to use four main pollutants (permanganate index, chemical oxygen demand, ammonia nitrogen and total phosphorus) to characterize the overall water environment in Heilongjiang Province at this stage.

In recent years, research on the influencing factors of regional water environment quality has become an emerging topic at home and abroad. In terms of research methods, principal component analysis was used to analyze the water environment quality of the basin [8–11], and this method can be combined with other methods; the selection of influencing factors also enriches its research perspective. From the perspective of the scale of the study area, a small-scale study was also carried out, usually for a certain city. For example, Shexia Zhan et al. (2021) discussed the impact of natural factors and human activities on the source water quality in Macao based on the obtained statistical results [12]. The change of the regional water environment is a complex process affected by the comprehensive action of natural social and economic factors, and is the result of the interaction of the three systems of society, economy and ecology [13–15]. Factors such as climate, population, economy, transportation, energy consumption, water resources, agriculture and forestry are the main drivers of changes in the water environment. In the analysis of the correlation between water environmental quality and the three major systems, it tends to be large in scale, time and space, and the analysis direction changes from single target to multi-target, and develops from single factor to multi-factor, from static to dynamic, from the natural environment system to a complex natural and social environment system.

To sum up, many researchers have used different theories and methods to evaluate the regional water environment, making contributions to ecological and environmental protection, and the research results have a certain utility. However, most of the existing studies are limited to the influence of a certain factor on a single indicator or pollutant, and less attention is paid to the provincial perspective. Finally, there is a lack of analysis of the relationship between different water quality indicators and natural factors, socio-economic changes, etc., and a lack of tracking of driving factors. Therefore, this study attempts to answer the following research questions: (1) How will the water environment quality change in Heilongjiang Province during the 13th Five-Year Plan? Compared with the "Twelfth Five-Year Plan", is the water quality better or worse? (2) What factors drive the change of water environment quality in Heilongjiang Province? What factors dominate? Exploring the changes in water environment quality and its influencing factors in the development process of the "Thirteenth Five-Year Plan" is not only of great practical significance for realizing the high-quality development of Heilongjiang Province, but also has certain reference value for other developing countries.

## 2. Materials and Methods

### *2.1. Study Area*

The water system in Heilongjiang Province is well developed. There are four major river systems in Heilongjiang Province, which belong to the four major river systems of Heilongjiang, Songhua River, Ussuri River and Suifen River. Among them, Heilongjiang and Ussuri River are the international boundary rivers, Xingkai Lake is the international boundary lake, Suifen River directly enters the Sea of Japan, and Songhua River and Nen River run through Heilongjiang Province. Heilongjiang Province has 2881 rivers with a

drainage area of more than 50.0 $Km^2$, 93 rivers between 1000 and 10,000 $Km^2$, and 18 rivers with an area of more than 10,000 $Km^2$. There are 253 lakes with an annual water surface area of 1.0 $Km^2$ and above, including 241 freshwater lakes and 12 saltwater lakes, with a total water surface area of 3037.0 $Km^2$ (excluding the overseas area of transboundary lakes). The main lakes are Xingkai Lake, Jingbo Lake and Lianhuan Lake.

### *2.2. Research Methods and Data*

The grey relational analysis belongs to the grey system theory, and it further studies the degree of correlation between the indicators through the similarity of the changes in the geometric shapes of the indicators [16]. The basic idea is to use the quantitative analysis of the dynamic process to calculate the correlation degree between the reference index and each comparison index in the system, and to determine the important factors that affect the reference index. It can describe the degree of correlation between variables despite incomplete information. The larger the correlation coefficient is, the closer is the relationship between the reference index and the comparison index, helping analysis of the positive factors that are conducive to the development of the system. Grey relational analysis is not limited by sample size and distribution, and is also applicable to data with short time span and irregularity. Since there is much unknown information about the mechanism of the impact of other factors on the quality of the ecological environment, it conforms to the characteristics of the grey system.

According to the characteristics of the data, it was divided into seven categories: climate, population, economy, energy, water resources, forestry, and agriculture. Among them, for the climate index we selected the average temperature, the average annual precipitation and the annual sunshine hours; for the population index we selected the population of Heilongjiang Province; for the economic index we selected the per capita GDP, primary industry, secondary industry, tertiary industry and local environmental protection expenditure; for the energy index we selected the elastic coefficient of energy consumption; for the water resources index we selected the surface water resources, the total surface water supply, the total groundwater supply, the total agricultural water use, the total ecological water consumption and the per capita water consumption; for the forestry index we selected the area of artificial afforestation in the current year; for the agricultural index we selected the pure amount of agricultural nitrogen fertilizer application, the pure amount of agricultural phosphorus fertilizer application, the pure amount of agricultural potassium fertilizer application and the amount of pesticide use; for the water environment index we selected permanganate index, chemical oxygen demand, ammonia nitrogen, total phosphorus and excellent water body proportion. The calculation method of grey relational analysis was adopted, and the details can be seen in Junli Li et al. (2020) research [17]. The data comes from the 2016–2020 "Eco-environmental Quality Status of Heilongjiang Province" [18–22], "Heilongjiang Province Eco-Environmental Quality Status Bulletin" [23–27], "Heilongjiang Ecological Environment Statistical Annual Report" [28–30], "China Statistical Yearbook 2020" [31]. The flow chart is shown in Figure 2. We can obtain the reference sequence and the comparison sequence, respectively:

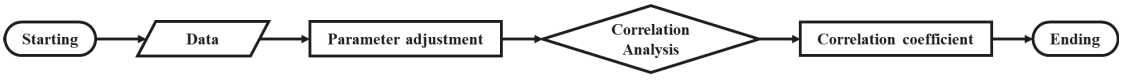



**Figure 2.** The flow chart of this study.

Reference sequence: statistics related to climate, population, economy, energy, water resources, forestry and agriculture in Heilongjiang Province from 2006 to 2020.

Comparative sequence: the annual average values of permanganate index, chemical oxygen demand, ammonia nitrogen, total phosphorus and the proportion of excellent water bodies in Heilongjiang Province from 2006 to 2020.

In addition, in order to understand the impact of pollution discharge on water environment quality in Heilongjiang Province, the Spearman correlation analysis method was used to analyze the permanganate index, chemical oxygen demand, ammonia nitrogen

and industrial source chemical oxygen demand discharge, industrial sources of ammonia nitrogen emissions, domestic sources of chemical oxygen demand emissions and domestic sources of ammonia nitrogen emissions in Heilongjiang Province from 2006 to 2019. The pollution emission data comes from the "2019 China Ecological Environment Statistical Yearbook" [32].

## 3. Results

### 3.1. Changes in the Environmental Quality of Surface Water during the "13th Five-Year Plan" Period

The proportions of water quality categories and changes in major pollutants in three different water stages of rivers in Heilongjiang Province From 2011 to 2020 can be seen in Tables 1 and 2. Among them, during the "13th Five-Year Plan" period, the proportion of water quality of class I–III in the dry season is 52.0–74.0%, the proportion of water quality in class I–III in the normal water period is 62.6–69.2%, and the proportion of water quality in class I–III in the high water period is 26.4–64.5%. From 2016 to 2020, the change of the proportion of water quality of class I–III in each water period is in a trend of fluctuation, but from 2011 to 2020, except for the wet season, the change trend of the proportion of water quality of class I–III in other water periods is a significant increase. During the "Thirteenth Five-Year Plan" period, the proportion of water quality of inferior Class V in the dry season is 3.9–9.2%, the proportion of water quality of inferior Class V in the normal water season is 2.8–7.4%, and the proportion of water quality of inferior Class V in the wet season is 0.9–5.8%. From 2016 to 2020, the change trend of the dry season and the flat water season is a fluctuating one, and the proportion of water quality of inferior Class V in the wet season has dropped significantly. The proportion of water quality of inferior Class V during the dry season from 2011 to 2020 fluctuates, and the proportion of water quality of inferior Class V during the normal and wet seasons decreases significantly. The above results show that although the water quality in the dry, flat and wet periods does not improve significantly during the "13$^{th}$ Five-Year Plan" period, the number of water bodies with poor water quality in the wet period is significantly reduced. Compared with the "Twelfth Five-Year Plan" period (2011–2015), during the "Thirteenth Five-Year Plan" period, the water quality in the dry and flat water periods is significantly improved, and the number of water bodies with poor water quality in the flat and wet periods is significantly reduced.

**Table 1.** 2011–2020 Proportion and change trend of water quality grades I–III in different water stages of rivers in Heilongjiang Province.

| Year | Dry Season | Normal Water Season | High Water Season |
|---|---|---|---|
| 2011 | 39.2% | 31.3% | 38.6% |
| 2012 | 47.1% | 48.9% | 51.1% |
| 2013 | 50.0% | 43.5% | 30.4% |
| 2014 | 58.6% | 61.1% | 55.6% |
| 2015 | 57.0% | 56.7% | 51.1% |
| 2016 | 70.9% | 66.9% | 61.2% |
| 2017 | 74.0% | 69.2% | 64.5% |
| 2018 | 52.0% | 62.6% | 26.4% |
| 2019 | 74.0% | 64.2% | 41.9% |
| 2020 | 68.9% | 67.3% | 55.1% |
| 2016–2020 rank correlation coefficient rs | −0.350 | −0.100 | −0.500 |
| Trend | volatility | volatility | volatility |
| 2011–2020 rank correlation coefficient rs | 0.770 | 0.842 | 0.200 |
| Trend | Significant increase | Significant increase | Significant increase |

**Table 2.** The proportion and trend of water quality of inferior Class V in each water stage of Heilongjiang Province from 2011 to 2020.

| Year | Dry Season | Normal Water Season | High Water Season |
|---|---|---|---|
| 2011 | 7.6% | 7.2% | 8.4% |
| 2012 | 13.8% | 6.7% | 6.7% |
| 2013 | 6.8% | 7.6% | 7.6% |
| 2014 | 3.4% | 5.6% | 5.6% |
| 2015 | 5.8% | 5.6% | 5.6% |
| 2016 | 5.5% | 7.4% | 5.8% |
| 2017 | 5.0% | 2.8% | 3.7% |
| 2018 | 9.2% | 5.6% | 2.8% |
| 2019 | 7.0% | 3.8% | 1.9% |
| 2020 | 3.9% | 2.8% | 0.9% |
| 2016–2020 rank correlation coefficient rs | −0.200 | −0.400 | −1.000 |
| Trend | volatility | volatility | Significant decrease |
| 2011–2020 rank correlation coefficient rs | −0.333 | −0.648 | −0.939 |
| Trend | volatility | Significant decrease | Significant decrease |

The trend of major pollutants in rivers in Heilongjiang Province from 2011 to 2020 is shown in Table 3. During the "Thirteenth Five-Year Plan" period, the four major pollutants in Heilongjiang Province show a fluctuating trend. Compared with the end of the "Twelfth Five-Year Plan" (2015), the main pollution indicators of rivers in Heilongjiang Province, the permanganate index, chemical oxygen demand, ammonia nitrogen and total phosphorus pollution concentration decrease by 16.7%, 18.2%, 34.7% and 33.3% respectively, and the permanganate index, chemical oxygen demand, ammonia nitrogen and total phosphorus show a significant downward trend from 2011 to 2020. The above results show that the downward trend of the concentration of major pollutants during the "13$^{th}$ Five-Year Plan" period is not obvious, but compared with the "12$^{th}$ Five-Year Plan" period, the concentration of major pollutants has dropped significantly.

**Table 3.** The major pollutants in rivers of Heilongjiang Province from 2011 to 2020.

| Year | Permanganate Index (mg/L) | Chemical Oxygen Demand (mg/L) | Ammonia Nitrogen (mg/L) | Total Phosphorus (mg/L) |
|---|---|---|---|---|
| 2011 | 6.7 | 25 | 0.746 | 0.16 |
| 2012 | 6.3 | 23 | 0.722 | 0.15 |
| 2013 | 6.6 | 23 | 0.683 | 0.15 |
| 2014 | 6.4 | 22 | 0.561 | 0.14 |
| 2015 | 6.6 | 22 | 0.575 | 0.15 |
| 2016 | 5.7 | 20 | 0.571 | 0.12 |
| 2017 | 5.4 | 19 | 0.593 | 0.12 |
| 2018 | 5.9 | 21 | 0.622 | 0.13 |
| 2019 | 5.8 | 21 | 0.495 | 0.10 |
| 2020 | 5.5 | 18 | 0.427 | 0.10 |
| 2016–2020 rank correlation coefficient rs | 0.1 | −0.2 | −0.6 | −0.6 |
| Trend | volatility | volatility | volatility | volatility |
| 2011–2020 rank correlation coefficient rs | −0.782 | −0.879 | −0.770 | −0.927 |
| Trend | Significant decrease | Significant decrease | Significant decrease | Significant decrease |

### *3.2. Correlation Analysis between Pollution Discharge and Surface Water Environmental Quality*

The Spearman correlation coefficient was used to indicate the strength of the correlation between pollution discharge and major pollutants in the surface water environment, as shown in Figure 3. Among them, X1: surface water permanganate index, X2: surface water chemical oxygen demand, X3: surface water ammonia nitrogen, X4: industrial source chemical oxygen demand discharge, X5: industrial source ammonia nitrogen discharge, X6: living source chemical oxygen demand emissions, X7: ammonia nitrogen emissions from living sources. The results show that surface water permanganate index, chemical oxy-

gen demand, ammonia nitrogen and industrial source chemical oxygen demand, ammonia nitrogen emission, living source chemical oxygen demand and ammonia nitrogen emission are all positively correlated. Among them, the permanganate index has a significant positive correlation with the chemical oxygen demand of industrial sources, chemical oxygen demand of living sources and ammonia nitrogen emissions from living sources, and has a very significant positive correlation with ammonia nitrogen emissions from industrial sources. The chemical oxygen demand of surface water has a very significant positive correlation with industrial source chemical oxygen demand, ammonia nitrogen, and living source chemical oxygen demand and ammonia nitrogen. There was a significant positive correlation between surface water ammonia nitrogen and chemical oxygen demand of industrial sources and ammonia nitrogen emissions from domestic sources.

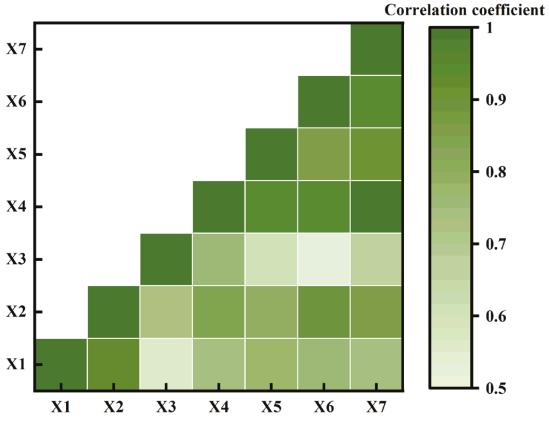



**Figure 3.** Spearman correlation coefficients between pollutant emissions and major pollutants in surface water.

### *3.3. Correlation Analysis between Surface Water Environmental Quality and Other Factors*

The grey correlation degree and its ranking of the surface water environmental quality comparison series are shown in Table S2 (Supporting Information). Specifically, the influencing factors with high correlation with the permanganate index are: the tertiary industry, the population of Heilongjiang Province, and net application amount of agricultural compound fertilizer. The influencing factors with a high degree of correlation with chemical oxygen demand are: tertiary industry, the area of artificial afforestation in the current year and the annual sunshine hours. The influencing factors with a high degree of correlation with ammonia nitrogen are: total surface water supply, pure amount of agricultural nitrogen fertilizer application, and per capita water consumption. The influencing factors with high correlation with total phosphorus are: total surface water supply, pure nitrogen fertilizer application and per capita water consumption. The influencing factors with high correlation with the proportion of excellent water quality are: primary industry, annual sunshine hours and tertiary industry.

In order to comprehensively evaluate the correlation between surface water environmental quality and various factors, the average method was used to evaluate the correlation index. There are 3 factors in the category of high correlation degree, which are population, forestry and agriculture in descending order of average degree of correlation, and 3 factors in the category of medium correlation degree, which are economy, meteorology and water resources in descending order of average degree of correlation. The factor in the category of low correlation degree is energy. The correlation statistics of each reference sequence are shown in Figure 4, and the average value is shown in Figure 5.

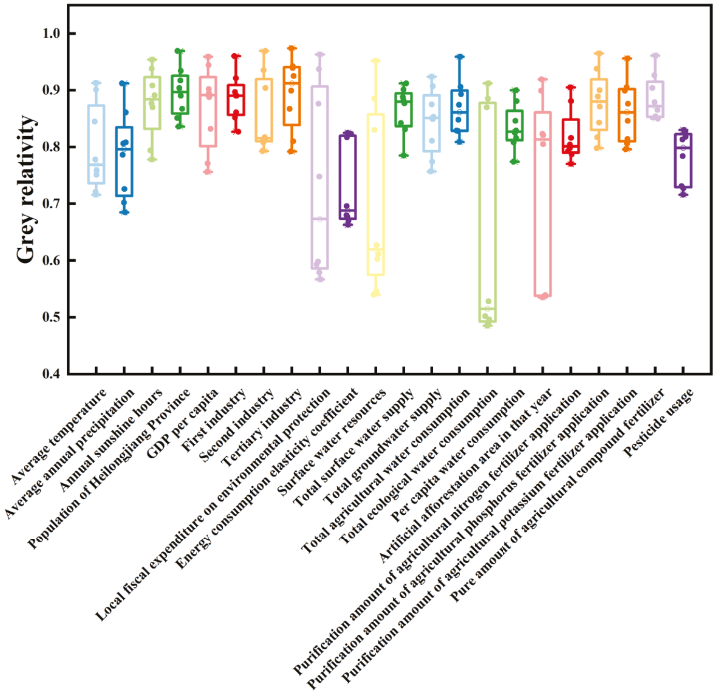



**Figure 4.** Statistical map of the correlation between each reference sequence and the environmental quality of surface water.

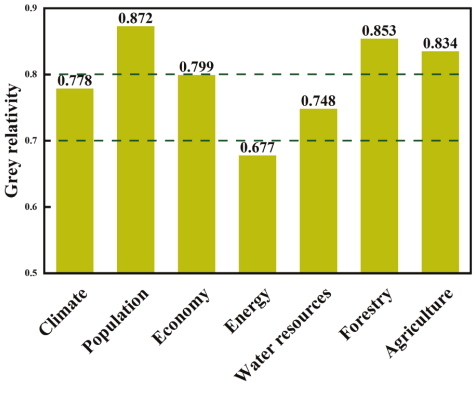



**Figure 5.** The average value of the correlation between each reference series and the environmental quality of surface water.

### High correlation factor

The average grey correlation between population factors and water environment quality in Heilongjiang Province is 0.872, ranking first among the seven categories of factors, showing a very high correlation. Among the seven categories of factors, population has the greatest impact on water quality. The average grey correlation between forestry factors and water environment quality in Heilongjiang Province is 0.853, ranking second among the seven categories of factors, showing a very high correlation. The average grey correlation between agricultural factors and water environment quality in Heilongjiang

Province is 0.834, ranking third among the seven categories of factors, showing a very high correlation.

**Medium correlation factor**

The average grey correlation degree between economic factors and water environment quality is 0.799, and the correlation degree is moderate. Among them, the primary industry, the tertiary industry, the per capita GDP, the secondary industry and the water environment quality are highly correlated, and the local financial environmental protection expenditure is low. The average grey correlation degree between climatic factors and water environment quality is 0.778, and the correlation degree is moderate. Among the climatic factors, the correlation degree of annual sunshine hours is greater than average temperature and precipitation, which is similar to ambient air quality. The average grey correlation degree between water resource factors and water environment quality is 0.748, and the correlation degree is medium. The average grey correlation degree of each factor in the water resources factor and the water environment quality is in descending order: total surface water supply, per capita water consumption, total agricultural water consumption, total groundwater water supply, surface water resources and total ecological water consumption quantity. Among them, the total surface water supply, per capita water consumption, total agricultural water supply and total groundwater supply are highly correlated with the environmental quality of surface water, while surface water resources and total ecological water are poorly correlated.

**Low correlation factor**

The average grey correlation between energy factors and water environment quality is 0.677, which is low. Among them, the energy consumption elasticity coefficient has a low correlation with the permanganate index, chemical oxygen demand, ammonia nitrogen, total phosphorus and the proportion of good water quality.

## 4. Discussion

During the "Thirteenth Five-Year Plan" period, the proportion of water quality of Class I–III increased, and the proportion of water quality of inferior Class V and the concentration of major pollutants decreased, indicating that the Heilongjiang Provincial Government's continuous "clear water defense war" has achieved remarkable results, mainly including continuous encryption monitoring and special inspection of law enforcement to ensure that inferior water bodies such as Ash River, Waken River, and Indus River do not rebound. In addition, through inter-departmental linkages to carry out special actions and cross-monitoring of water quality, the treatment of 44 black and odorous water bodies has been completed [33].

The primary industry has the greatest impact on the proportion of good water quality, and the primary industry refers to farmers and agriculture, forestry, animal husbandry, fishery, etc. As a major agricultural province, Heilongjiang Province has four major water systems flowing through a large amount of farmland. The surface water near the farmland is greatly affected by agricultural activities, resulting in the pollution of downstream waters [34]. This is consistent with the research results that the application rate of agricultural chemical fertilizers has a great influence on the content of ammonia nitrogen and total phosphorus in surface water. Heilongjiang Province is located in the highest latitude area in China, with four distinct seasons of precipitation, and the precipitation in the wet season is much higher than that in other water seasons. Fluctuations in water quality of Class I–III during the wet season and industrial wastewater discharge have little effect on ammonia nitrogen concentration in surface water; domestic wastewater discharge and primary industry have a greater impact on surface water ammonia nitrogen concentration. These results show that the effluent from farmland water pollution caused by surface water runoff and pollution caused by domestic wastewater discharge are the main sources of ammonia nitrogen and total phosphorus pollution in surface water in Heilongjiang Province. Therefore, the surface water pollution caused by living sources and agricultural sources

should be listed as the key work of the future surface water environment management in Heilongjiang Province [35].

The tertiary industry has a greater impact on the permanganate index, chemical oxygen demand and the ratio of good water quality. China's tertiary industry is other than the primary and secondary industries, and includes water conservancy, environment and public facilities management, etc. [36]. Due to the increasing investment in environmental protection in Heilongjiang Province, this may be related to the increase in the number of sewage treatment plants [28–30], which means that the increase in sewage treatment capacity indirectly affects the quality of surface water in Heilongjiang Province [37]. This is consistent with the research results that local financial environmental protection expenditures have low correlations with permanganate index, chemical oxygen demand, ammonia nitrogen and total phosphorus, but are correlated with the proportion of good water quality. It shows that although the local financial environmental protection expenditure has little effect on the content of pollutants in surface water, it can affect the comprehensive situation of surface water quality in Heilongjiang Province.

Forestry factors are highly correlated with the environmental quality of surface water in Heilongjiang Province. The regulatory effect of forests on surface water is mainly due to the good water storage function and hydrological effect of the forest litter layer and soil layer. These effects will promote the improvement of water environment quality. Although there are certain differences between different forest lands, the regulatory effect of natural mixed forest is better than that of pure forest or artificial afforestation.

As further studies on the impact of land use patterns on surface water systems in recent years suggest that the contribution of forest drainage to surface water eutrophication may be greater than previously estimated [38,39], changes in forest surface runoff and the exposure of understory organic and inorganic layers can affect the concentrations of phosphorus, nitrogen, and dissolved organic carbon in surface waters [40]. The study by Lepistö et al. (2021) shows that the percentage of forest drainage is positively correlated with the total organic nitrogen in forest streams, which in turn is correlated with the total organic carbon concentration [41]. This shows that the impact of forests on the environmental quality of surface water is not only positive, but may have some negative effects, especially on chemical oxygen demand, which is consistent with our findings. In addition, some natural factors such as annual sunshine hours and total surface water supply also have a major impact on the surface water environment, which shows that in addition to human activities, the role of natural conditions cannot be ignored.

## 5. Conclusions

During the "Thirteenth Five-Year Plan" period, the annual average concentration of major pollutants in surface water in Heilongjiang Province has dropped significantly, the proportion of water quality of Class I–III has increased, the proportion of water quality inferior to Class V has decreased, and the overall environmental quality of surface water has improved. Research on the driving factors of water quality change shows that nitrogen and phosphorus pollutants in farmland surface water runoff and domestic sewage are the main sources leading to ammonia nitrogen and total phosphorus pollution in surface water in Heilongjiang Province. The increase of sewage treatment plants has a greater impact on the permanganate index, chemical oxygen demand and the proportion of good water quality, which indirectly affects the overall water quality of Heilongjiang Province. It is worth noting that the impact of forests on the environmental quality of surface water in Heilongjiang Province is complex and may lead to increased chemical oxygen demand in surface water. In addition, the influence of natural factors on the surface water environment, such as the total annual water supply and surface water supply in Rizhao City cannot be ignored.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/w14152367/s1, Table S1: Basic items of surface water environmental quality※ Standard limit. Table S2: Grey correlation degree and ranking of surface water environmental quality comparison series.

**Author Contributions:** Conceptualization, B.L. and W.C.; methodology, C.F.; software, C.F.; resources, W.C.; writing—original draft preparation, C.F.; writing—review and editing, Y.B. and M.Z.; visualization, C.F.; supervision, Y.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** We would like to thank the Heilongjiang Provincial Department of Ecology and Environment for providing data support for this study, and Fengying Zhang from China Environmental Monitoring Station for providing technical guidance for this paper.

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

## References

1. Qiao, M.; Qi, W.; Liu, H.; Qu, J. Oxygenated polycyclic aromatic hydrocarbons in the surface water environment: Occurrence, ecotoxicity, and sources. *Environ. Int.* **2022**, *163*, 107232. [CrossRef] [PubMed]
2. Zou, H.-Y.; He, L.-Y.; Gao, F.-Z.; Zhang, M.; Chen, S.; Wu, D.; Liu, Y.; He, L.; Bai, H.; Ying, G. Antibiotic resistance genes in surface water and groundwater from mining affected environments. *Sci. Total Environ.* **2021**, *772*, 145516. [CrossRef] [PubMed]
3. Henning, N.; Wick, A.; Ternes, T.A. Biotransformation of pregabalin in surface water matrices and the occurrence of transformation products in the aquatic environment—Comparison to the structurally related gabapentin. *Water Res.* **2021**, *203*, 117488. [CrossRef] [PubMed]
4. Galafassi, S.; Nizzetto, L.; Volta, P. Plastic sources: A survey across scientific and grey literature for their inventory and relative contribution to microplastics pollution in natural environments, with an emphasis on surface water. *Sci. Total Environ.* **2019**, *693*, 133499. [CrossRef] [PubMed]
5. Sammut, G.; Sinagra, E.; Helmus, R.; de Voogt, P. Perfluoroalkyl substances in the Maltese environment—(I) surface water and rain water. *Sci. Total Environ.* **2017**, *589*, 182–190. [CrossRef]
6. Sur, K.; Verma, V.K.; Pateriya, B. Surface water estimation at regional scale using hybrid techniques in GEE environment—A case study on Punjab State of India. *Remote Sens. Appl. Soc. Environ.* **2021**, *24*, 100625. [CrossRef]
7. Azadi, S.; Amiri, H.; Mooselu, M.G.; Liltved, H.; Castro-Muñoz, R.; Sun, X.; Boczkaj, G. Network design for surface water quality monitoring in a road construction project using Gamma Test theory. *Water Resour. Ind.* **2021**, *26*, 100162. [CrossRef]
8. Haghnazar, H.; Johannesson, K.H.; González-Pinzón, R.; Pourakbar, M.; Aghayani, E.; Rajabi, A.; Hashemi, A.A. Groundwater geochemistry, quality, and pollution of the largest lake basin in the Middle East: Comparison of PMF and PCA-MLR receptor models and application of the source-oriented HHRA approach. *Chemosphere* **2022**, *288*, 132489. [CrossRef]
9. Abdelaziz, S.; Gad, M.I.; el Tahan, A.H.M.H. Groundwater quality index based on PCA: Wadi El-Natrun, Egypt. *J. Afr. Earth Sci.* **2020**, *172*, 103964. [CrossRef]
10. Abuzaid, A.S.; Jahin, H.S. Combinations of multivariate statistical analysis and analytical hierarchical process for indexing surface water quality under arid conditions. *J. Contam. Hydrol.* **2022**, *248*, 104005. [CrossRef]
11. Elkorashey, R.M. Utilizing chemometric techniques to evaluate water quality spatial and temporal variation. A case study: Bahr El-Baqar drain—Egypt. *Environ. Technol. Innov.* **2022**, *26*, 102332. [CrossRef]
12. Zhan, S.; Zhou, B.; Li, Z.; Li, Z.; Zhang, P. Evaluation of source water quality and the influencing factors: A case study of Macao. *Phys. Chem. Earth Parts A/B/C* **2021**, *123*, 103006. [CrossRef]
13. Wang, S.; Fu, B.; Zhao, W.; Liu, Y.; Wei, F. Structure, function, and dynamic mechanisms of coupled human–natural systems. *Curr. Opin. Environ. Sustain.* **2018**, *33*, 87–91. [CrossRef]
14. Ferro-Azcona, H.; Espinoza-Tenorio, A.; Calderón-Contreras, R.; Ramenzoni, V.C.; de las Mercedes Gómez País, M.; Mesa-Jurado, M.A. Adaptive capacity and social-ecological resilience of coastal areas: A systematic review. *Ocean. Coast. Manag.* **2019**, *173*, 36–51. [CrossRef]
15. Griffin, M.T.; Montz, B.E.; Arrigo, J.S. Evaluating climate change induced water stress: A case study of the Lower Cape Fear basin, NC. *Appl. Geogr.* **2013**, *40*, 115–128. [CrossRef]
16. Deng, J. Introduction to Grey System. *J. Grey Syst.* **1989**, *1*, 1–24.
17. Li, J.; Song, H.; Sun, W.; Sun, P.; Hao, J. Measuring Performance and its influence factors of National Sustainable Development Pilot Zones in Shandong, China. *J. Clean. Prod.* **2020**, *289*, 125620. [CrossRef]
18. Environmental Quality of Heilongjiang Province in 2016. Available online: http://sthj.hlj.gov.cn/hjzlbg/16863.jhtml (accessed on 1 April 2017).
19. Environmental Quality Status of Heilongjiang Province in 2017. Available online: http://sthj.hlj.gov.cn/hjzlbg/16862.jhtml (accessed on 24 May 2018).
20. Environmental Quality Status of Heilongjiang Province in 2018. Available online: http://sthj.hlj.gov.cn/hjzlbg/16876.jhtml (accessed on 1 February 2019).

21. The Quality of Ecological Environment in Heilongjiang Province in 2019. Available online: http://sthj.hlj.gov.cn/hjzlbg/16875.jhtml (accessed on 10 February 2020).
22. Eco-Environmental Quality Status of Heilongjiang Province in 2020. Available online: http://sthj.hlj.gov.cn/hjzlbg/18011.jhtml (accessed on 1 February 2021).
23. 2016 Heilongjiang Province Environmental Status Bulletin. Available online: http://sthj.hlj.gov.cn/hjzlzkgb/19153.jhtml (accessed on 8 August 2017).
24. 2017 Heilongjiang Province Environmental Status Bulletin. Available online: http://sthj.hlj.gov.cn/hjzlzkgb/19154.jhtml (accessed on 5 June 2018).
25. 2018 Heilongjiang Province Ecological Environment Bulletin. Available online: http://sthj.hlj.gov.cn/hjzlzkgb/19155.jhtml (accessed on 4 June 2019).
26. 2019 Heilongjiang Province Ecological Environment Bulletin. Available online: http://sthj.hlj.gov.cn/hjzlzkgb/19156.jhtml (accessed on 3 June 2020).
27. 2020 Heilongjiang Province Ecological Environment Bulletin. Available online: http://sthj.hlj.gov.cn/hjzlzkgb/19492.jhtml (accessed on 4 June 2021).
28. Heilongjiang Province Environmental Statistics Annual Report 2016. Available online: http://sthj.hlj.gov.cn/hjtj/12362.jhtml (accessed on 27 February 2018).
29. Heilongjiang Province Environmental Statistics Annual Report 2017. Available online: http://sthj.hlj.gov.cn/hjtj/12364.jhtml (accessed on 18 March 2019).
30. 2018 Annual Report of Ecological Environment Statistics of Heilongjiang Province. Available online: http://sthj.hlj.gov.cn/hjtj/12366.jhtml (accessed on 19 January 2020).
31. National Bureau of Statistics. *China Statistical Yearbook 2020*; China Statistics Press: Beijing, China, 2021.
32. National Bureau of Statistics, Ministry of Ecology and Environment. *2019 China Environmental Statistical Yearbook*; China Statistics Press: Beijing, China, 2021.
33. 2021 Government Work Report. Available online: https://www.hlj.gov.cn/n200/2021/0224/c68-11014966.html (accessed on 24 February 2021).
34. Komariah, I.; Matsumoto, T. Application of Hydrological Method for Sustainable Water Management in the Upper-Middle Ciliwung (UMC) River Basin, Indonesia. *J. Water Environ. Technol.* **2019**, *17*, 203–217. [CrossRef]
35. Mokarram, M.; Saber, A.; Sheykhi, V. Effects of heavy metal contamination on river water quality due to release of industrial effluents. *J. Clean. Prod.* **2020**, *277*, 123380. [CrossRef]
36. Muhammad, S.; Pan, Y.; Agha, M.H.; Umar, M.; Chen, S. Industrial structure, energy intensity and environmental efficiency across developed and developing economies: The intermediary role of primary, secondary and tertiary industry. *Energy* **2022**, *247*, 123576. [CrossRef]
37. Masuda, S.; Sato, T.; Mishima, I.; Maruo, C.; Yamazaki, H.; Nishimura, O. Impact of nitrogen compound variability of sewage treated water on N2O production in riverbeds. *J. Environ. Manag.* **2021**, *290*, 112621. [CrossRef]
38. Nieminen, M.; Sallantaus, T.; Ukonmaanaho, L.; Nieminen, T.M.; Sarkkola, S. Nitrogen and phosphorus concentrations in discharge from drained peatland forests are increasing. *Sci. Total Environ.* **2017**, *609*, 974–981. [CrossRef]
39. Leena, F.; Ahti, L.; Kristian, K.; Antti, R.; Laura, H.; Markus, H.; Samuli, J.; Pirkko, K.; Tuija, M.; Sirpa, P.; et al. Drainage for forestry increases N, P and TOC export to boreal surface waters. *Sci. Total Environ.* **2021**, *762*, 144098. [CrossRef]
40. Menberu, M.W.; Marttila, H.; Tahvanainen, T.; Kotiaho, J.S.; Hokkanen, R.; Kløve, B.; Ronkanen, A. Changes in pore water quality after peatland restoration: Assessment of a large-scale, replicated before-after-control-impact study in Finland. *Water Resour. Res.* **2017**, *53*, 8327–8343. [CrossRef]
41. Lepistö, A.; Räike, A.; Sallantaus, T.; Finér, L. Increases in organic carbon and nitrogen concentrations in boreal forested catchments—Changes driven by climate and deposition. *Sci. Total Environ.* **2021**, *780*, 146627. [CrossRef]

water

MDPI

*Article*

# Long-Term Study of Monitoring History and Change Trends in Surface Water Quality in China

**Fengying Zhang [1,2,†], Lanyu Lin [1,†], Wenpan Li [1], Dekun Fang [1], Zhuo Lv [1], Mingsheng Li [1], Guangwen Ma [1], Yeyao Wang [1], Li Wang [3,*] and Lihuan He [1,*]**

[1] China National Environmental Monitoring Centre, Beijing 100012, China; zhangfy@cnemc.cn (F.Z.); linly@cnemc.cn (L.L.); liwp@cnemc.cn (W.L.); fangdk@cnemc.cn (D.F.); lvzhuo@cnemc.cn (Z.L.); lims@cnemc.cn (M.L.); magw@cnemc.cn (G.M.); wangyy@cnemc.cn (Y.W.)
[2] Department of Health, Ethics & Society, CAPHRI Care and Public Health Research Institute, Faculty of Health, Medicine and Life Sciences, Maastricht University, 6200 MD Maastricht, The Netherlands
[3] Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
* Correspondence: wangli@igsnrr.ac.cn (L.W.); helh@cnemc.cn (L.H.)
† These authors contributed equally to this work.

**Abstract:** To investigate the monitoring history and long-term change trends in surface water quality in China since the reform and opening up, the history of surface water environment monitoring is summarized, including monitoring scope, monitoring methods, and technical requirements. Temporal and spatial patterns of surface water quality in China were analyzed based on the monitoring results. In the past 40 years, the monitoring targets for surface water quality have been continuously improved, the frequency of monitoring has become more science-based, and the monitoring indicators are now comprehensive. Overall, the temporal change trend in surface water quality has followed a "fluctuating changes stage—rapid deterioration stage—fluctuations stalemate stage—rapid improvement stage" pattern. However, the current regional surface water quality is still in a polluted status, and there is a gap between surface water quality status and the goal of building a well-off society. At present, China's surface water pollution is prone to high numbers of incidents and the treatment of surface water pollution has entered a crucial stage. The potential for the continuous reduction of major pollutant discharges has become more challenging, and the marginal cost for pollution control has increased. It is very difficult to comprehensively solve the outstanding water environment problems. In addition to strengthening the existing work on surface water quality control, it is also necessary to strengthen the work of risk identification, early warning, and regulation implementation of the surface water environment. During the 14th year plan period (2021–2025), the overall planning on water resources, water ecology, and water quality will be implemented, and beautiful rivers and lakes will be created.

**Keywords:** surface water; monitoring history; change trends in surface water quality; water quality protection

**Citation:** Zhang, F.; Lin, L.; Li, W.; Fang, D.; Lv, Z.; Li, M.; Ma, G.; Wang, Y.; Wang, L.; He, L. Long-Term Study of Monitoring History and Change Trends in Surface Water Quality in China. *Water* **2022**, *14*, 2134. https://doi.org/10.3390/w14132134

Academic Editors: George Arhonditsis and Danny D. Reible

Received: 21 April 2022
Accepted: 28 June 2022
Published: 4 July 2022



## 1. Introduction

Environmental monitoring is an important cornerstone of environmental protection and an important support for the construction of an ecological civilization and beautiful China [1–3]. Water quality monitoring is an important branch of environmental monitoring [2], which refers to the process of sampling and measuring various characteristic indexes of water to grasp the water environment quality status and the dynamic changes of pollutants in the water system, as well as to record the process [2]. Monitoring is the basis for water pollution control, environmental management, and scientific research. Through water quality monitoring, we can master the dynamic changes in the water environment and provide first-hand scientific data to support decisions regarding the prevention and

control of water pollution and the formulation of environmental protection policies and environmental legislation [2,4].

Many scholars have conducted a great deal of water quality research in China [5–7], but most studies conducted to date have focused on one aspect [8,9] or on a single river basin or lake [10–13]. Studies focused on long-term sequences across the country are rare [12]. The 40th anniversary of China's reform and opening up was in 2018; accordingly, it is important to summarize the history of water quality monitoring and water environmental protection during the past 40 years, as well as to analyze the trends in surface water quality during this time, which could provide support for precise pollution control and environmental management [11,12].

In this study, the history of surface water quality monitoring and surface water environmental protection is summarized, temporal and spatial variations in surface water quality are analyzed, and the current existing problems and pressure on surface water quality are proposed. The analyses conducted in this study are based on the Eco-Environmental Quality Report of China from 1980 to 2020 [14], the Report on the State of the Eco-environment in China from 1989 to 2020 [15], and other eco-environmental quality reports, related policy norms, and data from the government, combined with water quality monitoring data. The results presented herein will provide a foundation and scientific research support that will facilitate pollution prevention and control and enable the realization of an ecologically friendly civilization in China.

## 2. Materials and Methods

Due to the availability and integrity of historical monitoring data, our study area focuses on mainland China. All data used in this study were derived from environmental reports, statistical year books, government reports, relevant literature, or professional websites. Environmental reports included the China Eco-Environmental Quality Report (1980–2020) [14], Report on the State of Ecology and the Environment (1989–2020) [15], and the Annual Statistics Report on the Environment in China (1998–2015) [16]. Statistical yearbooks and government reports investigated included the China Statistical Yearbook [17], China Environmental Yearbook [18], government work reports, etc. Relevant literature and professional websites, such as academic literature from the China National Knowledge Internet (CNKI), Elsevier, and relevant data, policies, specifications, and systems published on governmental networks by the Ministry of Ecology and Environment (MEE), China National Environmental Monitoring Center (CNEMC), and the provincial-/city-level Ecology Environment Agency were also investigated.

The spatial distribution of surface water quality monitoring sites and surface water quality were evaluated by ArcGIS 10.0 with a license from the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, and temporal patterns in water quality and pollutants were summarized by Origin 2018.

## 3. Results and Discussion

### *3.1. History of Surface Water Quality Monitoring*

Surface water quality monitoring in China began in the 1980s. Although this was much later than developed countries, there has been considerable progress in the past 40 years. Currently, China's water environment monitoring technology covers large rivers, lakes, and other areas such as reservoirs, etc. Moreover, the monitoring techniques and means have been improved year by year.

China's water quality monitoring is conducted in a radiation mode. Specifically, the water quality monitoring center is the core, and the monitoring points are used as nodes to form the national surface water quality monitoring network, forming an integrated monitoring network covering the national-, provincial-, municipal-, and county-level surface water quality monitoring, and can meet the needs of water quality monitoring in different areas and different regions. At the same time, a combination of fixed-point sampling and mobile sampling is adopted to ensure the real-time accuracy of monitoring data.

3.1.1. Monitoring Scope

In 1988, the former National Environmental Protection Agency (NEPA) first established a national surface water quality monitoring network consisting of 353 sections. In 1993, the NEPA re-examined and certified the monitoring sections in the national control monitoring network and confirmed that the national surface water quality monitoring network consisted of 313 national control sections [14,15].

In 2003, the NEPA further adjusted the national environmental monitoring network and monitoring sections and identified the national surface water quality monitoring network, which includes the Yangtze River, Yellow River, Haihe River, Liaohe River, Songhua River, Pearl River, and Huaihe River, the "Three Lakes" (Taihu Lake, Dianchi Lake, and Chaohu Lake), and the three regional rivers (rivers in Zhejiang and Fujian Province, rivers in northwestern China, and rivers in southwestern China). Overall, the monitoring network has 759 monitoring sections (604 rivers and 155 lakes) covering 320 rivers and 28 lakes [14,15].

In 2012, the NEPA released a new national environmental monitoring basin network consisting of 972 monitoring sections covering 420 rivers and 62 lakes. Water monitoring sections were set in the main stream of China's main water systems, the primary and secondary tributaries with an annual runoff of more than 500 million cubic meters, national border rivers and provincial rivers with an annual runoff of more than 300 million cubic meters, and large water conservancy facilities [14,15].

In July of 2015, the State Council issued the "Eco-Environmental Monitoring Network Construction Plan," which clearly defined the national surface water monitoring network during the "13th Five-Year Plan". The number of national control sections was increased from 972 in the "12th Five-Year Plan" to 1589 in 2015, including 1246 sections in 1366 rivers and 343 sections in 139 lakes. These sections (Figure 1) include 1940 assessment and ranking sections, 195 control sections for evaluation in the estuaries (of which, 85 are evaluation, assessment, and ranking sections), and 717 sections for scientific research. The "13th Five-Year" National Surface Water Quality Monitoring Network covers both the main rivers of the country and important primary and secondary tributaries, as well as the third and fourth tributaries of the key areas, key lakes/reservoirs, etc. Therefore, it has good regional spatial representation and can comprehensively, accurately, and objectively reflect the water quality and temporal–spatial distribution characteristics of pollutants in the water system or region [14,15].

In summary, China's national surface water quality monitoring is an organic whole based on monitoring of surface water quality, adherence to water and land planning, land and sea planning, river and lake planning, upper and lower planning, urban and rural planning, and comprehensive monitoring of the national aquatic environment ecosystem.

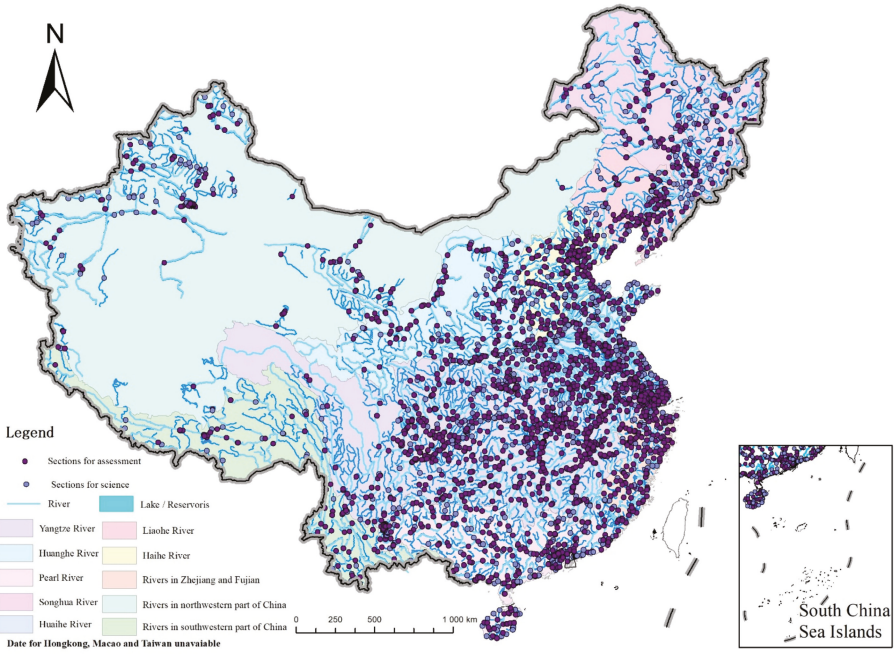



**Figure 1.** China national surface water quality monitoring network in 2020.

3.1.2. Monitoring Indicators

The monitoring indicators during the past 40 years were listed in Appendix A. Before 2011, there were 11 monthly monitoring indicators pertaining to water quality in rivers (water temperature, pH, conductivity, dissolved oxygen, permanganate index, five-day biochemical oxygen demand, ammonia nitrogen, petroleum, volatile phenols, mercury, and lead). When monitoring the water quality in lakes and reservoirs, total phosphorus, total nitrogen, chlorophyll a, transparency, and water level were also included [14,15].

After the Measures for the Evaluation of Surface Water Quality was issued in 2011, the monthly monitoring was conducted in accordance with the 24 indicators, as shown in the Surface Water Quality Standards (GB3838-2002) [14,15] (Appendix B).

3.1.3. Monitoring Frequency

The monitoring frequency of surface water quality has increased obviously [14,15]. Before 2003, the monitoring frequency of surface water quality was generally low (about six times per year), and it was monitored according to the water periods, including dryness, flatness, and abundance.

Monthly monitoring has been conducted since the establishment of the monthly reporting mechanism based on the national water quality monitoring system in 2003. Monitoring is conducted for the first 10 days of the month.

Since October 2017, the sampling and laboratory analysis separation mode has been fully implemented, the monitoring frequency has been increased, and the monitoring work now is being conducted quarterly, monthly, weekly, daily, and, even, one time per four hours.

3.1.4. Monitoring Method

Sampling for surface water quality monitoring is mainly conducted manually. Since the beginning of the 21st Century, the state has built 150 automatic surface water quality monitoring stations in the provincial boundary sections of major rivers and important

border rivers to provide early warnings regarding water quality. From July 1st 2009, real-time water quality data from national water quality automatic monitoring stations have been released to the public and published online [14,15].

Since October 2017, the 1940 national surface water assessment sections have fully implemented the sampling and laboratory analysis separation mode. By the end of July 2018, 2050 automatic surface water quality monitoring stations were built. Future water quality monitoring will be based on automatic monitoring supplemented by manual monitoring, leading to comprehensive realization of sampling and laboratory analysis separation, automatic monitoring, and data sharing [14,15].

#### 3.1.5. Surface Water Quality Standard

Over the past 40 years, China's surface water environmental quality standards have undergone four major changes [19]. In 1983, the Environmental Quality Standard for Surface Water (GB 3838-83) was promulgated and implemented for the first time. In 1988, the Environmental Quality Standard for Surface Water was revised to the Environmental Quality Standard for Surface Water (GB 3838-88). In 1999, this was revised to the Environmental Quality Standard for Surface Water (GHZB 1-1999) and, in 2002 to the Environmental Quality Standard for Surface Water (GB 3838-2002) [14,15].

### *3.2. Spatial Temporal Trends in Surface Water Quality*

#### 3.2.1. General Temporal Change Trends

Figure 2 showed the general temporal change trends of surface water quality. In the past four decades, the national surface water quality has shown a trend of fluctuating changes–rapid deterioration–volatility–rapid improvement.

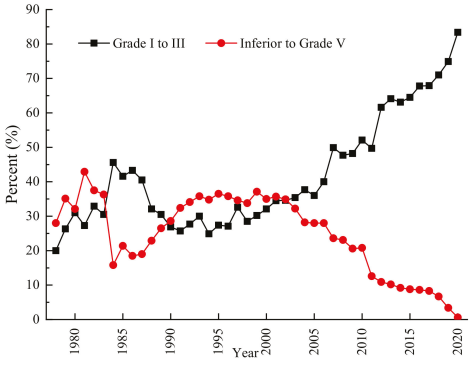



**Figure 2.** Annual ratio of water quality in China during 1978–2020.

The period of fluctuating changes was from 1978 to 1983. During this period, the annual ratios of grade I–III sections (sections with water quality between grade I and grade III) ranged from 20.0% to 30.5% and the annual proportions of inferior grade V sections (sections in which water quality failed to meet grade V) were from 28.0% to 36.3%. In this period, the overall situation of China's surface water quality transitioned from basically clean to partially deteriorating.

The stage of rapid deterioration was from 1984 to 1990. During this time, the annual ratio of grade I–III sections decreased by 18.7% from 45.6% in 1984 to 26.9% in 1990, while the proportion of inferior grade V sections increased by 12.8%, from 15.8% in 1984 to 28.6% in 1990. This stage corresponded to the beginning of reform and opening up, accompanied with fast economic and social development. The eastern part of China, including the Yangtze River Delta and the Pearl River Delta region, had begun to undergo rapid development, foreign enterprises gradually moved in, and local enterprises developed everywhere. The environmental effects caused by the rapid development of industry were

gradually emerging, especially in the 7th Five-Year Period (1986–1990), and pressure on the surface water quality in China began to increase. In this period, although the status and role of environmental protection in social and economic development were clarified, the relationship between economic construction and environmental protection had not been rationalized. As a result, the surface water quality generally evolved from partial deterioration to general deterioration.

The phase of volatility was from 1991 to 2001. During this period, the annual ratio of sections that met the water quality standard (water quality in grade I to grade III, grades I–III for short) was from 25.7% to 34.5%, and the proportion of inferior grade V sections was from 32.4% to 35.7%.

The rapid improvement period was from 2002 to 2020, during which time, the annual ratio of grade I–III sections increased by 48.8%, from 34.6% in 2002 to 83.4% in 2020, while the proportion of inferior grade V sections decreased by 34.3% from 34.9% in 2002 to 0.6% in 2020. During the "10th Five-Year Plan" period, water pollution in the country was initially curbed and environmental quality improved in some areas. These improvements were mainly attributable to the initial recognition of the relationship between economic development and the implementation of water environmental protection, strict industrial structure adjustment policies, urban sewage centralized treatment, total control, and key river basin water pollution prevention and control planning systems [20].

During the "11th Five-Year Plan" period, chemical oxygen demand and ammonia nitrogen were introduced as binding indicators in the environmental protection target, and the implementation of the environmental protection target responsibility system greatly improved the construction of pollution control facilities and promoted the improvement of the level of the conventional pollution indicators. During this period, the overall surface water quality of the national water environment was stable, and the water quality improved. The water quality of the main stream in the key river basin was obviously improved. The concentration of the main pollution indicators of the tributaries dropped drastically, and water pollution prevention and control work in the basin made remarkable achievement.

#### 3.2.2. Trends for Major Pollution Indicators

Among the major pollution indicators, the concentrations of permanganate and ammonia nitrogen showed similar temporal patterns (Figure 3), first increasing, then decreasing.

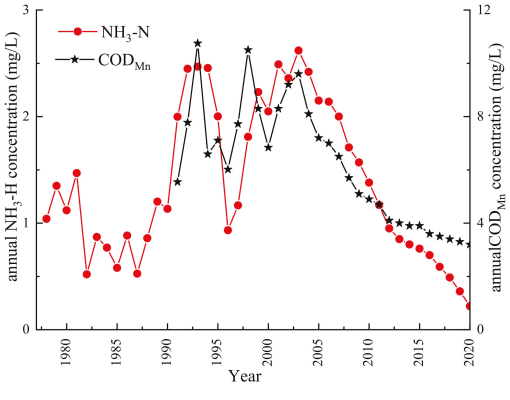



**Figure 3.** The annual concentration of $NH_3$-N and $COD_{Mn}$ during 1978–2020.

Among these, the annual ammonia nitrogen ($NH_3$-N for short) concentration showed an increasing trend from 1978–1981, followed by a sharp decrease in 1982, then fluctuations in 1983 and 1986 and an obviously increasing trend during 1987 and 1995, which increased from the lowest value of 0.53 mg/L to the highest value of 2.00 mg/L. From 1996 to 2002,

the annual concentration fluctuated from 0.93 mg/L to 2.36 mg/L, while it improved from 2003 to 2020, by a decrease in annual concentration from 2.62 mg/L to 0.22 mg/L.

The annual permanganate index ($COD_{Mn}$ for short) concentration fluctuated from 1991 to 2002, with annual levels ranging from 5.5 mg/L to 9.2 mg/L, while the concentration improved from 2003 to 2020 when annual concentrations decreased from 9.6 mg/L to 3.2 mg/L.

#### 3.2.3. Spatial Temporal Changes in the Seven Major River Basins

In general, the water quality in the Yangtze River and Pearl River was good comparing with the rest of the basins. Haihe River had the worst water quality, and it was also the only water basin with grade V sections. The water quality rankings in 2020 were Yangtze River > Pearl River > Yellow River > Songhua River > Huaihe River > Liaohe River > Haihe River.

Figure 4 showed the changes of water quality in the major river basins during 2007–2020. During 2007 and 2020, the water quality in Yangtze River experienced an improvement–fluctuating changes–improvement pattern. The Yellow River, Pearl River. and Huaihe River have experienced improvement–deterioration–improvement changing patterns. The Songhua River, Haihe River and Liaohe River have showed fluctuating improvement. Since 2017, the water quality in the seven major river basins showed a rapid improvement, with an obvious increase for the ratio of grade I to III sections and a decreasing trend for the ratio of grade V sections.

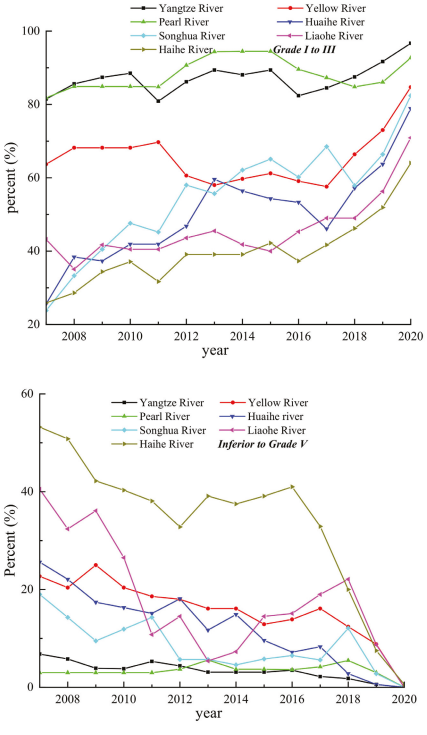



**Figure 4.** Changes of water quality in the major river basins during 2007–2020.

### *3.3. Surface Water Pollution Prevention and Control in China*

Analysis of the correlation between river environmental quality and pollutant emissions, national economic development level, and environmental pollution control investment revealed that the reduction of ammonia nitrogen and COD emissions [20] and the

increase in GDP and environmental pollution control investment are the main reasons for the improvements that have occurred in surface water quality [21,22].

China's surface water environmental protection work has been developed from scratch and expanded in scale [23]. During the past four decades, the intensity of surface water pollution control has increased, and the surface water ecological protection work has been strengthened, which has greatly improved the surface water environment in China [24].

Overall, in the past 40 years, China's environmental protection has experienced four development stages consisting of the initial "three wastes" governance, steady development of pollution prevention, a total quality control stage, and now, a new stage of environmental protection with improvement of environmental quality as the primary goal [25]. Since the founding of the country till the end of the 1970s, during the initial stage of development, China's surface water protection work has been in a disorderly status. The first national environmental protection work conference held in 1973 marked the beginning of water environment protection in China.

In the steady development stage from the 1980s to the early 21st Century, China's economy developed rapidly, while its water environmental protection work had achieved remarkable progress. During the "9th Five-Year Plan" period, the "Huaihe River Basin Water Pollution Prevention and Control Plan" was approved and implemented, marking China's entry into a stage of large-scale water pollution control. During this period, the country launched a comprehensive campaign targeting water pollution control of the Three Rivers and Three Lakes (Huaihe River, Haihe River, and Liaohe River and Taihu Lake, Dianchi Lake, and Chaohu Lake, respectively). Since then, organic matter pollution of surface water and the eutrophication of lakes have been effectively curbed, and the water environmental protection work has made significant progress. The total control led to a quality improvement from 2002 to 2014. In 2002, China issued the "Clean Production Promotion Law", which officially marked the transition from end-of-pipe treatment to total process control for China's water environment pollution control, and China's river environmental protection work entered a new stage.

In the same year, the MEP promulgated a new "Environmental Quality Standard for Surface Water" (GB3838-2002), among which the water quality evaluation indicator increased from 75 to 109, which was more in line with the status quo of surface water environmental protection in China. Since the "11th Five-Year Plan", target accountability for water quality control was introduced with more stringent and clear targets. The capacity of water quality control improved since then, and the water quality, in particular for the main stream in key river basins, improved significantly.

In the new stage of environmental protection with the improvement of environmental quality as the core (since 2015), the "13th Five-Year Plan" has entered a stage of double constraint on quality improvement and total control. This was marked by the 18th National Congress of the Communist Party of China (CPC). The ecological civilization's construction is included in the overall layout of the "five in one" plan, and water quality has become an important part of improving people's livelihood and building a well-off society. As a program of action for the prevention and control of water pollution in China, the "Water pollution prevention action plan" clarifies the quality improvement objectives of various water bodies in different periods, which was, by 2020, for the seven major river basins, the water quality with a good (grades I–III) ratio being above 70%, and by 2030, the ratio should reach 75%. The "13th Five-Year Plan for National Economic and Social Development Outline" also clarifies the dual-binding indicators for water quality improvement and total control, while the "13th Five-Year Plan for Ecological Environmental Protection" further proposes specific requirements for systematic treatment with quality improvement as the core.

The theory of harmonious coexistence between man and nature was introduced for development in China during the "13th Five-Year Plan", and the Yangtze River economic belt development and high-quality development and ecological protection for the Yellow River Basin were introduced. For the Yangtze River, the eco-environment restoration

was put in a stronger position compared to development, and for the Yellow River, a comprehensive management of the eco-environment system, including the environment quality of mountains, rivers, forests, fields, lakes, and grasslands, and the pollution sources was planned, to promote high-quality development.

## 4. Problems and Pressures

Although water quality in China has improved significantly and an environmental turning point has begun to appear, water quality in some areas is still not improving. As a result, there is still a gap between the quality of the water environment and the goal of building a well-off society [7,10,26,27].

### *4.1. Surface Water Is Still Polluted*

In 2020, the national surface water quality was fairly good. The main pollution indicators were chemical oxygen demand (COD), total phosphorus (TP), and the permanganate index, and the over-standard rates were 9.4%, 7.5%, and 5.8%, respectively. The proportion of the grade I–III section was 83.4%, which indicates that surface water quality in 16.6% of the sections was worse than grade III, while the proportion of sections that were worse than grade V was 0.6%. The Liaohe River and Haihe River were slightly polluted [14,15].

### *4.2. Total Phosphorus Pollution Has Increased*

In 2020, the rate of sections with TP exceeding the national surface water standard was 7.5%, more than that of $COD_{Mn}$, and became the second over-standard indicator affecting the national surface water quality. The TP concentration in 20 lakes and one reservoir exceeded the standard; the TP pollution in the lake was higher than in the reservoir [14,15].

### *4.3. High Pollutant Emission Intensity and Increasing Pollution from Residential Sources*

China's primary source of water pollution has changed from industrial pollution to domestic pollution, and there is now great pressure to reduce domestic sewage. With increased urbanization, the contribution of domestic pollution sources has become increasingly prominent, and it is now the main source of water pollution. As a result, pollution control of rural life at the township level needs to be strengthened. In 2015, the total discharge of wastewater was 73.53 billion tons in China, which was 2.7% greater than in 2014. However, industrial wastewater discharge was 19.95 billion tons, which was 2.8% lower than in the previous year, accounting for 27.1% of the total wastewater discharge. Domestic sewage discharge was 53.52 billion tons, which represented an increase of 4.9% over 2014 and accounted for 72.8% of the total wastewater discharge. Finally, the discharge of wastewater from centralized pollution control facilities was about 0.6 billion tons, accounting for 0.1% of the total wastewater discharge [14,15].

### *4.4. The Problem of Lake Eutrophication Is Obvious*

In 2020, 76.8% of the 110 major lakes and reservoirs across the country met the grade I–III standard, while 15.2% and 2.7% met the grade IV and V standards, respectively, and 5.4% failed to meet the grade V standard. The main pollution indicators were TP, COD, and $COD_{Mn}$. Of the 110 lakes (reservoirs) being monitored for nutritional status, 10 were oligotrophic, 68 were mesotrophic, 26 were under slight eutrophication, six were under intermediate eutrophication, and one was under heavy eutrophication [14,15].

### *4.5. Heavy Metal Levels in Surface Water Have Exceeded the Standard*

In 2019, there were 28 cases of heavy metals exceeding the standard in 23 surface water sections. The main indicators exceeding the standards were mercury, arsenic, and selenium. Among these, 12 cases exceeded the mercury limit, seven exceeded the arsenic limit, and three exceeded the selenium limit. From the perspective of the river basin, the over-standard sections were mainly distributed in the Yellow River, Haihe River, Yangtze River, and the Pearl River Basin, with 8 cases, 6 cases, 5 cases, and 4 case of heavy metals

exceeding the standard, respectively. At the provincial level, the sections exceeding the standard are mainly distributed in Hebei, Inner Mongolia, Hubei, Hunan, Shanxi, and other areas.

## 5. Conclusions

In the past 40 years, China's environmental monitoring has made great progress. The monitoring of water quality is developing towards the continuous improvement of monitoring targets, the frequency of monitoring is becoming more reasonable, the monitoring indicators are being comprehensively covered, the technical methods are becoming scientific, and the quality standards are becoming increasingly stringent.

During this same period, the national surface water quality has shown a trend of fluctuating changes–rapid deterioration–volatility–rapid improvement. From 2002 to 2020, the water quality in the main watersheds was generally good, showing a gradual improvement trend, and the pollution was reduced. The main change has been that the proportion of water quality of grade I–III increased annually, while the proportion of inferior water quality of grade V has decreased. During 2002–2020, the proportion of grade I–III water sections increased from 34.6% to 83.4%, while the proportion of inferior grade V sections decreased from 34.9% to 0.6%. Furthermore, the water quality in the main river basins has improved.

The increase in GDP and environmental pollution control investment are the main reasons China's surface water environmental protection system has been developed from scratch and expanded, the intensity of surface water pollution control has been strengthened, and the surface water ecological protection work has increased, which has greatly improved the surface water environment in China.

However, with the current rapid development of the economy and high consumption of materials and energy, the number of surface water pollution incidents in China has been high, and the treatment of surface water environmental pollution has entered a crucial stage. Moreover, the potential for the continuous reduction of total pollutant discharges has been narrowing, while the marginal costs of governance are continually increasing. Accordingly, it has become very difficult to achieve comprehensive improvement of the water environment quality by focusing on solving the outstanding water environment problems in a short period of time. Therefore, it is necessary to improve risk identification, early warning systems, and regulation of surface water environments.

In the year 2022, the Chinese government released the "14th five-year plan for ecological environment protection", which enhances the regulations on water quality control. In this plan, specific measures are outlined for water quality monitoring, water ecology monitoring, and water pollution source monitoring. For water quality monitoring, a national water monitoring network will cover key basins and prefecture-level cities, supported mainly by automatic monitoring. The local water monitoring network should cover the major water bodies, major cities and towns, large industrial clusters, planting and breeding areas, and key boundaries. For water ecology monitoring, a water ecology monitoring framework should be established and implemented in the Yangtze River Basin. For water pollution source monitoring, the "water cross section-water quality-pollution source" monitoring traceability technology and three-dimensional monitoring network for surface sources should be established and forecast and warning capacity on water environment should be achieved.

**Author Contributions:** F.Z.: Conceptualization, Software, Formal analysis, Writing—Original Draft, Review & Editing, Supervision, L.L.: Conceptualization, Software, Formal analysis, Writing—Original Draft, W.L.: Resources, Data Curation, D.F.: Resources, Data Curation, Z.L.: Data Curation, Formal analysis, M.L.: Resources, Data Curation, Visualization, G.M.: Resources, Data Curation, Y.W.: Resources, Supervision, L.W.: Conceptualization, Writing—Review & Editing, Supervision, Funding acquisition, L.H.: Conceptualization, Writing—Review & Editing, Supervision. All authors have read and agreed to the published version of the manuscript.

**Funding:** The present study was funded by the Science and Technology Project of Beautiful China Ecological Civilization Construction (No. XDA23100400) and the National Natural Science Foundation of China (No. U2243206 and No. 42007414).

**Data Availability Statement:** All sources of data are cited throughout the paper.

**Acknowledgments:** The authors wish to thank all the staff members at the China National Environmental Monitoring Center for their strong support of this study.

**Conflicts of Interest:** All authors declare that they have no conflicts of interest to disclose in the context of this study.

## Appendix A. Monitoring Indicators during the Past 40 Years

| Year / Items | Since 1983 | Since 1988 | Since 1999 | Since 2002 |
|---|---|---|---|---|
| | **GB3838-83** | **GB3838-88** | **GHZB 1-1999** | **GB3838-2002** |
| 1 | pH | pH | pH | pH |
| 2 | DO | DO | DO | DO |
| 3 | $BOD_5$ | $BOD_5$ | $BOD_5$ | $BOD_5$ |
| 4 | COD | CODMn | CODMn | CODMn |
| 5 | $NH_3$-N | / | $NH_3$-N | $NH_3$-N |
| 6 | $NO_2$-N | Soluble iron | Soluble iron | COD |
| 7 | Volatile phenols | Volatile phenols | Volatile phenols | Volatile phenols |
| 8 | CN- | / | / | / |
| 9 | As | As | As | As |
| 10 | Hg | Hg | Hg | Hg |
| 11 | Cr6+ | Cr | Cr | Cr |
| 12 | Pb | Pb | Pb | Pb |
| 13 | Cd | Cd | Cd | Cd |
| 14 | Cu | Cu | Cu | Cu |
| 15 | Petroleum | Petroleum | Petroleum | Petroleum |
| 16 | Fecal coliform | Fecal coliform | Fecal coliform | Fecal coliform |
| 17 | | TP | TP | TP |
| 18 | | Zn | Zn | Zn |
| 19 | | nitrite | nitrite | TN |
| 20 | | Anionic surfactant | Anionic surfactant | Anionic surfactant |
| 21 | | Cyanide | Cyanide | Cyanide |
| 22 | | Se | Se | Se |
| 23 | | Sulfide | Sulfide | Sulfide |
| 24 | | temperature | temperature | temperature |
| 25 | | $COD_{Cr}$ | $COD_{Cr}$ | / |
| 26 | | Kjeldahl nitrogen | Kjeldahl nitrogen | / |
| 27 | | Mn | Mn | / |
| 28 | | Chloride | Chloride | / |
| 29 | | un-ionized ammonia | un-ionized ammonia | / |
| 30 | | Total fluoride | | $F^-$ |
| 31 | | Benzo (a) pyrene | total cyanide | / |

**Appendix B. Monitoring Indicators for GB3838-2002**

| Indicator | Grade I | Grade II | Grade III | Grade IV | Grade V |
|---|---|---|---|---|---|
| temperature | The artificial change in water temperature shall be limited to a weekly average maximum temperature rise ≤ 1 and a weekly average maximum temperature drop ≤ 2 | | | | |
| pH | 6~9 | | | | |
| DO≥ | 7.5 | 6 | 5 | 3 | 2 |
| $COD_{Mn}$ (mg/L)≤ | 2 | 4 | 6 | 10 | 15 |
| COD (mg/L)≤ | 15 | 15 | 20 | 30 | 40 |
| $BOD_5$ (mg/L)≤ | 3 | 3 | 4 | 6 | 10 |
| $NH_3$-N (mg/L)≤ | 0.15 | 0.5 | 1 | 1.5 | 2 |
| TP (mg/L)≤ | 0.02 | 0.1 | 0.2 | 0.3 | 0.4 |
| TN (mg/L)≤ | 0.2 | 0.5 | 1 | 1.5 | 2 |
| Cu (mg/L)≤ | 0.01 | 1 | 1 | 1 | 1 |
| Zn (mg/L)≤ | 0.05 | 1 | 1 | 2 | 2 |
| $F^-$ (mg/L)≤ | 1 | 1 | 1 | 1.5 | 1.5 |
| Se (mg/L)≤ | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 |
| As (mg/L)≤ | 0.05 | 0.05 | 0.05 | 0.1 | 0.1 |
| Hg (mg/L)≤ | 0.00005 | 0.00005 | 0.0001 | 0.001 | 0.001 |
| Cd (mg/L)≤ | 0.001 | 0.005 | 0.005 | 0.005 | 0.01 |
| Cr (mg/L)≤ | 0.01 | 0.05 | 0.05 | 0.05 | 0.1 |
| Pb (mg/L)≤ | 0.01 | 0.05 | 0.05 | 0.05 | 0.1 |
| Cyanide (mg/L)≤ | 0.005 | 0.05 | 0.2 | 0.2 | 0.2 |
| Volatile phenol (mg/L)≤ | 0.002 | 0.002 | 0.005 | 0.01 | 0.1 |
| Petroleum (mg/L)≤ | 0.05 | 0.05 | 0.05 | 0.5 | 1 |
| Anionic surfactant (mg/L)≤ | 0.2 | 0.2 | 0.2 | 0.3 | 0.3 |
| Sulfide (mg/L)≤ | 0.05 | 0.1 | 0.2 | 0.5 | 1 |
| Fecal coliform (number/L)≤ | 200 | 2000 | 10,000 | 20,000 | 40,000 |

## References


1. Bai, Y.; Niu, H.; Wen, X. The strategy study on the development and innovations of Chinese environmental monitoring: A comparison study. *Procedia Environ. Sci.* **2012**, *13*, 2458–2463. [CrossRef]
2. Li, D.; Liu, S. Chapter 1—Sensors in Water Quality Monitoring, Chapter 3—System and Platform for Water Quality Monitoring. In *Water Quality Monitoring and Management*; Li, D., Liu, S., Eds.; Academic Press: Beijing, China, 2019; pp. 1–112.
3. Sun, A.Y.; Zhong, Z.; Jeong, H.; Yang, Q. Building complex event processing capability for intelligent environmental monitoring. *Environ. Modell. Softw.* **2019**, *116*, 1–6. [CrossRef]
4. Wang, X.; Berman, E.M.; Chen, D.-Y.; Niu, X. Strategies to improve environmental networks for pollution control: Evidence from eco-compensation programs in China. *J. Environ. Manag.* **2019**, *234*, 387–395. [CrossRef]
5. Zhang, C.; Tian, D.; Yi, X.; Zhang, T.; Ruan, J.; Wu, R.; Chen, C.; Huang, M.; Ying, G. Occurrence, distribution and seasonal variation of five neonicotinoid insecticides in surface water and sediment of the Pearl Rivers, South China. *Chemosphere* **2019**, *217*, 437–446. [CrossRef]
6. Zhang, J.; Li, S.; Dong, R.; Jiang, C.; Ni, M. Influences of land use metrics at multi-spatial scales on seasonal water quality: A case study of river systems in the Three Gorges Reservoir Area, China. *J. Clean. Prod.* **2019**, *206*, 76–85. [CrossRef]
7. Zhu, Y.; Price, O.R.; Kilgallon, J.; Qi, Y.; Tao, S.; Jones, K.C.; Sweetman, A.J. Drivers of contaminant levels in surface water of China during 2000–2030: Relative importance for illustrative home and personal care product chemicals. *Environ. Int.* **2018**, *115*, 161–169. [CrossRef]
8. Tian, Y.; Jiang, Y.; Liu, Q.; Dong, M.; Xu, D.; Liu, Y.; Xu, X. Using a water quality index to assess the water quality of the upper and middle streams of the Luanhe River, northern China. *Sci. Total Environ.* **2019**, *667*, 142–151. [CrossRef]

9. Wang, J.; Fu, Z.; Qiao, H.; Liu, F. Assessment of eutrophication and water quality in the estuarine area of Lake Wuli, Lake Taihu, China. *Sci. Total Environ.* **2019**, *650*, 1392–1402. [CrossRef]
10. Martinsen, G.; Liu, S.; Mo, X.; Bauer-Gottwein, P. Joint optimization of water allocation and water quality management in Haihe River basin. *Sci. Total Environ.* **2019**, *654*, 72–84. [CrossRef]
11. Nukapothuala, S.; Chen, C.; Wu, J. Long-term distribution patterns of remotely sensed water quality variables in Pearl River Delta, China. *Estuar. Coast. Shelf Sci.* **2019**, *221*, 90–103. [CrossRef]
12. Wang, J.-H.; Yang, C.; He, L.-Q.-S.; Dao, G.-H.; Du, J.-S.; Han, Y.-P.; Wu, G.-X.; Wu, Q.-Y.; Hu, H.-Y. Meteorological factors and water quality changes of Plateau Lake Dianchi in China (1990–2015) and their joint influences on cyanobacterial blooms. *Sci. Total Environ.* **2019**, *665*, 406–418. [CrossRef]
13. Xiao, J.; Wang, L.; Deng, L.; Jin, Z. Characteristics, sources, water quality and health risk assessment of trace elements in river water and well water in the Chinese Loess Plateau. *Sci. Total Environ.* **2019**, *650*, 2004–2012. [CrossRef]
14. MEE. *Annual Report on Environmental Quality in China (1981 to 2019)*; China Environment Publishing Group: Beijing, China, 2020.
15. MEE. *Report on the State of the Ecology and Environment in China 1986–2020*; MEE: Beijing, China, 2021. Available online: http://english.mee.gov.cn/Resources/Reports/soe/index.shtml (accessed on 20 April 2022).
16. MEE. *Annual Statistics Report on Environment in China*; China Environment Publishing Group: Beijing, China, 2016.
17. NBSPRC. *China Statistical Yearbook 1982–2020*; China Statistical Press: Beijing, China, 2021.
18. Editorial board of China Environment Yearbook. *China Environmental Yearbook 1989–2020*; China Forum of Environmental Journalists: Beijing, China, 2021.
19. Su, J.; Ji, D.; Lin, M.; Chen, Y.; Sun, Y.; Huo, S.; Zhu, J.; Xi, B. Developing surface water quality standards in China. *Resour. Conserv. Recycl.* **2017**, *117*, 294–303. [CrossRef]
20. Wang, C.; Wu, J.; Zhang, B. Environmental regulation, emissions and productivity: Evidence from Chinese COD-emitting manufacturers. *J. Environ. Econ. Manag.* **2018**, *92*, 54–73. [CrossRef]
21. Chen, H.; Hao, Y.; Li, J.; Song, X. The impact of environmental regulation, shadow economy, and corruption on environmental quality: Theory and empirical evidence from China. *J. Clean. Prod.* **2018**, *195*, 200–214. [CrossRef]
22. Zhou, Y.; Ma, J.; Zhang, Y.; Qin, B.; Jeppesen, E.; Shi, K.; Brookes, J.D.; Spencer, R.G.M.; Zhu, G.; Gao, G. Improving water quality in China: Environmental investment pays dividends. *Water Res.* **2017**, *118*, 152–159. [CrossRef]
23. Guttman, D.; Young, O.; Jing, Y.; Bramble, B.; Bu, M.; Chen, C.; Furst, K.; Hu, T.; Li, Y.; Logan, K.; et al. Environmental governance in China: Interactions between the state and "nonstate actors". *J. Environ. Manag.* **2018**, *220*, 126–135. [CrossRef]
24. Zhang, B.; Cao, C.; Hughes, R.M.; Davis, W.S. China's new environmental protection regulatory regime: Effects and gaps. *J. Environ. Manag.* **2017**, *187*, 464–469. [CrossRef]
25. Wang, D.Q.; Changbo, M.L.; Wang, J. Research in Reconstructing the System of National Water Quality Management in China. *Environ. Prot.* **2017**, *8*, 49–56. (In Chinese)
26. Cheng, X.; Chen, L.; Sun, R.; Jing, Y. Identification of regional water resource stress based on water quantity and quality: A case study in a rapid urbanization region of China. *J. Clean. Prod.* **2019**, *209*, 216–223. [CrossRef]
27. Li, K.; Fang, L.; He, L. How population and energy price affect China's environmental pollution? *Energy Policy* **2019**, *129*, 386–396. [CrossRef]

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

*Water* Editorial Office
E-mail: water@mdpi.com
www.mdpi.com/journal/water

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34

www.mdpi.com


ISBN 978-3-0365-7147-8