remote sensing

MDPI

Article

Imagery Time Series Cloud Removal and Classification Using Long Short Term Memory Neural Networks

Francisco Alonso-Sarria *, Carmen Valdivieso-Ros and Francisco Gomariz-Castillo

Instituto Universitario del Agua y del Medio Ambiente, Universidad de Murcia, 30100 Murcia, Spain; mcarmen.valdivieso@um.es (C.V.-R.); fjgomariz@um.es (F.G.-C.)
* Correspondence: alonsarp@um.es

Abstract: The availability of high spatial and temporal resolution imagery, such as that provided by the Sentinel satellites, allows the use of image time series to classify land cover. Recurrent neural networks (RNNs) are a clear candidate for such an approach; however, the presence of clouds poses a difficulty. In this paper, random forest (RF) and RNNs are used to reconstruct cloud-covered pixels using data from other next in time images instead of pixels in the same image. Additionally, two RNN architectures are tested to classify land cover from the series, treating reflectivities as time series and also treating spectral signatures as time series. The results are compared with an RF classification. The results for cloud removal show a high accuracy with a maximum RMSE of 0.057 for RNN and 0.038 for RF over all images and bands analysed. In terms of classification, the RNN model obtained higher accuracy (over 0.92 in the test data for the best hyperparameter combinations) than the RF model (0.905). However, the temporal–spectral model accuracies did not reach 0.9 in any case.

Keywords: Sentinel-2; cloud removal; LSTM

check for updates

Citation: Alonso-Sarria, F.; Valdivieso-Ros, C.; Gomariz-Castillo, F. Imagery Time Series Cloud Removal and Classification Using Long Short Term Memory Neural Networks. *Remote Sens.* **2024**, *1*, 2150. https://doi.org/10.3390/rs16122150

Academic Editors: Emilio Guirado and Javier Blanco-Sacristán

Received: 27 March 2024
Revised: 7 June 2024
Accepted: 10 June 2024
Published: 13 June 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Land use, land cover and their changes are among the most important environmental variables with a clear influence on environmental processes at both global and regional scales [1,2]. They are also important in planning [3,4], urban mapping [5,6] or ecosystem services valuation [7].

The need to understand the distribution of land use and land cover, as well as their temporal evolution, has led to the rapid application of various statistical and machine learning methods to extract this information from satellite imagery. The European Union has even promoted the monitoring of agricultural parcels using remote sensing techniques [8]. Among the most commonly used supervised classification algorithms in remote sensing are k-nearest neighbors, support vector machines, random forests and neural networks [9]. Deep learning (DL) is an approach to neural networks that is capable of inferring features appropriate to the task at hand from the input variables [10], reducing the preprocessing step and producing higher accuracies than classical machine learning models [11]. The accuracy of DL approaches usually improves other machine learning methods in RS classification [11–13]. They have been applied to various tasks whose input is sequential data, such as speech recognition [14], human activity classification from body sensor data [15] or air quality prediction [16]. They can also be used in the temporal domain of RS [17] or even in the spectral domain [18]. Recent reviews on DL and its application in RS are provided by [19,20], or [21]. Convolutional (CNNs) and recurrent (RNNs) neural networks are among the most commonly used neural network architectures [9,22].

Time series of medium spatial resolution imagery have demonstrated a significant ability to characterise environmental phenomena, describing trends as well as discrete changes [23]. Seasonality is a key characteristic of vegetation, and thus, multi-temporal remote sensing is useful for monitoring growth dynamics for vegetation classification [24,25]. For these

reasons, time series have been increasingly used to map Land Use/Land Cover (LULC) and to identify the nature of land cover change. Time series can outperform single date classifications [26–28]. The easiest way to perform a multi-temporal land cover classification is to include multiple images from significant dates, as intra-annual phenological diversity can help discriminate between different land covers [28–31]. Another option would be to extract temporal features (simple statistics) or phenological metrics from the time series. A more complex approach is to summarise the time series using mathematical models, such as Fourier transform, wavelet transform, spline fitting, sigmoid function or Kalman filter. The drawbacks of these approaches are that they require manual work and use predefined models [11].

In recent years, new satellites with higher spatio-temporal resolution, but also shorter revisit times, have allowed higher accuracy using machine learning methods. The European Space Agency's (ESA) Sentinel programme is one of the latest missions focusing on many different aspects of Earth observation, including land monitoring [32]. Sentinel-2 (S2) consists of two twin-polar orbiting satellites (Sentinel-2A and 2B) and provides higher spatial and temporal resolution: 5 days at the equator and 2–3 days under cloud-free conditions in the mid-latitudes [33,34], although it is fully interoperable with the previous Landsat satellite programme [33]. Its higher temporal resolution allows for a greater number of images per month, increasing the likelihood that several of them will be cloud-free for a given period of time.

In an RNN, the response (y) to a given set of predictors (x) depends not only on this set of predictors but also on the previously processed sets, which have given the network an internal state (h) that allows the response to vary depending on the previous (past) sets [35]. The definition of an RNN is thus based on the concept of a cell, an operation with two inputs and two outputs. The inputs would be the value of each new register to be analysed and the network state, and the outputs would be the output for that register and the updated network state. It is not possible to train an RNN directly with the backpropagation algorithm; instead, time-based backpropagation is used, which is based on "unrolling" an RNN in as many steps as the set sequence length, but if this is too long, problems of vanishing and explosion of the gradient can occur.

The long short term memory (LSTM) cell, proposed by [36], is a unit used in RNN that is capable of learning dependencies between nearby points and with respect to distant points. It performs this by connecting the input, the previous state and the output through logic gates (linear operations) that decide which information is retained. The linearity of this process limits the disappearance/explosion of the gradient (Figure 1).

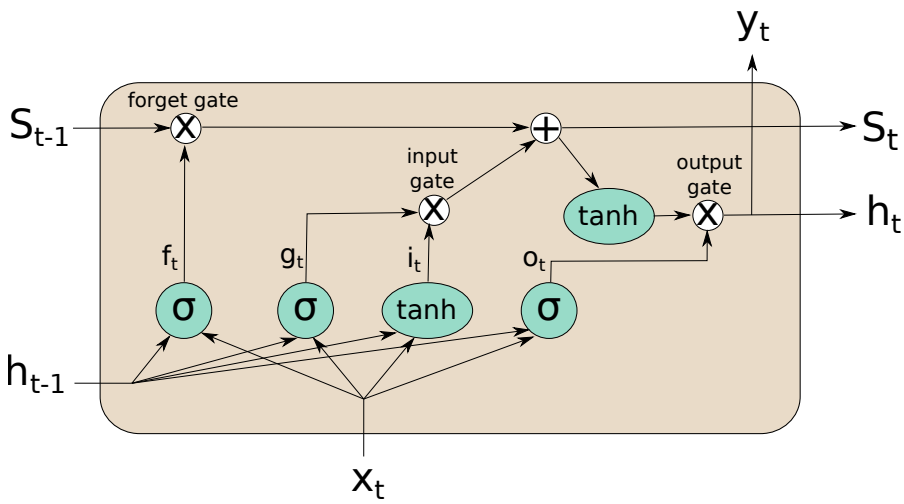

Figure 1. LSTM cell schema.

Deep learning algorithms allow models to be automatically applied to time series of images collected by satellites over a crop cycle [22]. Ref. [37] used an LSTM to extract spec-

tral and temporal features for change detection. Other variants use CNN; one-dimensional CNNs may be used to analyse sequences of data as RS time series [38–41]. It is also possible to extract spatial features from the image using two-dimensional CNN layers that are used as inputs to RNN layers [18]. Therefore, the authors of [42] used Conv2D layers to extract spatial features and a bidirectional sequential encoder to process the resulting time series, and [43] used CNNs with very-high-resolution aerial orthoimagery and NDVI calculated from Sentinel-2 data. Bidirectional LSTM cells allow the time series to be processed in both directions, so the authors of [9] proposed a recurrent neural network model with 2 Bi-LSTM layers and a fully connected layer to classify 10 crops using Sentinel-2 imagery reflectance and NDVI. Ref. [44] achieved the best results with a Bi-LSTM neural network and random forest by classifying active or abandoned crops and eight subclasses. More complex DL models such as transformers have also recently been used to classify time series of satellite images [10].

The presence of clouds and their shadows severely limits the ability to use satellite imagery for land use classification. This problem affects all spectral bands of passive sensors and is particularly severe when trying to integrate images from different dates, especially when considering pixel classification from reflectivity time series [45,46]. NASA or ESA image search and download applications allow for the inclusion of a cloud threshold, and the high temporal resolution of Sentinel-2 makes it possible to obtain several images with a low cloud percentage. However, even images with low cloud cover often contain small cumulus clouds and their shadows. If a time series of images is to be used, the union of cumulus clouds and their shadows in each image may represent a significant portion of the study area. In cloudy areas, the number of images available can be significantly reduced. Best-available-pixel composites are an option for creating proper images [23,47]; however, this approach might introduce noise into the composition, which also means a reduction in the number of available images.

The presence of clouds actually poses two problems: first, the localisation of clouds, and second, the filling of clouds by estimating ground reflectivity values. Several techniques have been proposed for cloud detection, and even a comparison exercise has recently been carried out [48]. In almost all cases, indices are used that highlight the differences in spectral response between clouds and ground cover. The detection of shadows is therefore much more complex, as the differences in reflectivity are not as relevant. The models used generally prefer to generate false positives rather than false negatives, including the use of buffers from the detected clouds or shadows.

On the other hand, several techniques have been proposed to estimate the ground reflectivity in cloud pixels to complete the image. Ref. [49] divided these techniques into three categories: spatial, temporal and mixed. The former are generally based on spatial interpolations, e.g., nearest neighbour or geostatistical. These methods require that the area of interest is not very large and that there is a high degree of regularity in the coverages present in the image and their spatial distribution. Temporal methods use before and after images to estimate the reflectivity in the missing pixels assuming a certain constancy in the coverages. Logically, this will depend on the temporal distance between the image to be reconstructed and those used as predictors. It is clear that better results are obtained by using images from sensors with high temporal resolution. In general, the reflectance of each band is predicted from the reflectance of the same band in other pixels of the same image or from images of different dates. Ref. [49] proposed using all bands as predictors, thus also introducing spectral information, using Random Forest as a prediction model.

The most common spatial techniques used to estimate the values of pixels without data include interpolation [50–53] using ordinary kriging [54] or co-kriging [55], among others. Multi-spectral information from cloud-insensitive bands [56–59] and multi-temporal information from cloud-free images taken by other sensors at a similar date [60] or without filling gaps [11] have also been used, followed by the further development of methods using multi-sensor data fusion [61,62]. Ref. [52] proposed the method called the Neighborhood Similar Pixel Interpolator (NSPI), which achieved variable RMSEs on the accuracy of the

filled results depending on the date of the image. However, the algorithms developed are based on the assumption of the availability of another image in a short time interval, a condition that is not always met.

Recent attempts to use deep learning for cloud filling in satellite images include work by the authors of [63], who proposed an architecture integrating convolutional cells to use information from close in time images to fill the encountered gaps in Landsat 8 and Sentinel 2 images; Ref. [49] concatenated the outputs of convolutional layers with small kernel sizes in a LSTM cell, applying them to restore cloudy pixels in MODIS images. Generative adversarial networks (GANs) have been also used to attain this objective, for instance, in work by [64], where two variations of GAN, Pix2Pix and Spatiotemporal GAN, are used to produce a cloud-free image for those dates where the available image has cloud cover. The model is trained on earlier and later cloud-free images. Ref. [65] used a multi-modal and multi-temporal 3D convolutional neural network and sequence-to-sequence translation model using Sentinel1 SAR images to fill Sentinel2 cloudy images. Finally, Ref. [66] proposed the use of conditional generative adversarial networks and convolutional long short-term memory networks to fill cloud pixels. These studies share the use of a large global data set of Sentinel2 images from around the world, covering all the Earth's biomes.

The aim of this study is to test a method to simultaneously detect and correct the presence of clouds in the image series that will later be subjected to land cover classification. The proposed method is a temporal one that uses recurrent neural networks to generate a model from the pixels without cloud or shadow. This model is then used to estimate the reflectance of the pixels suspected of being cloudy or shaded. The similarity between the value in the image and the estimated value makes it possible to determine in each case whether the pixel is a cloud/shadow or not. In the first case, the estimated value is used, and in the second case, the original value is used. The results are compared with the use of random forest with the reflectances from the previous and subsequent images, as in [49], the average of the previous and subsequent reflectance values, and a linear interpolation. This strategy was tested in a study area characteristic of the semi-arid Mediterranean climate, where such problems are common, with data from 2018 using all Sentinel-1 and Sentinel-2 data available in that year.

As a secondary objective, the land cover of the study area in 2018 was classified using all available Sentinel-1 and Sentinel-2 data in that year. Two neural network models were used for this purpose. One is a combination of two bidirectional LSTM layers to extract temporal features of reflectance values, followed by two dense layers. The second combines an LSTM layer to extract spectral features from the spectral signatures, which are the input to a temporal LSTM layer to extract temporal features from them. Finally, two dense layers are used to process these features and produce a prediction. The results of both models are compared with a random forest model, which is used as a baseline model. Although RNNs have been used to process hyperespectral data [18], to the best of our knowledge, there have been no previous attempts to integrate LSTM cells in a single model to process spectral and temporal information.

We propose a simple LSTM model as a method to fill cloudy pixels using image time series. We want to verify the ability of simple models to perform remote sensing tasks. On the other hand, we propose, to the best of our knowledge, a new neural network model that treats spectral information as a series using LSTM. In our study, we have assumed that cloud removal would increase classification accuracy. In order to test this hypothesis, a set of image series with different percentages of cloud images would be needed to take into account how accuracy changes in response to two factors: image availability and correction/non-correction.

2. Study Area

Both objectives were carried out on a set of Sentinel-2 images corresponding to the Campo de Cartagena or Mar Menor watershed (1275 km^2) in south-eastern Spain (Figure 2).

The climate is semi-arid Mediterranean (mean annual precipitation 300–350 mm), with high rainfall irregularity, but high temporal variability, leading to a regular alternation of extreme droughts and floods. Temperatures are warm throughout the year. Despite the low rainfall, the characteristics of the soil, temperature and orography make the area very suitable for agricultural purposes. Cultivated since ancient times, in the last fifty years, there has been a gradual change from dry farming to irrigated farming, using water transferred from the Tagus River, water from desalination plants and also groundwater. In terms of natural vegetation, there is a great diversity and heterogeneity of vegetation, mainly Mediterranean scrub, although there are also patches of Mediterranean woodland. The other main use in this area is urban; many large urbanised areas, whose summer population increase is difficult to quantify, are located along the coastline delimiting the lagoon.

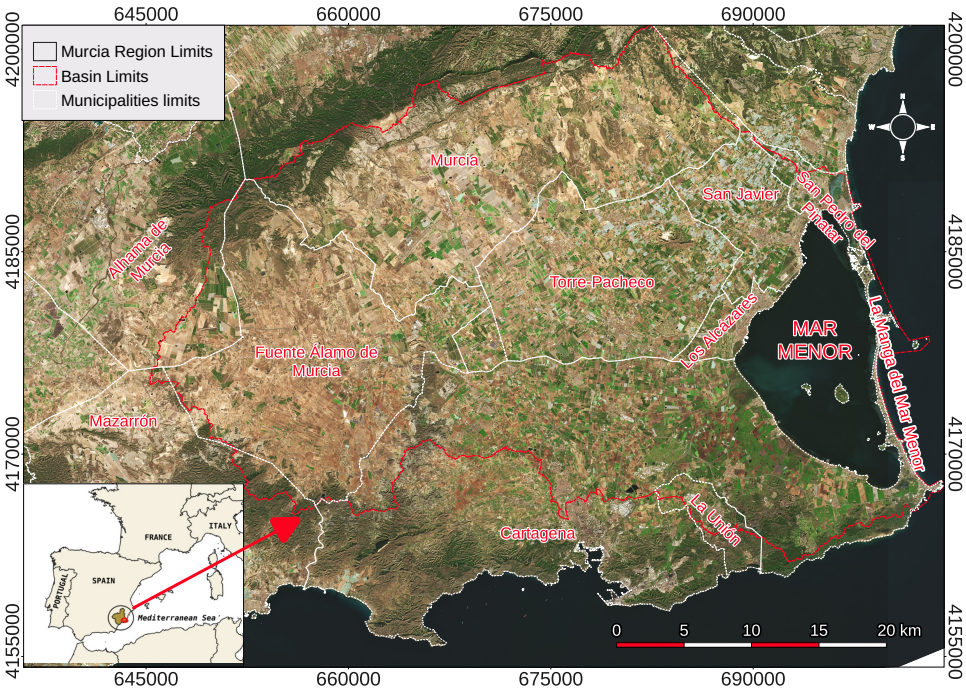

Figure 2. Study area: Mar Menor Basin. Source: BTN25 2006–2019 CC-BY 4.0 ign.es (administrative data); Data derived from MDT25 2015 CC-BY 4.0 ign.es.

Semi-arid Mediterranean regions represent a major challenge for land cover classification using remote sensing imagery. This is due to the high spatial irregularity associated with specific socio-economic and physical characteristics. This irregularity includes a high diversity of spatial patterns, fragmentation and a wide range of vegetation cover [5,67,68].

Although semi-arid, this is a coastal area, so the appearance of cumulus clouds is frequent due to the inflow of moist air from the Mediterranean Sea. For this reason, this study area is a good candidate for testing cloud removal algorithms.

3. Methodology

3.1. Data Set and Preprocessing

Figure 3 shows the methodological workflow of the study. It starts with the preprocessing of Sentinel 1 and 2 images and then continues with cloud and shadow detection to calibrate the cloud filling models before the classification process that concludes the workflow. Sentinel-2 images were only downloaded when the cloud cover was less than 2%. These images are shown in Figure 4. There are more S1 images than S2 because the former are not affected by clouds, and the figure only shows the data used. For models using both S1 and S2 images, each S2 image was paired with the closest S1 image in time.

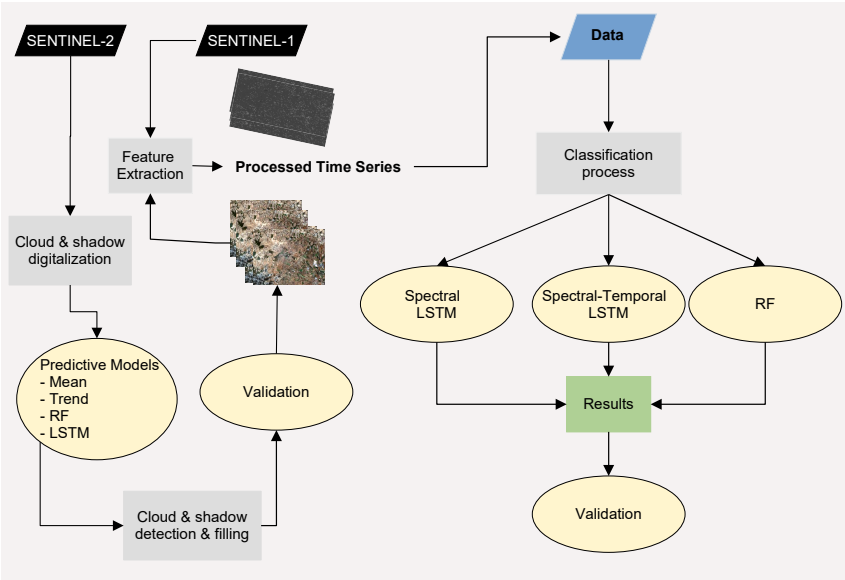

Figure 3. Methodological workflow of the study.

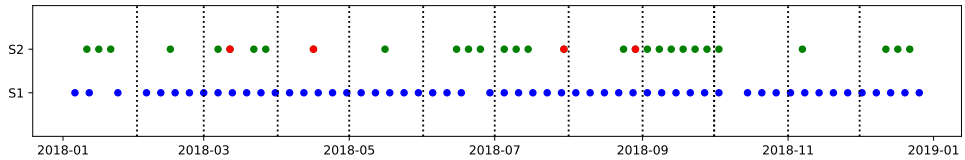

Figure 4. Available S1 (blue) and S2 (green) images. The S2 images that presented cloudiness are shown in red.

S2 Top of Atmosphere (TOA) data (L1C product) were acquired and corrected using ACOLITE [69], with the interface and code freely distributed by the Remote Sensing and Ecosystem Modelling (REMSEM) team, part of the Royal Belgian Institute of Natural Science (RBINS) and the Operational Directorate Natural Environment (OD Nature). This correction method resulted in a more accurate classification than other atmospheric correction algorithms with part of this data set [31].

3.2. Detection and Filling of Missing Pixels Due to Clouds or Shadows

Four S2 images appeared with small cumuli, although the percentage of cloudiness was below the threshold: 12 March, 16 April, 30 July and 29 August. The four before and after images were used to detect cloud pixels and estimate their ground reflectivities.

The starting point is a rough digitisation of the areas affected by clouds; the aim is not to be precise, but to include all clouds and shadows present in the image, leaving enough pixels outside the affected areas to calibrate a model. In a second step, a large set of training pixels, outside the cloud-affected area, is selected for a predictive model. The response variables of this model are the reflectivities in the image from which the clouds are to be removed and the predictors are the same reflectivities in the unaffected images. In reality, several images can be corrected simultaneously as long as the union of the cloud-affected areas in each of these images leaves a sufficiently large and diverse space to obtain a good sample. This procedure is only valid when the cloud cover to be removed does not occupy a large percentage of the image in question, although we believe that it can give better results than spatial filling methods when the percentage occupied is intermediate.

The procedures used to estimate the reflectivities in the unclouded pixels were as follows: the mean of the reflectivities on the before and after dates, an estimate from the before and after reflectivity values, a random forest model using all the bands of the before

and after dates, similar to the procedure proposed by [49] and a recurrent neural network (LSTM) using all the reflectivities on four before and four after dates as predictors. This network (Figure 5) starts with two branches, one for the previous values and the other for the posterior values, which are processed in reverse order. Both series pass through an LSTM layer and two dense layers. The outputs of the last two dense layers are then concatenated and pass through 3 more dense layers.

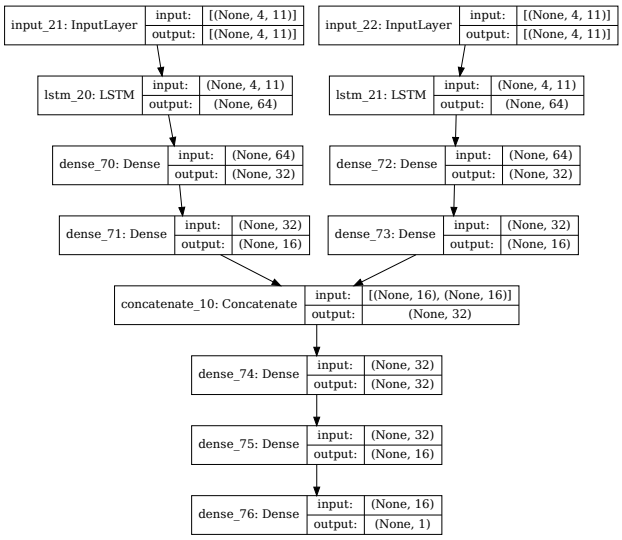

Figure 5. Architecture of the LSTM model for cloud filing.

To determine the accuracy of the model, the set of sample points obtained was divided into training/validation and test points. A random point 15 km from any cloud was chosen as the centre of a 7.5 km circle within which the test pixels were sampled. This decision ensures that the test areas are independent of the training and validation areas. The root mean square error (RMSE) and the coefficient of determination (R^2) were used as accuracy statistics. These are are classical statistics for measuring goodness of fit in regression problems. Several new metrics have been proposed from computer vision to measure the fit between an original image and a reconstructed version after adding some noise: Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM) and Spectral Angle Mapper (SAM) [64–66,70]. Such metrics assume that an original noiseless image is available. This is not the case when trying to remove clouds from remote sensing images, unless clouds are simulated on a cloudless image. However, simulating a realistic cloud on an image is not a trivial task. In this study, we did not simulate clouds but attempted to remove real clouds. We adapted the PSNR to be calculated from the MSE obtained in cross validation and the maximum reflectivity computed in the cloudless portion of the image. The PSNR equation is

$$PSNR = 10 \cdot log_{10}\left(\frac{MAX_I^2}{MSE}\right) \quad (1)$$

where MAX_I^2 is the squared maximum cloud-free value in the image, and MSE is the mean squared error. The objective of such a metric is to re-escalate MSE in terms of the real range of values in the image, allowing a fairer comparison between bands or study areas. The results are log-transformed to obtain orders of magnitude.

In the case of the neural network, the training set of pixels was divided into training and validation as is usually done when training neural network models. Figure 6 shows the training, validation and test points for the July image. The whole procedure was developed in Python 3.9.18 using the numpy 1.26.4, scikit-learn 1.3.0 and tensorflow 2.4.1 libraries.

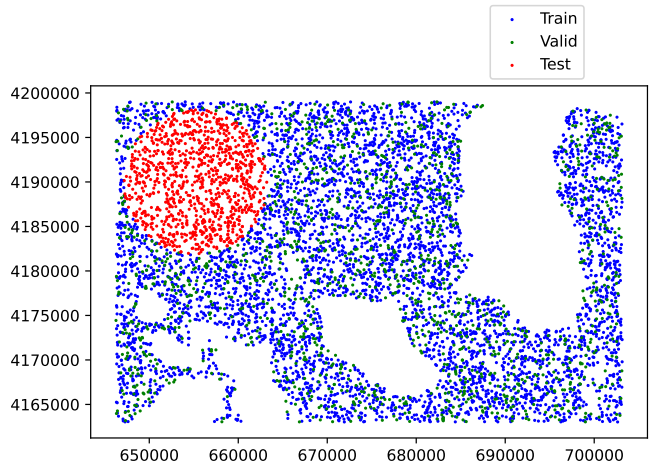

Figure 6. Training, validation and test points for the July cloud-filling model. Areas without any point correspond to the clouds in the image to be removed.

Once a model has been calibrated from the pixels outside the potentially cloudy areas, the difference between the original reflectivities and those estimated by the model in the potentially cloudy areas is evaluated. Only if the difference exceeds a number of standard deviations from the mean (a threshold to be set a priori) is the pixel considered to be a cloud or a shadow, and the original value is replaced by the estimated value. A high threshold may leave some shadow pixels undetected, but a low threshold may modify pixels that do not need it. In any case, the accuracy of the models is so high that this would not be a serious problem.

3.3. Reflectance Time-Series Classification

Once the cloud-filling process was completed, the classification of the entire time series of images was carried out.

3.3.1. Training Areas and Classification Scheme

The training areas were digitised using aerial orthophotography from the Spanish National Aerial Orthophotography Plan (PNOA) acquired in 2016 and 2019. A stratified sampling design guarantees an adequate presence of all classes. The representativeness of the data set was improved using Isolation Forest with the methodology proposed in [71], resulting in a total of 131 polygons (Table 1). The "netting" class represents recent agricultural practices with a spectral response different from that of greenhouses or irrigated crops. It consists of nets covering trees for their protection. Although there are some remnants of rainfed plots, they are not included as most of them are being converted to irrigation or are abandoned.

Table 1. Land covers taken into account in the classification, including the number of polygons for training and validation and number of total pixels per class.

Class	Description	Polygons	Pixels
Forest	Mediterranean forest	10	1000
Scrub	Scrubland	12	1200
Dense tree crops	Fruit and citrus trees	18	1800
Irrigated grass crops	Mainly horticultural crops	10	1000
Impermeable	All artificial surfaces	18	1639
Water	Water bodies, including artificial reservoirs	12	1158
Bare soil	Uncovered or low-vegetation covered land	11	1055
Greenhouses	Irrigated crops surfaces under plastics structures	26	2600
Netting	Irrigated tree and vegetables crops covered by nets	14	1400
Total		131	12,852

3.3.2. Features

The use of indices to recognise biophysical patterns is a common practice [72–75]. The following four indices were calculated from the reflectivity layers for each image:

- The Normalised Difference Vegetation Index (NDVI) [76] (Equation (2)).

$$NDVI = \frac{B8A - B4}{B8A + B4} \tag{2}$$

  where B8A is a narrow near-infrared (NIR) band for vegetation detection, and B4 is the red band (R) from S2A MSI.
- Tasseled cap brightness (TCB) [77] attempts to highlight spectral information from satellite imagery that detects variations in soil reflectance (Equation (3)):

$$\begin{aligned} TCB = (0.3037 \cdot B2) + (0.2793 \cdot B3) + (0.4743 \cdot B4) + (0.5585 \cdot B8) \\ +(0.5082 \cdot B11) + (0.1863 \cdot B12) \end{aligned} \tag{3}$$

  where B1, B2 and B3 are the blue (B), green (G) and red (R) bands, respectively; B8 is the NIR band; and B11 and B12 are the SWIR of the S2A MSI. A more recent formulation can be found in [78].
- The Normalised Difference Built-up Index (NDBI) [79] is used to distinguish built-up surfaces, which receive positive values, from bare soils (Equation (4)):

$$NDBI = \frac{B11 - B8}{B11 + B8} \tag{4}$$

  where B11 is the SWIR band, and B8 is the NIR band.
- The Modified Normalised Difference Water Index (MNDWI) [80] was proposed to detect water surfaces. However, it can also be used to detect water in vegetation or soil covers (Equation (5)):

$$MNDWI = \frac{B3 - B11}{B3 + B11} \tag{5}$$

  where B3 is the green band, and B11 is the SWIR band.

These indices were selected from the large number of indices available because they cover the main features in the study area that affect reflectivity values and coverage.

Other metrics obtained from optical data are texture features. Haralick et al. [81] proposed the use of a gray-level co-occurrence matrix (GLCM) as a method of quantifying the spatial relationship between neighbouring pixels in an image. Haralick texture features, computed from the GLCM, are widely used due to their simplicity and intuitive interpretations, and have been successfully applied in remote sensing [82]. From the large number of Haralick metrics, we used only the three metrics recommended in [82]: Contrast, Entropy and Angular Second Moment (Equations (6)–(8)).

$$Contrast = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \tag{6}$$

$$Entropy = \sum_{i,j=0}^{N-1} P_{i,j}(-logP_{i,j}) \tag{7}$$

$$AngularSecondMoment = \sum_{i,j=0}^{N-1} P_{i,j}^2 \tag{8}$$

where $P_{i,j}$ is the probability of i and j values occurring in adjacent pixels, and i and j are the column and row labels of the GLCM, respectively. Such metrics were calculated on two summary layers per date: the first principal component of the spectral layers (i.e., an albedo layer) and the NDVI.

The Sentinel-1 SAR images were selected in the Interferometric Wide (IW) mode with a full swath of 250 km and a spatial resolution of 5 × 20 m in a single view. This is the main Terrain Observations by Progressive Scans SAR (TOPSAR) acquisition mode, with three merged sub-swaths consisting of a series of overlapping bursts steering the beam from back to front in azimuth direction. The angle of incidence used for the IW mode ranges from 29.1° to 46°. The final Ground Range Detected (GRD) product is resampled to the common pixel spacing range, meaning that the data have been focused, multi-looked and projected to ground range using an Earth ellipsoid model.

For this study, all SAR images available for a full year were used, so there are images acquired in both ascending and descending directions, with angles of incidence ranging from about 30 to 46 degrees. This implies possible differences in backscatter intensity due to the effect of the local angle of incidence on the pixel area, although such geometric issues were probably not very serious due to the low slope of the study area. However, the possible geometric issues were corrected by applying a radiometric terrain correction in the pre-processing workflow [83].

Pre-processing of the S1 SAR images was performed in SNAP 9.0.6 in batch mode, and included the following steps: (1) orbit file application, (2) GRD boundary noise removal, (3) calibration, (4) terrain correction, (5) resampling, (6) conversion to dB and (7) speckle filtering. Terrain correction was performed using the SRTM 3Sec HGT as Digital Elevation Model (DEM) with bilinear interpolation and resampling to 10 m with the nearest neighbour model using the same projection of S2 data. The intensity bands in co- and cross-polarization (VV and VH) were used for this study.

In addition, the Dual Polarization SAR Vegetation Index (DPSVI) was calculated to separate bare soil from vegetation:

$$DPSVI = \frac{\sigma^0 VV + \sigma^0 VH}{\sigma^0 VV} \tag{9}$$

This means three features (VV, VH and DPSVI) per date, giving a total of twelve features.

3.3.3. Models

Two RNN models are tested. The first one is a simple LSTM layer followed by a dense layer (Figure 7), and the second one is a spectral–temporal model. This model combines an LSTM layer to extract features from the spectral signatures, followed by another LSTM layer to extract temporal features from the previous one, and finally, a dense layer (Figure 8). Finally, temporal SAR data are introduced, when necessary, as a secondary input branch to the previous model.

Calibration was performed using the Adam optimiser and 100 epochs, but with early stopping and a patience value of 20. All combinations of four batch sizes (8, 16, 32 and 64) and nine drop rates (0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7 and 0.8) were tested. Batch size and drop rate are two important hyperparameters in the training of a neural network. Batch size refers to the number of training examples used in an iteration of the model training process. Common batch sizes are powers of 2 for efficient memory allocation on GPUs. Larger batch sizes allow for more efficient use of hardware resources, potentially speeding up the training process. They also provide a more accurate estimate of the gradient but may result in poorer generalisation. Smaller batch sizes may introduce more noise into the training process, but this noise may act as a regulariser that helps the model avoid overfitting. The drop rate refers to the probability that an individual neuron is set to zero during each training forward pass through the network. This can prevent overfitting by ensuring that the model does not become overly reliant on any particular set of neurons. The model is then forced to learn more robust features that are useful in combination with many different subsets of the other neurons. Fine-tuning these hyperparameters based on the specific problem and data set is essential for optimal model performance.

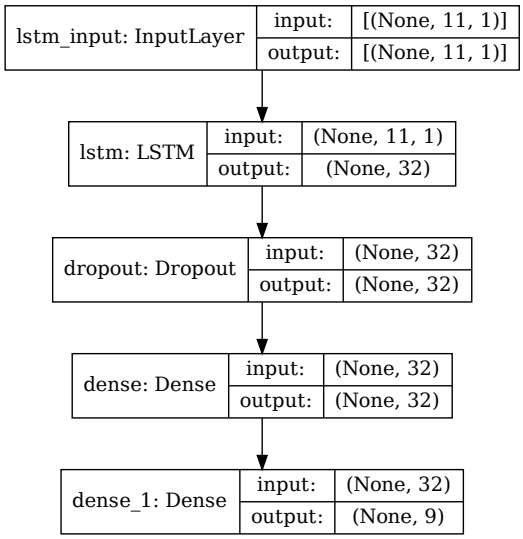

Figure 7. Architecture of the temporal RNN model.

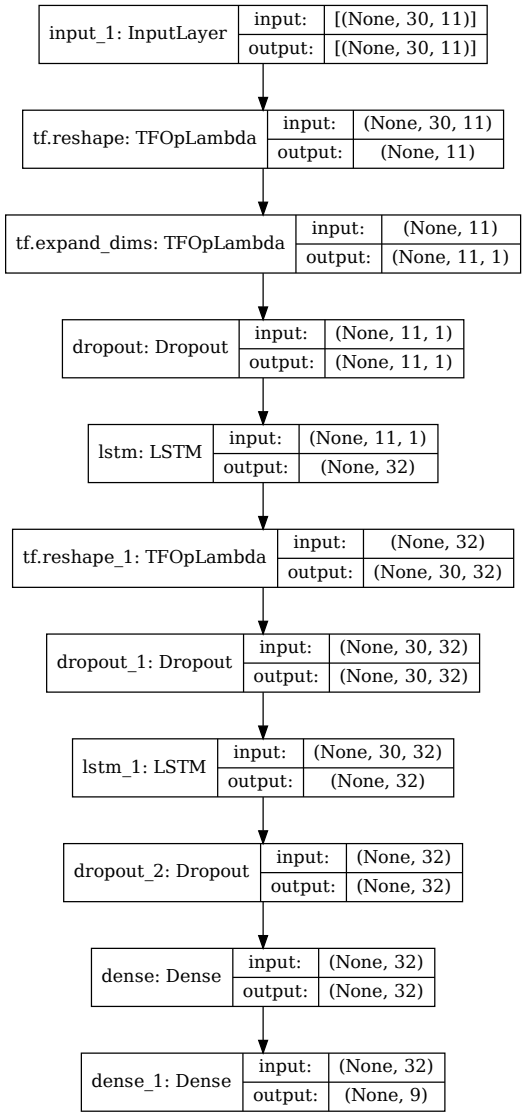

Figure 8. Architecture of the spectral–temporal RNN model.

3.4. Validation

The three models were validated using 5-fold cross-validation. Each of the 5 groups contained complete polygons to avoid splitting them into training and test, thus preserving the independence between training and test pixels. As the calibration process of a neural network requires the use of a validation set, one of the 4 groups dedicated to training was used for this task. Accuracy was the chosen accuracy statistic and its standard deviation was calculated [84] to assess the significance of the differences between models and data.

4. Results

4.1. Cloud Removal

Table 2 shows the R^2, RMSE and PSNR values for the four cloud removal models tested and for all bands in the four images analysed. The most accurate models in each case are highlighted. In cases where the difference between LSTM and RF was considered negligible, both models are highlighted. The results are more accurate when applied to the summer images than to the spring images. This is probably due to the fact that there are more and closer in time images available for summer than spring (see Figure 4). It is worth noting that the R^2 values might be misleading, as in several cases, the best model in terms of R^2 is not the best model in terms of RMSE, indicating a generalized bias in the prediction. Another pattern that emerges from this analysis is that the shorter the wavelength of the band, the more accurate the models are. The PSNR results give the same information as the RMSE in terms of model comparison. The only differences in RMSE are between bands, as the maximum reflectivity changes for different wavelengths.

For the summer images, the best results are obtained with the LSTM network; the next best results are those of the random forest; those of the averages or the interpolation are slightly worse. In any case, all models achieve very high accuracy statistics. The accuracy of the model based on the mean of the reflectivities in the before and after image achieves RMSE values between 0.007 and 0.02 and R^2 values between 0.942 and 0.987. In the case of the random forest model, the values are between 0.005 and 0.019 for RMSE and 0.953 and 0.992 for R^2. Finally, the model based on the LSTM network obtains RMSE values between 0.005 and 0.011 and R^2 values between 0.986 and 0.991. It is clear that these are very high values in all cases, but, in any case, the results of the LSTM network achieve a significant percentage improvement in relation to the margin of improvement allowed by the other models. Figure 9 shows the whole image, original (right) and corrected (left). Figure 10 shows a detail of images from different dates. On the left is the original image; the red polygons show the pixels modified by the algorithm. The corrected image is shown on the right.

In spring, the results of random forest and RNN are similar for the April image, whereas in the case of the March image, RF gives more accurate predictions for all bands except for the first one, where the LSTM model is more accurate. For the other three bands, the differences between RF and LSTM are negligible. The RMSE values for RF are between 0.12 and 0.38, while those of the LSTM are between 0.01 and 0.057.

Table 2. Accuracy metrics of the cloud removal models for each band in the four images analysed. The best models are highlighted in bold. When the difference between two models is considered negligible, both models are highlighted.

Date	Band	NN			Mean			Trend			RF		
		r^2	RMSE	PSNR	r^2	RMSE	PSNR	r^2	RMSE	PSNR	r^2	RMSE	PSNR
03/12/2018	1	**0.867**	**0.010**	34.566	0.776	0.022	27.717	0.872	0.018	29.460	0.815	0.019	28.991
03/12/2018	2	0.823	**0.019**	33.872	0.618	0.037	28.083	0.687	0.031	29.620	**0.894**	**0.020**	33.427
03/12/2018	3	0.818	**0.024**	32.396	0.611	0.049	26.196	0.696	0.039	28.179	**0.904**	**0.023**	32.765
03/12/2018	4	0.732	0.048	26.375	0.633	0.072	22.853	0.758	0.054	25.352	**0.935**	**0.027**	31.373
03/12/2018	5	0.637	0.047	26.121	0.632	0.060	24.000	0.795	0.046	26.307	**0.898**	**0.030**	30.020
03/12/2018	6	0.845	**0.029**	30.632	0.775	0.040	27.839	0.778	0.044	27.011	**0.907**	**0.030**	30.338
03/12/2018	7	0.782	0.043	27.318	0.726	0.049	26.184	0.765	0.048	26.363	**0.924**	**0.031**	30.160
03/12/2018	8	0.810	0.042	27.535	0.710	0.052	25.680	0.737	0.053	25.514	**0.924**	**0.033**	29.630

Table 2. *Cont.*

Date	Band	NN			Mean			Trend			RF		
		r^2	RMSE	PSNR	r^2	RMSE	PSNR	r^2	RMSE	PSNR	r^2	RMSE	PSNR
03/12/2018	8A	0.771	0.046	26.492	0.758	0.047	26.306	0.789	0.048	26.123	**0.924**	**0.032**	29.645
03/12/2018	11	0.780	0.047	26.558	0.624	0.086	21.310	0.835	0.054	25.352	**0.900**	**0.035**	29.119
03/12/2018	12	0.728	0.057	24.448	0.693	0.082	21.289	0.893	0.050	25.586	**0.894**	**0.038**	27.970
04/16/2018	1	0.877	**0.011**	32.729	0.460	0.024	25.953	0.707	0.022	26.709	**0.927**	**0.012**	31.974
04/16/2018	2	0.862	**0.016**	34.151	0.440	0.034	27.604	0.630	0.033	27.863	**0.925**	**0.016**	34.151
04/16/2018	3	0.886	**0.016**	34.665	0.450	0.041	26.492	0.620	0.037	27.383	**0.933**	**0.017**	34.138
04/16/2018	4	0.876	**0.024**	31.725	0.329	0.065	23.071	0.569	0.053	24.844	**0.925**	**0.026**	31.030
04/16/2018	5	0.863	**0.022**	31.847	0.421	0.051	24.545	0.624	0.043	26.027	**0.917**	**0.024**	31.092
04/16/2018	6	0.778	**0.026**	30.453	0.559	0.043	26.083	0.522	0.047	25.310	**0.880**	**0.026**	30.453
04/16/2018	7	0.773	**0.030**	29.309	0.402	0.058	23.583	0.430	0.057	23.734	**0.855**	**0.032**	28.749
04/16/2018	8	0.790	**0.031**	29.442	0.387	0.060	23.706	0.406	0.062	23.421	**0.872**	**0.032**	29.166
04/16/2018	8A	0.769	**0.032**	28.987	0.400	0.060	23.527	0.442	0.057	23.973	**0.861**	**0.033**	28.720
04/16/2018	11	0.846	**0.030**	30.458	0.411	0.067	23.479	0.586	0.055	25.193	**0.898**	**0.033**	29.630
04/16/2018	12	0.864	**0.023**	32.330	0.336	0.077	21.835	0.578	0.061	23.858	**0.915**	**0.032**	29.462
07/30/2018	1	**0.991**	**0.005**	37.614	0.987	0.007	34.692	0.988	0.010	31.594	**0.992**	**0.005**	37.614
07/30/2018	2	**0.986**	**0.007**	39.036	0.942	0.012	34.354	0.944	0.013	33.659	0.953	0.011	35.110
07/30/2018	3	**0.992**	**0.007**	39.622	0.954	0.015	33.003	0.956	0.015	33.003	0.964	0.013	34.246
07/30/2018	4	**0.992**	**0.010**	37.377	0.957	0.020	31.356	0.959	0.021	30.932	0.965	0.019	31.802
07/30/2018	5	**0.991**	**0.010**	37.114	0.973	0.015	33.593	0.974	0.017	32.505	0.979	0.014	34.192
07/30/2018	6	**0.988**	**0.010**	36.835	0.970	0.015	33.313	0.971	0.016	32.753	0.975	0.014	33.912
07/30/2018	7	**0.990**	**0.010**	37.538	0.971	0.016	33.455	0.973	0.016	33.455	0.974	0.015	34.016
07/30/2018	8	**0.986**	**0.011**	36.184	0.950	0.020	30.991	0.953	0.021	30.567	0.958	0.019	31.437
07/30/2018	8A	**0.990**	**0.010**	38.169	0.973	0.016	34.087	0.975	0.016	34.087	0.975	0.015	34.647
07/30/2018	11	**0.995**	**0.011**	38.581	0.978	0.019	33.834	0.979	0.020	33.388	0.983	0.017	34.800
07/30/2018	12	**0.995**	**0.010**	39.676	0.981	0.019	34.100	0.982	0.019	34.100	0.985	0.017	35.067
08/29/2018	1	**0.989**	**0.005**	37.072	0.987	0.007	34.150	0.988	0.010	31.052	**0.991**	**0.005**	37.072
08/29/2018	2	**0.986**	**0.007**	40.149	0.942	0.012	35.467	0.944	0.013	34.772	0.961	0.010	37.051
08/29/2018	3	**0.990**	**0.007**	40.105	0.954	0.015	33.485	0.956	0.015	33.485	0.969	0.012	35.423
08/29/2018	4	**0.990**	**0.010**	37.488	0.957	0.020	31.468	0.959	0.021	31.044	0.974	0.016	33.406
08/29/2018	5	**0.991**	**0.009**	38.144	0.973	0.015	33.707	0.974	0.017	32.620	0.987	0.011	36.401
08/29/2018	6	**0.991**	**0.009**	38.166	0.970	0.015	33.729	0.971	0.016	33.168	0.985	0.011	36.423
08/29/2018	7	**0.990**	**0.010**	36.942	0.971	0.016	32.860	0.973	0.016	32.860	0.984	0.012	35.358
08/29/2018	8	**0.985**	**0.011**	37.326	0.950	0.020	32.134	0.953	0.020	32.134	0.973	0.014	35.232
08/29/2018	8A	**0.989**	**0.011**	36.542	0.973	0.016	33.288	0.975	0.016	33.288	0.983	0.012	35.786
08/29/2018	11	**0.996**	**0.009**	39.434	0.978	0.019	32.944	0.979	0.020	32.499	0.985	0.016	34.437
08/29/2018	12	**0.995**	**0.009**	38.372	0.981	0.019	31.882	0.982	0.019	31.882	0.988	0.014	34.534

Figure 9. *Cont.*

Figure 9. Real color composition of the 16 April 2018 image. The original image is shown above and the image corrected with LSTM below.

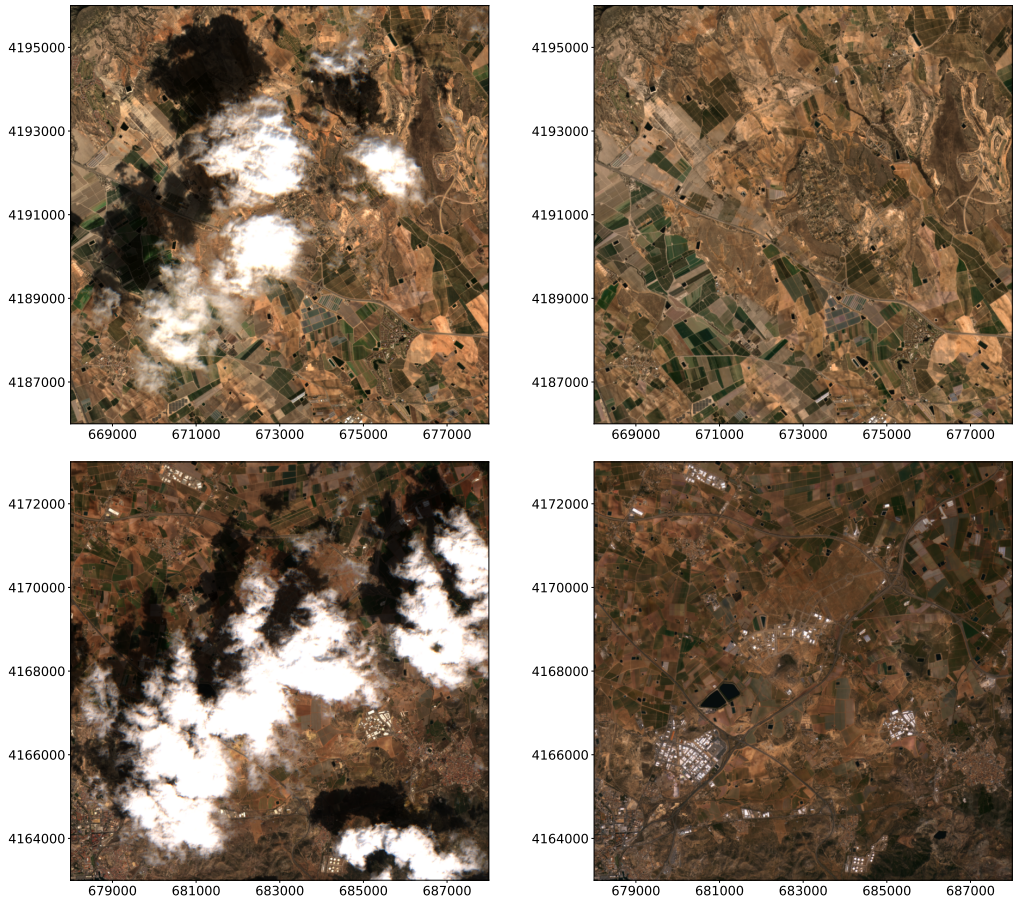

Figure 10. *Cont.*

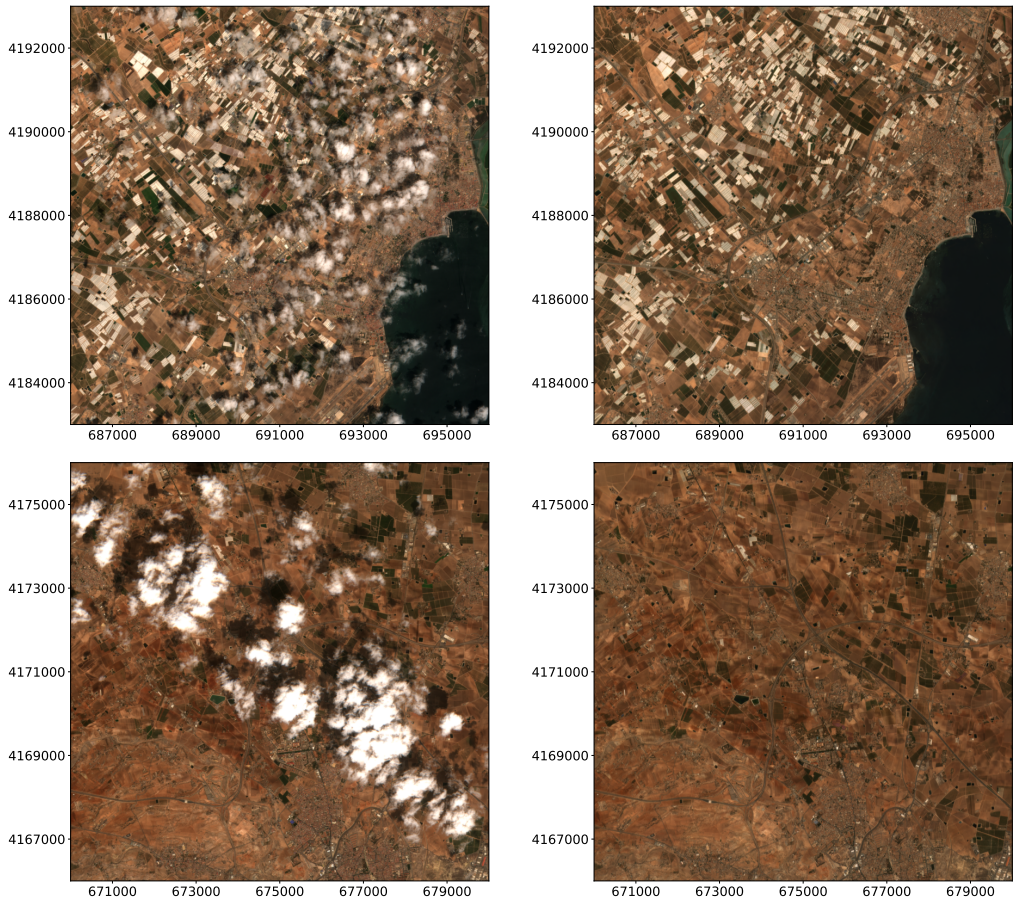

Figure 10. Zoom of the original (**left**) and LSTM corrected (**right**) images. The dates are, from top to bottom, 12 March 2018, 16 April 2018, 30 July 2018 and 29 August 2018.

4.2. Classification

Figure 11 shows the accuracy and kappa values obtained when classifying the temporal series of 2018 images using random forest with different data sets. The accuracy and kappa values are quite high, reaching a maximum accuracy of 0.9215 when using all data sets except SAR. However, the standard deviations, measured according to [84], show that the differences in accuracy using different data sets are not significant.

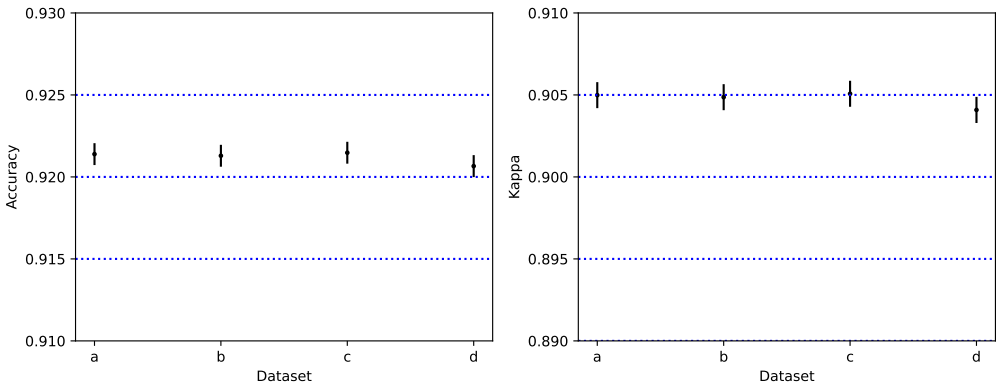

Figure 11. Results of the RF model with different data sets. (a) Spectral (b) Spectral + Indices; (c) Spectral +Indices + Texture; (d) Spectral + Indices + Texture + Radar. Bars show standard deviations.

Figure 12 shows the validation results of the temporal LSTM model using different data sets and different combinations of drop rate and batch size. In general, there is a clear increase in accuracy as the drop rate increases until a maximum accuracy is reached for values of 0.6–0.7. Accuracy also appears to increase with batch size, although the difference is not as clear when the drop rate is at the optimum values. It is only when batch size equals eight that the accuracy values obtained are sistematically lower. In terms of data sets, the best results are obtained when all data sets are used. It may be surprising that the second best option is to use only spectral data. The possible explanation is that the information contained in the indices is somehow retrieved by the weights in the neural network. In fact, the inclusion of texture values seems to actually reduce the accuracy.

Figure 13 shows the results of the temporal LSTM model on the test data. The general patterns are similar to those obtained with the validation data. However, there are differences that make it difficult to select the best model based on the validation data alone. In this case, the best validation accuracy results were obtained with the full data set, a batch size of 32 and a drop rate of 0.6; however, with the test data, the best results are obtained with the same data set and batch size, but a drop rate of 0.7.

The spectral–temporal LSTM model gives significantly lower accuracy results (Figures 14 and 15). In addition to the lower accuracy values, the relationship between batch size and accuracy is less clear than in the previous model. In addition, the pattern is reversed for the drop rate. Accuracy values decrease dramatically as the drop rate decreases. In fact, the best accuracy results for this model are obtained with drop rates of 0 and 0.1. An increase in accuracy as the drop rate increases reflects an over-parametrised model in which the accuracy improves by setting some of the weights equal to zero to avoid overfitting. Conversely, a maximum accuracy at zero drop rate may indicate that the model is actually under-parametrised.

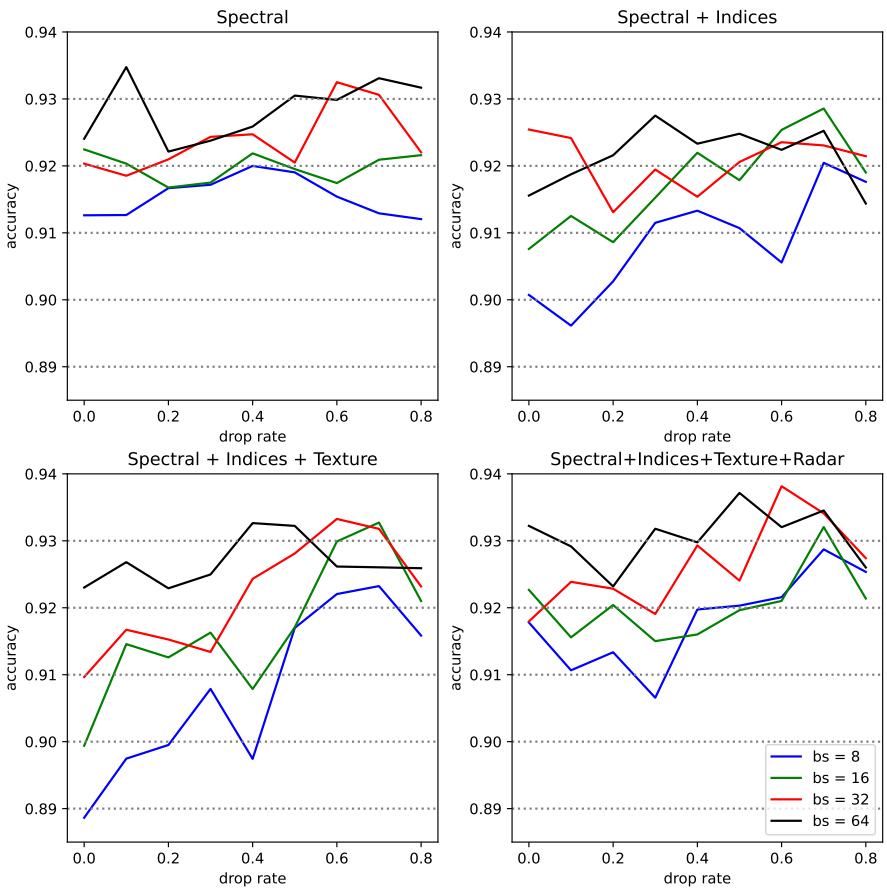

Figure 12. Validation results of the LSTM model with different batch sizes and drop rates.

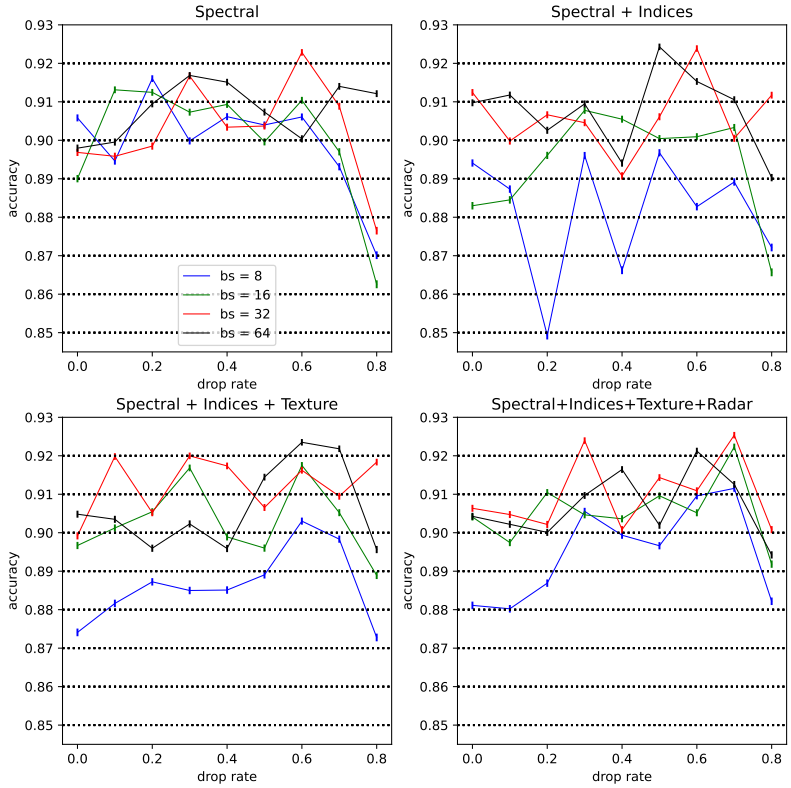

Figure 13. Test results of the LSTM model with different batch sizes and drop rates.

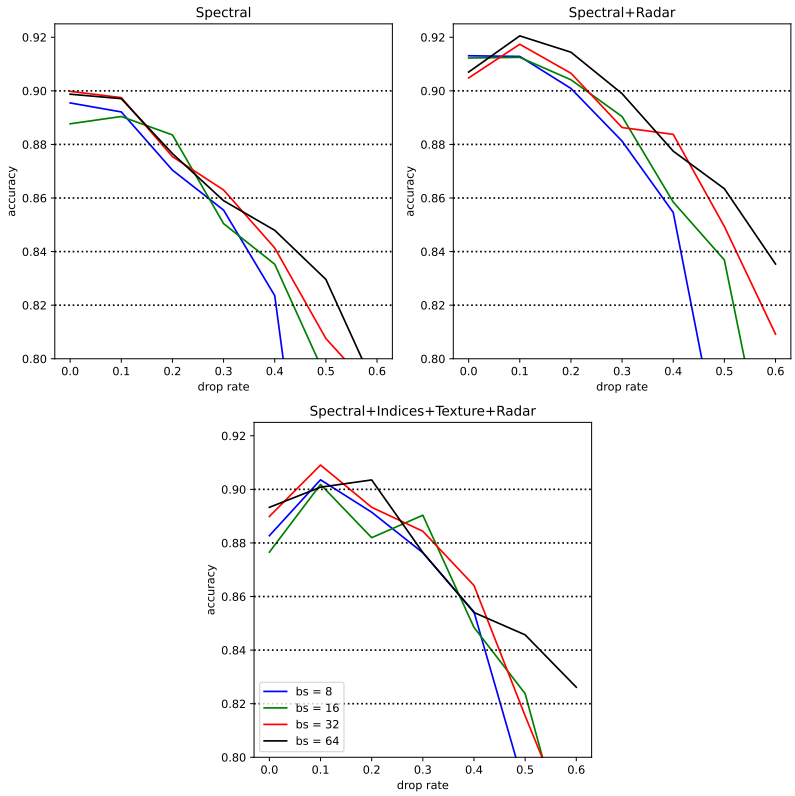

Figure 14. Validation results of the spectral–temporal LSTM model with different batch sizes and drop rates.

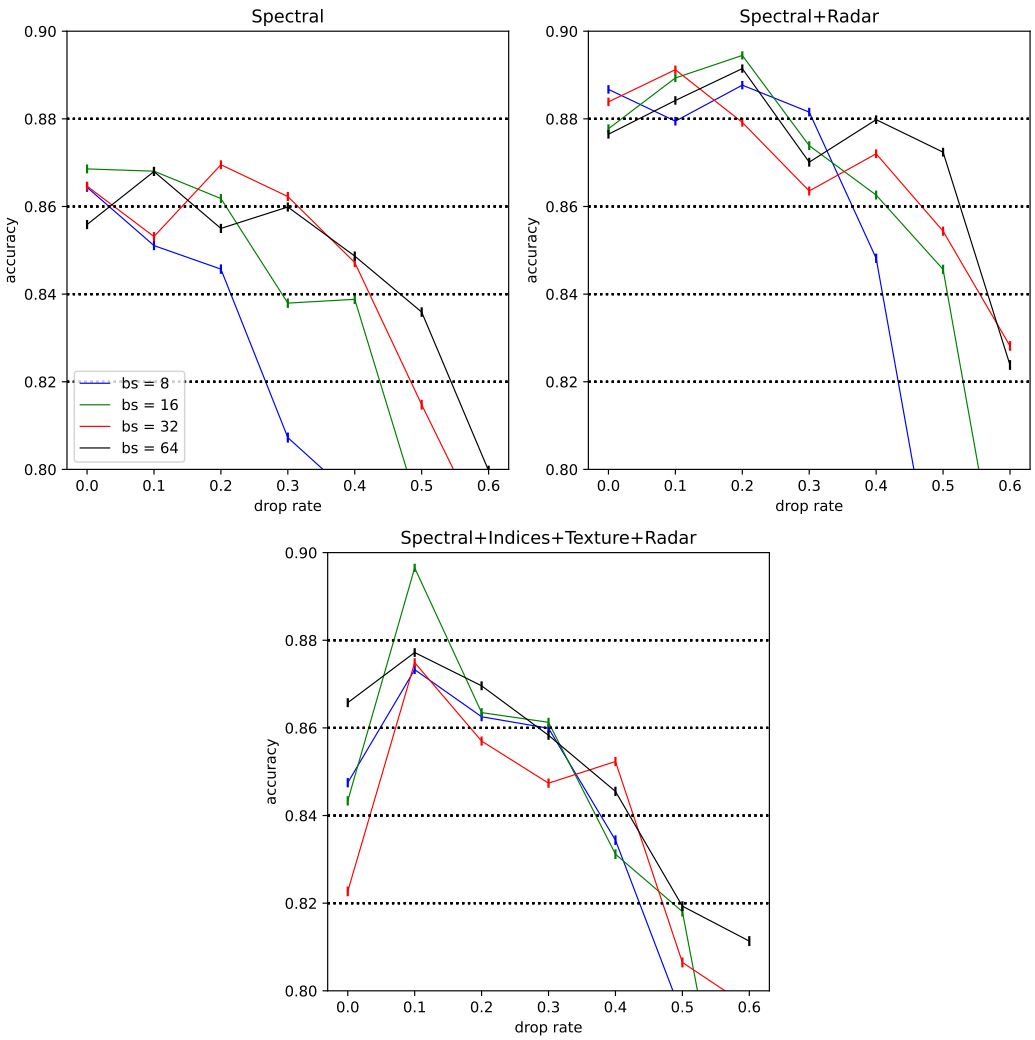

Figure 15. Test results of the spectral–temporal LSTM model with different batch sizes and drop rates.

5. Discussion

Currently, the use of deep learning continues to expand in remote sensing applications. There is a trend to consider applying different NN architectures to both LULC classification and the previous data preparation processes. The temporal dimension extracted from image time series would be compromised by missing information, mainly due to cloud contamination. Such a problem is commonly addressed by avoiding cloudy images as input to the models. However, this convention may result in the loss of some of the valuable information contained in the time series [45,46]. Although, in general, DL models are not able to work with missing or incorrect data, the LSTM cell seems to depict robust to low or moderate cloud cover in images [36,85].

There have been some attempts to use LSTM cells to fill in missing data information and not many attempts to use them in LULC classification. However, by applying LSTM-based NN, it is possible to overcome some of the common problems encountered when using traditional ML techniques. According to Ref. [52], therefore, the possibility of using a training, validation and test data set of a time series, in our case, eight images before and after the target image, and obtaining the training and test points directly from the unclouded pixels of the target image, avoids the problems of long time intervals between images [86].

Comparison of the results with previous work is limited by the different accuracy metrics used. Whereas some studies focus on classical regression statistics, such as R^2,

RMSE, or MAE [49,63,87], others focus more on computer vision metrics such as PNSR, SSIM or MAS [64–66]. As has been mentioned, in this study, we used R^2, RMSE and PNSR.

The CSTG method proposed by [87] achieves R^2 values between 0.86 and 0.93, with an average of 0.901 in the model they propose, whereas our proposed NN ranges between 0.637 and 0.995, with an average of 0.901 and RF ranges between 0.815 and 0.992, with an average of 0.939. Refs. [63,88] proposed the global-local loss function for gap filling. In the first step, they used the Fmask [89] and Multi-Scale Convolutional Feature Fusion (MSCFF) [90] methods to obtain cloud and shadow masks, with a buffer or two pixels at the edge to reduce the detection error. They use R^2 and RMSE as accuracy metrics, with r^2 ranging from 0.964 to 0.992 with an average of 0.981 and RMSE from 0.02 to 0.032 with an average of 0.0265. In our case, NN r^2 ranges from 0.637 to 0.995 with an average of 0.901, and RF ranges from 0.815 to 0.992 with an average of 0.939. RMSE ranges from 0.005 to 0.057 with an average of 0.0198 and with RF from 0.005 to 0.038 with an average of 0.0203. Ref. [66] reported a maximum PNSR value of 25.04 30.15 with average of 27.34 for cGAN and between 25.03 and 28.92 with average 27.17 for CLGN, without specifying bands. Ref. [65] reported PNSR values between 26.68 and 27.07 and an average of 26.78 without specifying bands. Ref. [64] reported PNSR values between of 22.894 for Pix2Pix and 26.186 for STGAN Res-Net without specifying bands. For these three studies where bands are not specified, we assume that the reported results are averages across bands.

Most of the PNSR values obtained in our study are larger. This does not necessarily mean that our model is better; the differences may be related to differences in the study area and the objectives of the research. The work of [64–66] is based on massive global data sets of Sentinel2 data. Training a model on global data is challenging. Our objective is less ambitious; we are trying to develop a method to fill small cloud patches in otherwise mostly cloud-free images on a small study area in order to enlarge the available data set for using LSTM networks to classify LULC. In addition, previous computer vision work has attempted to restore the whole image after modification, but our main aim is to use the restored image for other work, such as land cover classification, so we reject completely cloudy images as useless for our purposes.

On the other hand, we believe that RMSE is more useful than PSNR when the goal is to use the resulting images as input to a machine learning system. Human perception seems to respond logarithmically to light intensity [91]; so the $\log_{10}$ transformation in the PSNR equation provides an index that is more related to human perception of light intensity than RMSE. However, RMSE is more informative about the reflectance error that a cloudy image may introduce into a machine learning classification algorithm used to estimate LULC or some other environmental variable.

Among the few studies using LSTM for data filling and classification, ref. [53] compared five DL models for crop mapping with and without some training pixel information and showed how, by using a hybrid filling method composed of interpolation and a temporal filter, some architectures such as LSTM had a slight increase in overall accuracy with filled data, although they concluded that the improvement in accuracy was not worth the time spent reconstructing the series. We did not directly measure the improvement in this work, but comparing the accuracy results of the classifications obtained with RF with the previous work described in [31], where a maximum OA of 0.91 was obtained with a kappa of 0.89, we obtained an increase up to the 2% with reconstructed images. Regarding the classification method, both architectures with LSTM tested in [53] obtained lower accuracies among the others, conversely to ours, which obtained much higher values either LSTM or RF. However, both studies have similarities involving the use of optical data and LSTM as one of the methods applied, differences that made further comparison of results difficult.

In addition, several studies on the topic [21,92–96] found that the LSTM architecture can be successfully applied to increase the overall accuracy of time series sequence classification using single or multi-source data and data filled with many different methods.In relation to LULC classification, Ref. [93] proposed a deep learning architecture based on

the UNET model tested in 11 Sentinel-2 tiles using RF as the baseline model. Although the accuracy result of the DL model is higher, the differences between DL and RF are not that large. Ref. [95] integrated several bidirectional LSTM layers and used MLP and RF as baselines to classify six sites. Their results are very similar to both the former study and our study. The neural network model works better than the baseline models, but the differences are not particularly high. Ref. [42] achieved results of almost 0.91 of accuracy and 0.77 kappa with LSTM using medium resolution images, which where improved in [97]. Also using only optical images, ref. [98] achieved better results adding LSTM cells to other architectures of neural networks than using RF, as did the authors of [53]. Ref. [99] used SAR images and meteorological data for crop classification and achieved better OA results than RF when using time series. Hence, accuracies may be increased when using multi-source data [31,88,99]. In fact, our results point in the same direction, with the most accurate results achieved when the data set containing all the data from all the different sources are used together. It would also be interesting for future research to compare this approach with other network models, such as RCNN or GAN.

6. Conclusions

In terms of cloud removal, random forest and LSTM are clearly the best options; however, the former seems to be better for the March images; LSTM seems to be more accurate in summer; and both give similar results in April. In general, the models are more accurate in summer than in spring; this result is probably related to the higher availability of cloud-free images in summer. It is less certain, but also possible, that this is the reason why random forest achieves the same or better LSTM accuracy in spring. In any case, the differences between the different models are not significant in absolute terms, but they are significant in relative terms. Finally, the quality of the images produced is very good.

In relation to the final classification, the results obtained with random forest are similar to those obtained with the best RNN-based models. Random forest also has the advantage of being easy to use. In fact, it would be interesting to integrate the temporal evolution explicitly in RF, similar to the proposal of [100] to build a geographically weighted RF model. The model based on the spectral RNN did not give the expected results, although it was a priori an interesting model. Selecting a model based on validation results alone does not guarantee the best set of hyperparameters, although it may allow values to be obtained in the environment. In terms of combinations of data sets, it seems that the best option is to use all of them, although quite good results are also obtained using only spectral data.

In this particular case, it appears that spectral features alone are sufficient to achieve the highest accuracy. However, this is not something that could be assumed a priori. In any other case study, the level of accuracy achievable with different data combinations should be tested.

Comparing our results with others from similar studies, we think that it is not necessarily a good idea to train models with images from different biomes around the globe. The environmental differences may act as noise that the algorithms cannot process correctly, thus increasing the errors in the resulting images. This less ambitious approach allows to obtain accurate results even with simple, easy and quick to train models.

Author Contributions: Conceptualization, F.A.-S.; methodology, F.A.-S.; software, F.A.-S., C.V.-R. and F.G.-C.; validation, F.A.-S. and C.V.-R.; data curation, C.V.-R.; writing—original draft preparation, F.A.-S.; writing—review and editing, F.A.-S., C.V.-R. and F.G.-C.; funding acquisition, F.A.-S. and F.G.-C. All authors have read and agreed to the published version of the manuscript.

Funding: Grant TED2021-131131B-I00 funded by MICIU/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR.

Data Availability Statement: All source data can be downloaded from ESA websites.

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

Abbreviations

The following abbreviations are used in this manuscript:

CNN	Convolutional Neural Network
DL	Deep Learning
DPSVI	Dual Polarization SAR Vegetation Index
ESA	European Spatial Agency
GRD	Ground Range Detected
IW	Interferometric Wide
LSTM	Long Short Term Memory
LULC	Land Use and Land Cover
NDBI	Normalized Bald Index
NDVI	Normalized Vegetation Index
NDWI	Normalized Water Index
PNOA	Spanish Plan of National Ortophotography
RF	Random Forest
RNN	Recurrent Neural Network
RS	Remote Sensing
SAR	Synthetic Aperture Radar
TOA	Top Of the Atmosphere

References

1. Watson, R.; Noble, I.R.; Bolin, B.; Ravindranath, N.; Verardo, D.; Dokken, D. *Land Use, Land-Use Change and Forestry: A Special Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK, 2000.
2. Mason, P.; Manton, M.; Harrison, D.; Belward, A.; Thomas, A.; Dawson, D. The Second Report on the Adequacy of the Global Observing Systems for Climate in Support of the UNFCCC. Technical Report 82, 74, GCOS Rep. 2003. Available online: https://stratus.ssec.wisc.edu/igos/docs/Second_Adequacy_Report.pdf (accessed on 6 June 2024).
3. Naeem, S.; Cao, C.; Fatima, K.; Acharya, B. Landscape greening policies-based land use/land cover simulation for Beijing and Islamabad—An implication of sustainable urban ecosystems. *Sustainability* **2018**, *10*, 1049. [CrossRef]
4. Carranza-García, M.; García-Gutiérrez, J.; Riquelme, J. A Framework for Evaluating Land Use and Land Cover Classification Using Convolutional Neural Networks. *Remote Sens.* **2019**, *11*, 274. [CrossRef]
5. Wambugu, N.; Chen, Y.; Xiao, Z.; Wei, M.; Bello, S.; Junior, J.; Li, J. A Hybrid Deep Convolutional Neural Network for Accurate Land Cover Classification. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *103*, 102515. [CrossRef]
6. Yao, X.; Yang, H.; Wu, Y.; Wu, P.; Wang, B.; Zhou, X.; Wang, S. Land Use Classification of the Deep Convolutional Neural Network Method Reducing the Loss of Spatial Features. *Sensors* **2019**, *19*, 2792. [CrossRef] [PubMed]
7. Andrew, M.; Wulder, M.; Nelson, T. Potential contributions of remote sensing to ecosystem service assessments. *Prog. Phys. Geogr.* **2014**, *38*, 328–353. [CrossRef]
8. European Union. Commission implementing regulation (eu) 2018/746 of 18 May 2018 amending implementing regulation (eu) no 809/2014 as regards modification of single applications and payment claims and checks. *Off. J. Eur. Union* **2018**, *61*, 1–7.
9. Campos-Taberner, M.; García-Haro, F.; Martínez, B.; Izquierdo-Verdiguier, E.; Atzberger, C.; Camps-Valls, G.; Gilabert, M. Understanding deep learning in land use classification based on Sentinel-2 time series. *Sci. Rep.* **2020**, *10*, 17188. [CrossRef] [PubMed]
10. Yan, J.; Liu, J.; Wang, L.; Liang, D.; Cao, Q.; Zhang, W.; Peng, J. Land-Cover Classification With Time-Series Remote Sensing Images by Complete Extraction of Multiscale Timing Dependence. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 1953–1967. [CrossRef]
11. Zhong, L.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* **2019**, *221*, 430–443. [CrossRef]
12. Campos-Taberner, M.; Romero-Soriano, A.; Gatta, C.; Camps-Valls, G.; Lagrange, A.; Le Saux, B.; Tuia, D. Processing of extremely high-resolution lidar and RGB data: Outcome of the 2015 IEEE GRSS data fusion contest-part a: 2-d contest. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2016**, *9*, 5547–5559. [CrossRef]
13. Liu, T.; Abd-Elrahman, A.; Morton, J.; Wilhelm, V. Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. *GISci. Remote Sens.* **2018**, *55*, 243–264. [CrossRef]
14. Li, X.; Wu, X. Constructing Long Short-Term Memory based Deep Recurrent Neural Networks for Large Vocabulary Speech Recognition. *arXiv* **2015**, arXiv:1410.4281.
15. Murad, A.; Pyun, J.Y. Deep Recurrent Neural Networks for Human Activity Recognition. *Sensors* **2017**, *17*, 2556. [CrossRef] [PubMed]
16. Yan, R.; Liao, J.; Yang, J.; Sun, W.; Nong, M.; Li, F. Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering. *Expert Syst. Appl.* **2021**, *169*, 114513. [CrossRef]

17. Ndikumana, E.; Ho Tong Minh, D.; Baghdadi, N.; Courault, D.; Hossard, L. Deep recurrent neural network for agricultural classification using multitemporal SAR Sentinel-1 for Camargue, France. *Remote Sens.* **2018**, *10*, 1217. [CrossRef]
18. Mou, L.; Bruzzone, L.; Zhu, X. Learning Spectral-Spatial-Temporal Features via a Recurrent Convolutional Neural Network for Change Detection in Multispectral Imagery. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 924–935. [CrossRef]
19. Zhang, L.; Zhang, L.; Du, B. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 22–40. [CrossRef]
20. Zhu, X.; Tuia, D.; Mou, L.; Xia, G.; Zhang, L.; Xu, F.; Fraundorfer, F. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–36. [CrossRef]
21. Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 166–177. [CrossRef]
22. Simón-Sánchez, A.; González-Piqueras, J.; de la Ossa, L.; Calera, A. Convolutional Neural Networks for Agricultural Land Use Classification from Sentinel-2 Image Time Series. *Remote Sens.* **2022**, *14*, 5373. [CrossRef]
23. Gómez, C.; White, J.; Wulder, M. Optical remotely sensed time series data for land cover classification: A review. *Isprs J. Photogramm. Remote Sens.* **2016**, *116*, 55–72. [CrossRef]
24. Rogan, J.; Franklin, J.; Roberts, D. A comparison of methods for monitoring multitemporal vegetation change using Thematic Mapper imagery. *Remote Sens. Environ.* **2002**, *80*, 143–156. [CrossRef]
25. Xie, Y.; Sha, Z.; Yu, M. Remote sensing imagery in vegetation mapping: A review. *J. Plant Ecol.* **2008**, *1*, 9–23. [CrossRef]
26. Gómez, C.; White, J.; Wulder, M.; Alejandro, P. Historical forest biomass dynamics modelled with Landsat spectral trajectories. *ISPRS J. Photogramm. Remote Sens.* **2014**, *93*, 14–28. [CrossRef]
27. Franklin, S.; Ahmed, O.; Wulder, M.; White, J.; Hermosilla, T.; Coops, N. Large area mapping of annual land cover dynamics using multi-temporal change detection and classification of Landsat time-series data. *Can. J. Remote Sens.* **2015**, *41*, 293–314. [CrossRef]
28. Gomariz-Castillo, F.; Alonso-Sarria, F.; Cánovas-García, F. Improving Classification Accuracy of Multi-Temporal Landsat Images by Assessing the Use of Different Algorithms, Textural and Ancillary Information for a Mediterranean Semiarid Area from 2000 to 2015. *Remote Sens.* **2017**, *9*, 1058. [CrossRef]
29. Müller, H.; Rufin, P.; Griffiths, P.; Barros Siqueira, A.; Hostert, P. Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape. *Remote Sens. Environ.* **2014**, *156*, 490–499. [CrossRef]
30. Senf, C.; Leitão, P.; Pflugmacher, D.; van der Linden, S.; Hostert, P. Mapping land cover in complex Mediterranean landscapes using Landsat: Improved classification accuracies from integrating multi-seasonal and synthetic imagery. *Remote Sens. Environ.* **2015**, *156*, 527–536. [CrossRef]
31. Valdivieso-Ros, C.; Alonso-Sarria, F.; Gomariz-Castillo, F. Effect of the Synergetic Use of Sentinel-1, Sentinel-2, LiDAR and Derived Data in Land Cover Classification of a Semiarid Mediterranean Area Using Machine Learning Algorithms. *Remote Sens.* **2023**, *15*, 312. [CrossRef]
32. Aschbacher, J.; Milagro-Pérez, M. The European Earth monitoring (GMES) programme: Status and perspectives. *Remote Sens. Environ.* **2012**, *120*, 3–8. [CrossRef]
33. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [CrossRef]
34. Berger, M.; Moreno, J.; Johannessen, J.; Levelt, P.; Hanssen, R. ESA's sentinel missions in support of Earth system science. *Remote Sens. Environ.* **2012**, *120*, 84–90. [CrossRef]
35. Géron, A. *Hands-On Machine LEarning with Scikit-Learn, Keras, and TensorFlow*; O'Reilly: Sebastopol, CA, USA, 2019.
36. Hochreiter, S.; Schmidhuber, J. Long Short Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
37. Lyu, H.; Lu, H.; Mou, L. Learning a Transferable Change Rule from a Recurrent Neural Network for Land Cover Change Detection. *Remote Sens.* **2016**, *8*, 506. [CrossRef]
38. Ji, S.; Zhang, C.; Xu, A.; Shi, Y.; Duan, Y. 3D convolutional neural networks for crop classification with multi-temporal remote sensing images. *Remote Sens.* **2018**, *10*, 75. [CrossRef]
39. Zhao, H.; Chen, Z.; Jiang, H.; Jing, W.; Sun, L.; Feng, M. Evaluation of three deep learning models for early crop classification using sentinel-1A imagery time series—A case study in Zhanjiang, China. *Remote Sens.* **2019**, *11*, 2673. [CrossRef]
40. Liao, C.; Wang, J.; Xie, Q.; Baz, A.A.; Huang, X.; Shang, J.; He, Y. Synergistic use of multi-temporal RADARSAT-2 and Venμs data for crop classification based on 1D convolutional neural network. *Remote Sens.* **2020**, *12*, 832. [CrossRef]
41. Xue, R.; Bai, X.; Zhou, F. Spatial-temporal ensemble convolution for sequence SAR target classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 1250–1262. [CrossRef]
42. Rußwurm, M.; Korner, M. Temporal Vegetation Modelling Using Long Short-Term Memory Networks for Crop Identification from Medium-Resolution Multi-spectral Satellite Images. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21 July 2017; pp. 1496–1504.
43. Ruiz, L.; Almonacid-Caballer, J.; Crespo-Peremarch, P.; Recio, J.; Pardo-Pascual, J.E.; Sánchez-García, E. Automated classification of crop types and condition in a mediterranean area using a fine-tuned convolutional neural network. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *XLIII-B3-2020*, 1061–1068. [CrossRef]
44. Portalés-Julià, E.; Campos-Taberner, M.; García-Haro, F.; Gilabert, M. Assessing the sentinel-2 capabilities to identify abandoned crops using deep learning. *Agronomy* **2021**, *11*, 654. [CrossRef]

45. Che, Z.; Purushotham, S.; Cho, K.; Sontag, D.; Liu, Y. Recurrent Neural Networks for Multivariate Time Series with Missing Values. *Sci. Rep.* **2018**, *8*, 6085. [CrossRef] [PubMed]
46. Ng, M.K.P.; Yuan, Q.; Yan, L.; Sun, J. An Adaptive Weighted Tensor Completion Method for the Recovery of Remote Sensing Images with Missing Data. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3367–3381. [CrossRef]
47. White, J.; Wulder, M.; Hobart, G.; Luther, J.; Hermosilla, T.; Griffiths, P.; Coops, N.; Hall, R.; Hostert, P.; Dyk, A.; et al. Pixel-based image compositing for large-area dense time series applications and science. *Can. J. Remote Sens.* **2014**, *40*, 192–212. [CrossRef]
48. Skakun, S.E.A. Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2. *Remote Sens. Environ.* **2022**, *274*, 112990. [CrossRef]
49. Wang, Q.; Wang, L.; Zhu, X.; Ge, Y.; Tong, X.; Atkinson, P. Remote sensing image gap filling based on spatial-spectral random forests. *Sci. Remote Sens.* **2022**, *5*, 100048. [CrossRef]
50. Inglada, J.; Arias, M.; Tardy, B.; Hagolle, O.; Valero, S.; Morin, D.; Dedieu, G.; Sepulcre, G.; Bontemps, S.; Defourny, P.; et al. Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery. *Remote Sens.* **2015**, *7*, 12356–12379. [CrossRef]
51. Sadiq, A.; Sulong, G.; Edwar, L. Recovering defective Landsat 7 Enhanced Thematic Mapper Plus images via multiple linear regression model. *Iet Comput. Vis.* **2016**, *10*, 788–797. [CrossRef]
52. Chen, J.; Zhu, X.; Vogelmann, J.E.; Gao, F.; Jin, S. A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. *Remote Sens. Environ.* **2011**, *115*, 1053–1064. [CrossRef]
53. Zhao, H.; Duan, S.; Liu, J.; Sun, L.; Reymondin, L. Evaluation of Five Deep Learning Models for Crop Type Mapping Using Sentinel-2 Time Series Images with Missing Information. *Remote Sens.* **2021**, *13*, 2790. [CrossRef]
54. Pringle, M.J.; Schmidt, M.; Muir, J.S. Geostatistical interpolation of SLC-off Landsat ETM+ images. *Isprs J. Photogramm. Remote Sens.* **2009**, *64*, 654–664. [CrossRef]
55. Zhang, C.; Li, W.; Civco, D. Application of geographically weighted regression to fill gaps in SLC-off Landsat ETM+ satellite imagery. *Int. J. Remote Sens.* **2014**, *35*, 7650–7672. [CrossRef]
56. Makarau, A.; Richter, R.; Schlapfer, D.; Reinartz, P. Combined Haze and Cirrus Removal for Multispectral Imagery. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 379–383. [CrossRef]
57. Shen, Y.; Wang, Y.; Lv, H.; Qian, J. Removal of thin clouds using cirrus and QA bands of Landsat-8. *Photogramm. Eng. Remote Sens.* **2015**, *81*, 721–731. [CrossRef]
58. Shen, Y.; Wang, Y.; Lv, H.; Qian, J. Removal of Thin Clouds in Landsat-8 OLI Data with Independent Component Analysis. *Remote Sens.* **2015**, *7*, 11481–11500. [CrossRef]
59. Zhang, Y.; Guindon, B.; Cihlar, J. An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images. *Remote Sens. Environ.* **2002**, *82*, 173–187. [CrossRef]
60. Gladkova, I.; Grossberg, M.; Bonev, G.; Romanov, P.; Shahriar, F. Increasing the Accuracy of MODIS/Aqua Snow Product Using Quantitative Image Restoration Technique. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 740–743. [CrossRef]
61. Xin, Q.; Olofsson, P.; Zhu, Z.; Tan, B.; Woodcock, C.E. Toward near real-time monitoring of forest disturbance by fusion of MODIS and Landsat data. *Remote Sens. Environ.* **2013**, *135*, 234–247. [CrossRef]
62. Roy, D.P.; Ju, J.; Lewis, P.; Schaaf, C.; Gao, F.; Hansen, M.; Lindquist, E. Multi-temporal MODIS–Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sens. Environ.* **2008**, *112*, 3112–3130. [CrossRef]
63. Zhang, Q.; Yuan, Q.; Li, J.; Li, Z.; Shen, H.; Zhang, L. Thick cloud and cloud shadow removal in multitemporal imagery using progressively spatio-temporal patch group deep learning. *Isprs J. Photogramm. Remote Sens.* **2020**, *162*, 148–160. [CrossRef]
64. Sarukkai, V.; Jain, A.; Uzkent, B.; Ermon, S. Cloud removal from satellite images using spatiotemporal generator networks. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Snowmass Village, CO, USA, 1–5 March 2020; pp. 1796–1805.
65. Ebel, P.; Xu, Y.; Schmitt, M.; Zhu, X.X. SEN12MS-CR-TS: A Remote-Sensing Data Set for Multimodal Multitemporal Cloud Removal. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5222414. [CrossRef]
66. Sebastianelli, A.; Nowakowski, A.; Puglisi, E.; Rosso, M.P.D.; Mifdal, J.; Pirri, F.; Mathieu, P.P.; Ullo, S.L. Spatio-Temporal SAR-Optical Data Fusion for Cloud Removal via a Deep Hierarchical Model. *arXiv* **2022**, arXiv:2106.12226v3.
67. Berberoglu, S.; Curran, P.; Lloyd, C.; Atkinson, P. Texture classification of Mediterranean land cover. *Int. J. Appl. Earth Obs. Geoinf* **2007**, *9*, 322–334. [CrossRef]
68. Lasanta, T.; Vicente-Serrano, S. Complex Land Cover Change Processes in Semiarid Mediterranean Regions: An Approach Using Landsat Images in Northeast Spain. *Remote Sens. Environ.* **2012**, *124*, 1–14. [CrossRef]
69. Vanhellemont, Q. Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the Landsat and Sentinel-2 archives. *Remote Sens. Environ.* **2019**, *225*, 175–192. [CrossRef]
70. Wang, Z.; Bovik, A.; Sheikh, H.; Simoncelli, E. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef] [PubMed]
71. Alonso-Sarria, F.; Valdivieso-Ros, C.; Gomariz-Castillo, F. Isolation forests to evaluate class separability and the representativeness of training and validation areas in land cover classification. *Remote Sens.* **2019**, *11*, 3000. [CrossRef]
72. Klein, I.; Gessner, U.; Dietz, A.; Kuenzer, C. Global WaterPack—A 250 m resolution dataset revealing the daily dynamics of global inland water bodies. *Remote Sens. Environ.* **2017**, *198*, 345–362. [CrossRef]

73. Mostafiz, C.; Chang, N. Tasseled cap transformation for assessing hurricane landfall impact on a coastal watershed. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 736–745. [CrossRef]
74. Yang, X.; Qin, Q.; Grussenmeyer, P.; Koehl, M. Urban surface water body detection with suppressed built-up noise based on water indices from Sentinel-2 MSI imagery. *Remote Sens. Environ.* **2018**, *219*, 259–270. [CrossRef]
75. Hong, C.; Jin, X.; Ren, J.; Gu, Z.; Zhou, Y. Satellite data indicates multidimensional variation of agricultural production in land consolidation area. *Sci. Total Environ.* **2019**, *653*, 735–747. [CrossRef]
76. Rouse, J.; Haas, R.; Schell, J.; Deering, D. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; Technical report, Remote Sens. Center Tex. A&M Univ. Coll. Stn., A&M Univ. Coll. Stn.; NASA: Washington, DC, USA, 1973.
77. Kauth, R.; Thomas, G. The Tasselled-Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat. In Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, West Lafayette, IN, USA, 29 June–1 July 1976; Purdue University: West Lafayette, IN, USA, 1976; pp. 41–51.
78. Nedkov, R. Orthogonal transformation of segmented images from the satellite Sentinel-2. *Comptes Rendus L'Academie Bulg. Des Sci.* **2017**, *70*, 687–692.
79. Chen, X.; Zhao, H.; Li, P.; Yin, Z. Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. *Remote Sens. Environ.* **2006**, *104*, 133–146. [CrossRef]
80. Xu, H. Modification of normalized difference water index (NDWI) to enhanced open water features in remotely sensed imagery. *Int. J. Remote Sens.* **2006**, *27*, 3025–3033. [CrossRef]
81. Haralick, R. Statistical and structural approaches to texture. *Proc. IEEE* **1979**, *67*, 786–804. [CrossRef]
82. Hall-Beyer, M. Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. *Int. J. Remote Sens.* **2017**, *38*, 1312–1338. [CrossRef]
83. Filipponi, F. Sentinel-1 GRD preprocessing workflow. *Multidiscip. Digit. Publ. Inst. Proc.* **2019**, *18*, 11.
84. Rossiter, D. *Technical Note: Statistical Methods for Accuracy Assesment of Classified Thematic Maps*; Technical report; Department of Earth Systems Analysis International Institute for Geo-information Science & Earth Observation (ITC): Enschede NL, USA, 2004.
85. Podsiadlo, I.; Paris, C.; Bruzzone, L. A study of the robustness of the long short-term memory classifier to cloudy time series of multispectral images. In *Image and Signal Processing for Remote Sensing XXVI*; Notarnicola, C., Bovenga, F., Bruzzone, L., Bovolo, F., Benediktsson, J.A., Santi, E., Pierdicca, N., Eds.; SPIE: Bellingham, WA, USA, 2020; p. 59. [CrossRef]
86. Sun, L.; Chen, Z.; Gao, F.; Anderson, M.; Song, L.; Wang, L.; Hu, B.; Yang, Y. Reconstructing daily clear-sky land surface temperature for cloudy regions from MODIS data. *Comput. Geosci.* **2017**, *105*, 10–20. [CrossRef]
87. Wang, Y.; Zhou, X.; Ao, Z.; Xiao, K.; Yan, C.; Xin, Q. Gap-Filling and Missing Information Recovery for Time Series of MODIS Data Using Deep Learning-Based Methods. *Remote Sens.* **2022**, *14*, 4692. [CrossRef]
88. Avolio, C.; Tricomi, A.; Mammone, C.; Zavagli, M.; Costantini, M. A deep learning architecture for heterogeneous and irregularly sampled remote sensing time series. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Yokohama, Japan, 28 July 2019–2 August 2019; pp. 9807–9810. [CrossRef]
89. Zhu, Z.; Wang, S.; Woodcock, C.E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. *Remote Sens. Environ.* **2015**, *159*, 269–277. [CrossRef]
90. Li, Z.; Shen, H.; Cheng, Q.; Liu, Y.; You, S.; He, Z. Deep learning based cloud detection for medium and high resolution remote sensing images of different sensors. *Isprs J. Photogramm. Remote Sens.* **2019**, *150*, 197–212. [CrossRef]
91. Stevens, S.; Galanter, E. Ratio scales and category scales for a dozen perceptual continua. *J. Exp. Psychol.* **1957**, *54*, 377. [CrossRef] [PubMed]
92. Zhang, X.; Zhou, Y.; Luo, J. Deep learning for processing and analysis of remote sensing big data: A technical review. *Big Earth Data* **2022**, *6*, 527–560. [CrossRef]
93. Stoian, A.; Poulain, V.; Inglada, J.; Poughon, V.; Derksen, D. Land Cover Maps Production with High Resolution Satellite Image Time Series and Convolutional Neural Networks: Adaptations and Limits for Operational Systems. *Remote Sens.* **2019**, *11*, 1986. [CrossRef]
94. Parente, L.; Taquary, E.; Silva, A.P.; Souza, C.; Ferreira, L. Next Generation Mapping: Combining Deep Learning, Cloud Computing, and Big Remote Sensing Data. *Remote Sens.* **2019**, *11*, 2881. [CrossRef]
95. Xu, J.; Zhu, Y.; Zhong, R.; Lin, Z.; Xu, J.; Jiang, H.; Huang, J.; Li, H.; Lin, T. DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping. *Remote Sens. Environ.* **2020**, *247*, 111946. [CrossRef]
96. Machichi, M.A.; Mansouri, L.E.; Imani, Y.; Bourja, O.; Lahlou, O.; Zennayi, Y.; Bourzeix, F.; Houmma, I.H.; Hadria, R. Crop mapping using supervised machine learning and deep learning: A systematic literature review. *Int. J. Remote Sens.* **2023**, *44*, 2717–2753. [CrossRef]
97. Rußwurm, M.; Körner, M. Multi-Temporal Land Cover Classification with Sequential Recurrent Encoders. *ISPRS Int. J. Geo Inf.* **2018**, *7*, 129. [CrossRef]
98. Teloglu, H.K.; Aptoula, E. A Morphological-Long Short Term Memory Network Applied to Crop Classification. In Proceedings of the IGARSS 2022–2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022; pp. 3151–3154. [CrossRef]

99. Reuß, F.; Greimeister-Pfeil, I.; Vreugdenhil, M.; Wagner, W. Comparison of Long Short-Term Memory Networks and Random Forest for Sentinel-1 Time Series Based Large Scale Crop Classification. *Remote Sens.* **2021**, *13*, 5000. [CrossRef]
100. Georganos, S.; Grippa, T.; Gadiaga, A.; Linard, C.; Lennert, M.; Vanhuysse, S.; Mboga, N.; Wolff, E.; Kalogirou, S. Geographical random forests: A spatial extension of the random forest algorithm to address spatial heterogeneity in remote sensing and population modelling. *Geocarto Int.* **2021**, *36*, 121–136. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.