Article

Analysis of the Spatiotemporal Urban Expansion of the Rome Coastline through GEE and RF Algorithm, Using Landsat Imagery

Francesco Lodato , Nicola Colonna, Giorgio Pennazza, Salvatore Praticò, Marco Santonico, Luca Vollero and Maurizio Pollino

MDPI

Article

Analysis of the Spatiotemporal Urban Expansion of the Rome Coastline through GEE and RF Algorithm, Using Landsat Imagery

Francesco Lodato [1], Nicola Colonna [2] , Giorgio Pennazza [1] , Salvatore Praticò [3] , Marco Santonico [1], Luca Vollero [4] and Maurizio Pollino [2,*]

[1] Department of Science and Technology for Sustainable Development and One Health, Campus Bio-Medico University of Rome, 00128 Rome, Italy; f.lodato@unicampus.it (F.L.); g.pennazza@unicampus.it (G.P.); m.santonico@unicampus.it (M.S.)

[2] ENEA, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Casaccia Research Centre, 00123 Rome, Italy; nicola.colonna@enea.it

[3] Dipartimento di Agraria, Università Mediterranea di Reggio Calabria, Loc. Feo di Vito s.n.c., 89122 Reggio Calabria, Italy; salvatore.pratico@unirc.it

[4] Computational Systems and Bioinformatics Lab, Biomedical Engineering Faculty, Campus Bio-Medico University of Rome, 00128 Rome, Italy; l.vollero@unicampus.it

* Correspondence: maurizio.pollino@enea.it

Abstract: This study analyzes, through remote sensing techniques and innovative clouding services, the recent land use dynamics in the North-Roman littoral zone, an area where the latest development has witnessed an important reconversion of purely rural areas to new residential and commercial services. The survey area includes five municipalities and encompasses important infrastructure, such as the "Leonardo Da Vinci" Airport and the harbor of Civitavecchia. The proximity to the metropolis, supported by an efficient network of connections, has modified the urban and peri-urban structure of these areas, which were formerly exclusively agricultural. Hereby, urban expansion has been quantified by classifying Landsat satellite images using the cloud computing platform "Google Earth Engine" (GEE). Landsat multispectral images from 1985 up to 2020 were used for the diachronic analysis, with a five-yearly interval. In order to achieve a high accuracy of the final result, work was carried out along the temporal dimension of the images, selecting specific time windows for the creation of datasets, which were adjusted by the information related to the NDVI index variation through time. This implementation showed interesting improvements in the model performance for each year, suggesting the importance of the NDVI standard deviation parameter. The results showed an increase in the overall accuracy, being from 90 to 97%, with improvements in distinguishing urban surfaces from impervious surfaces. The final results highlighted a significant increase in the study area of the "Urban" and "Woodland" classes over the 35-year time span that was considered, being 67.4 km^2 and 70.4 km^2, respectively. The accurate obtained results have allowed us to quantify and understand the landscape transformations in the area of interest, with particular reference to the dynamics of urban development.

Keywords: change detection; Google Earth Engine; satellite imagery; machine learning; remote sensing; permeable areas

1. Introduction

Land cover change (LCc) mapping plays a crucial role in landscape and environmental planning and restoration [1]. LCc analysis represents key information to understand the relationship between human activities and environmental aspects [2,3]. Such studies assume an important value in the context of increasing soil exploitation and sealing [4], with alterations in the carbon cycle and a loss of ecosystem services [5]. According to the United Nations report "World Urbanization Prospects" [6], urban areas grow by about

20,000 football fields each day. Several groups, such as researchers, decision makers, and stakeholders, have recently showed a growing interest in urban LCc dynamics [7]. In the Italian context, the last decades have witnessed a process of strong urbanization, with the transformation of large areas with an agricultural vocation into residential, commercial, and industrial areas, leading to radical morphological and structural transformations [8].

A reliable tool for monitoring urban LCc dynamics is provided by remote sensing (RS) [3,9,10], operated by the various constellations of satellites that spatially and temporally cover vast portions of the Earth. Sexton et al. [11] processed an annual Landsat-based time series to monitor the urban growth of the Washington, D.C.—Baltimore, MD, metropolitan region. Hu and Ban [12] proposed an unsupervised algorithm to detect LCc in two large urban areas of Toronto (Canada) and Beijing (China) using satellite SAR data. Appiah et al. [13] proposed a satellite-based approach to analyze urban LCc in a peri-urban district of Ghana. Christensen and Mc Cord [14] exploited Landsat and MODIS satellite data to characterize China's urbanization dynamics. Mertes et al. [15] performed a MODIS-based time-series analysis that was coupled with enhanced vegetation index (EVI) observations to map the urban expansion in Southeast Asia. Zhou et al. [16] presented a new global urban change mapping approach based on night-time lights. Liu et al. [17] processed multi-temporal Landsat time-series mapping of global urban land dynamics. Huang et al. [18] used high-resolution commercial satellite WorldView images and deep learning algorithms to classify urban land uses. Gbanie et al. [19] coupled Landsat and SPOT satellite images to detect urban LCc trajectories in Sierra Leone. Reynolds et al. [20] used 20 years of Landsat-based time series and spectral indices to monitor the annual urban LCc in northwest Arkansas. Samal and Gedam [21] proposed an object-based image analysis approach to map the LCc that was associated with the urbanization process using Landsat and Indian remote sensing satellite multi-temporal images. Pandey et al. [22] processed a time series of MODIS satellite images to study the urban growth dynamic that was associated with agricultural land loss in India. Mahmoud and Divigalpitiya [23] processed 25 years of Landsat-based time series to monitor the spatiotemporal variation of urban expansion in an area in Egypt. Chai and Seto [24] mapped urban growth from Landsat time series to highlight micro-urbanization in Nigeria and the Democratic Republic of the Congo. These high-frequency temporal analyses need petabytes of data, commonly indicated by the neologism "Big-Data." Despite the large amount of data generated, optical satellite RS is affected by numerous operational problems, mainly related to cloud cover and to a considerable computational-time effort when trying to create a spatially and temporally extended cloud-free mosaic. To confront these kinds of problems, two different solutions are available. On one hand, there are cluster-based computing systems, which rely on the use of several computers in co-operation and, on the other hand, there are cloud-based computation platforms, which virtualize the infrastructures of supercomputers [25]. In this context, the Google Earth Engine (GEE) cloud computing platform [26] represents an effective benefit, through a user-friendly interface with the availability of numerous satellite and ancillary data catalogues [27,28], as well as a set of already implemented algorithms. It provides a viable and robust tool for pre- and post-processing of geospatial Big-Data, integrating GIS spatial analysis and management functions.

Currently, the GEE is the most widely used platform for geospatial data processing, and the number of academic publications is steadily increasing, with the main focus being on crop monitoring, water resources, and Land Use and Land Cover (LULC) mapping [29,30]. According to the review of Yang et al. [31], most of the GEE papers concern applications in areas of China, followed by the USA and Europe. In the specific Italian context, the GEE has been used in the framework of archaeological site monitoring [32,33], risk prevention and assessment [34], and crop and forest monitoring [35,36]. However, there is a lack of studies on LCc and, more specifically, on urbanization processes and dynamics. According to the review of Pérez-Cutillas et al. [37], the articles including urbanization analyses in the GEE environment are the least frequent, accounting for only 2% of all published papers, according to the review of Kumar and Mutanga [29]. Goldblatt et al. [38]

performed a pixel-based analysis in the GEE environment to monitor the urbanization processes in India. Sengupta et al. [39] proposed a study using the GEE and Landsat-5 satellite images to detect trajectories of coastal land reclamation in nine deltaic megacities. Zhang et al. [40] proposed a method that was based on Landsat-8 time series and vegetation indices in the GEE environment to map urban areas in three cities in China. Using the GEE code editor, Xue et al. [41] processed Landsat-5 and Landsat-8 time series to monitor the urban expansion of the mining city of Datong (China). Carneiro et al. [41] proposed a multi-source classification approach based on satellite images and the GEE environment to map urban land in a study area in Brazil.

The study that is described in the present paper is an improvement of previous work [4], further testing the potential of the new clouding tools that are available in this updated version. In particular, the capabilities of the GEE and the availability of Landsat imagery over a time span of about 40 years have been exploited to reconstruct the dynamics of the LCc in the North-Roman littoral. Using GEE aggregation functions [42] on each sampling year, it was possible to obtain a single cloud-free image composite, including all of the satellite data that were acquired in a specific time frame. This made the effects caused by randomness and seasonality less pronounced. In this specific case study, which was aimed at identifying urban surfaces, achieving maximum accuracy in the results was necessary, avoiding typical misclassifications due to the similarity between bare soils and urban soils. The authors, therefore, proposed integrating datasets with information related to the temporal variability of the pixels throughout the year, including the standard deviation of the NDVI index.

Thanks to the availability of cloud-based tools, operating simultaneously on multiple images that were organized in time series was possible. Such capability has allowed us to discriminate more efficiently the unchanged elements as urban, compared to soil, where changes during the year are related to water content or vegetation cover. Through machine learning algorithms, it has been possible to classify the study area into different land-use classes. Through diachronic analysis, we have evaluated the LCc, providing a significant improvement to the previous studies in terms of time interval and span.

2. Materials and Methods

2.1. Study Area

The study area (Figure 1) includes the following five municipalities that comprise the northern coastal fringe of Rome, located along the coastline of the Tyrrhenian Sea: Fiumicino, Ladispoli, Cerveteri, Santa Marinella, and Civitavecchia. They cover an area of about 500 km^2 along the N–S direction; starting from the south, with the municipality of Fiumicino near to the outlet of the Tiber River, to the northern limit, with the municipality of Civitavecchia bounded by the Mignone River. The area is mainly inside the boundary of the Metropolitan City of Rome Capital, an important area characterized by long-term landscape transformation [43].

The area comprises two distinct geomorphologic entities, a plain and a hilly sector. The plain sector extends from the municipality of Cerveteri to the Tiber River. In this area, it was necessary during the 20th Century to implement major land reclamation works [44] and, currently, this area is still mainly agricultural, with large portions devoted to arable crops in the northernmost area of the municipality of Fiumicino. The hilly sector is represented by the foothills of the Tolfa Mountains, which rise behind the municipalities of Cerveteri, Santa Marinella, and Civitavecchia. These areas, characterized by acid volcanism, are mainly represented by volcanic domes of low elevation [45], with marked slopes making them unsuitable for agricultural and urban purposes. Extensive coverage of Mediterranean maquis and broadleaf is still characterized by a coppice with a twenty-year cycle [46].
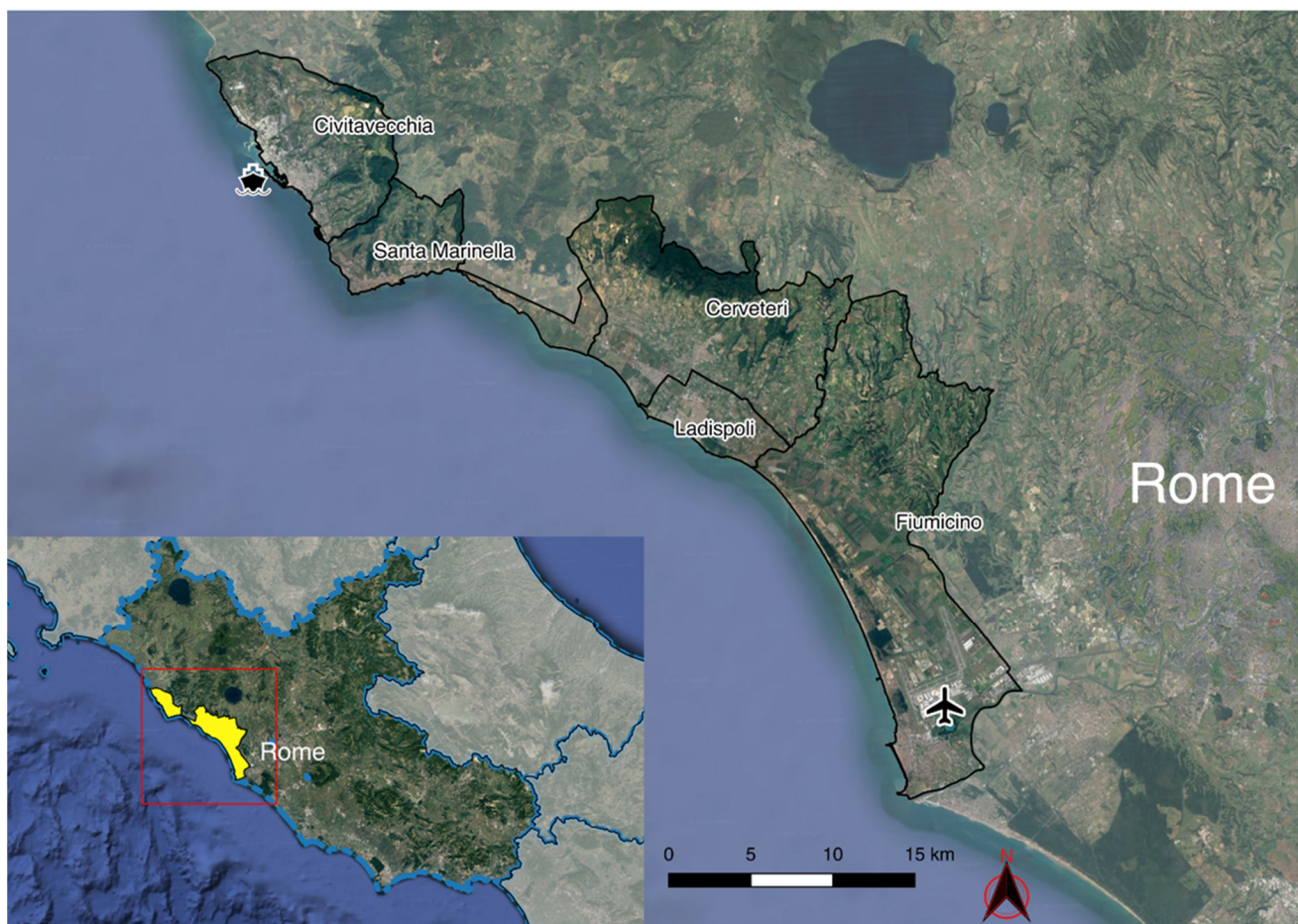
Figure 1. The geographic location of the five municipalities. In the bottom left: location of the survey area in central Italy, the area of interest (AOI) is highlighted in yellow.

The study area also includes important infrastructures, such as the "Leonardo Da Vinci" Airport (located in Fiumicino) and the harbor of Civitavecchia, which are connected to each other and to the city of Rome by a highway (A-12), a consular road (Via Aurelia SS1), and a railway (FL5 regional line), incentivizing urban development that is mainly linked to commuter movement. There is also a significant presence of coastal urban settlements that were built in the early 1970s to meet the demands of the seaside tourism sector linked to the city of Rome.

A polygon of administrative boundaries of the investigated municipalities was used for classification. We also considered a buffer of 20 m from the coastline, including the harbor areas.

2.2. Satellite Data

The choice of Landsat satellite images is impacted by the availability of a wide temporal coverage (1984–present) at a high and constant resolution (30 m). The images used in the present study are derived from surface reflectance data from Landsat-5 TM (1985, 1990, and 1995), Landsat-7 (2000, 2005, and 2010), and Landsat-8 OLI (2015 and 2020). The images have been imported as image collections in the GEE code editor. For each year of study, the data from all images surveyed in the phase of maximum vegetative vigor (from 05/01 to 09/30) were masked from the cloud through the quality assessment band and aggregated, resulting in a single median image using the *reducer* function of GEE. This time window coincides with the summer period, characterized by low levels of cloudiness in the survey area. Furthermore, this period falls in the deciduous trees' maximum vegetative vigor phenological phase, avoiding the misclassification of forested areas. Different dataset

accuracies were then evaluated by testing the classifiers with only the reflectance bands as the input (D1) and with the addition of two indices: *NDVI* (Equation (1)) [47] and *SwiRed* (Equation (2)) [48] (D2). These indices were selected based on their ability to distinguish between the macro-categories chosen for this study effectively.

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{1}$$

$$SwiRed = \frac{SWIR1 - RED}{SWIR1 + RED} \tag{2}$$

Finally, the model was tested by adding the NDVI standard deviation data (D3), calculated on a different image collection, based on the NDVI averages of each training pixel over an entire year (01/01 to 12/31) (Table 1). To avoid errors in the parameter calculation, the images with a cloud cover of more than 15% were removed from the collection. Using the SimpleCloudScore algorithm in the GEE environment, all pixels with cloud-likelihood scores were removed and >5 were masked.

Table 1. Composition of datasets used to train and validate the classification models.

	D1	D2	D3
05/01 to 30/09 (median)	BLUE GREEN RED NIR SWIR1 SWIR2	BLUE GREEN RED NIR SWIR1 SWIR2	BLUE GREEN RED NIR SWIR1 SWIR2
05/01 to 30/09 (median)		NDVI SwiRed	NDVI SwiRed
01/01 to 31/12			NDVI Standard Deviation
No. Variables	6	8	9

The hypothesis behind this parameter is related to the dynamism of agricultural land. In general, any area that is not an impervious cultivated area will have significant variations in NDVI throughout the year due to the tillage-crop cycle. In addition, bare lands not covered with vegetation may undergo variation in NDVI and spectral behavior due to varying water content throughout the year [49]. This parameter was therefore included to assess possible improvements in the performance of the classifier. In particular, this parameter was used to discriminate a stable surface area over time, such as urbanized surfaces, from a dynamic matrix, such as soils, a recurring problem in remote sensing due to the spectral similarity of these elements [50].

Therefore, the classifications reported here are no longer based on single images, but on a multitude of pixel-based information (Table 2), which the GEE allows to operate on the temporal dimension quickly and practically.

Due to the *filterBound* function, every image that partially or fully overlaps the study area is included in the collection; thus, the number of images in a collection can vary significantly due to the different orbits.

Table 2. Number of images in the collections for each year of sampling. The time span for the reflectance band and indices ranges from 05/01 to 09/30 of each sampling year, while the standard deviation of the NDVI was measured on all images (cloud cover < 15%) in the time span from 01/01 to 31/12 of the sampling year.

Year	1985	1990	1995	2000	2005	2010	2015	2020
05/01 to 09/30 (D1–D2–D3)	13	18	14	19	8	5 *	28	22
01/01 to 12/31 (NDVI st. Dev.)	11	17	16	19	12	6	25	36

* due to a shortage of images in 2010, the time span over which the medians for the spectral bands and indices were calculated was extended from 04/15 to 10/15.

2.3. Classification

Figure 2 depicts the overall workflow implemented and used in order to carry out the LCc analysis.
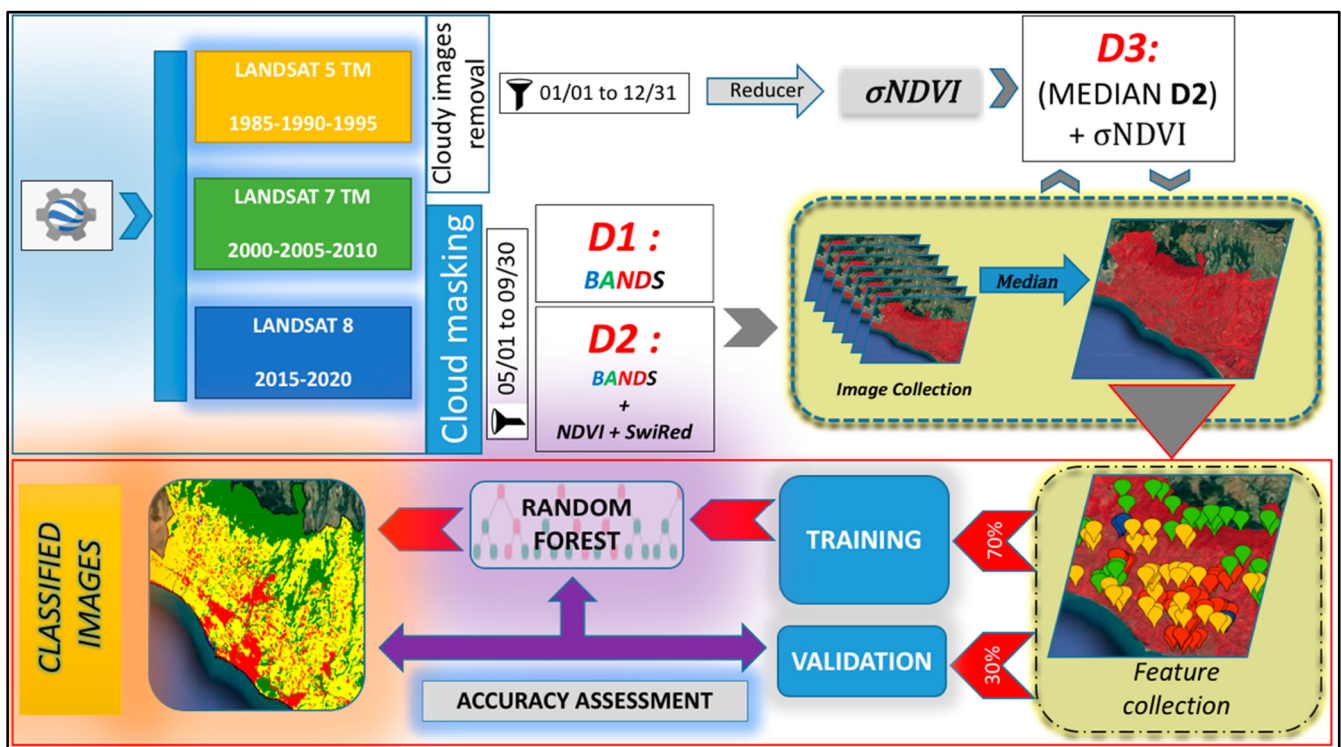

Figure 2. Image classification workflow for LCc analysis in GEE.

For the classification, considering the research goals, with a focus on urban expansion, the following four main macro categories were defined:

- Urban: artificial surfaces of various types, tiles and sheet metal roofs, asphalt roads, greenhouses, concrete, etc.;
- Woodland: area with dense and tall vegetation;
- Permeable Areas: non-impervious surface, bare soil, agricultural areas subjected to cyclical tillage, sparsely vegetated permeable areas, orchards, and seasonal or permanent crops. Considering the purposes of the present study, which focused on urban expansion, it was chosen to enclose all of those permeable areas that were purely agricultural, grazing lands, and wasteland in a single category;
- Water Bodies: reservoirs and rivers.

In the pixel-based stratified strategy, for each yearly median image, the sampling points were extracted as follows: *n* = 320 points were distributed for the macro classes "Ur-

ban", "Woodland", and "Permeable Areas"; and $n = 25$ for the class "Water Bodies", with approximately 1000 sampling points for each year of sampling (approximately 8000 sampling points in total), randomly selected on the AOI, to reduce spatial autocorrelation. The limited amount of "Water Bodies" sampling points is due to the spatially limited presence of this class in the survey area.

The sampling pixel points were chosen by integrating the historical images, land knowledge (through visual interpretation), and different false-color combinations. The samples were randomly divided, with 70% of the data used in training and 30% for validation, as a standard practice used in other similar work with satellite image classification [51] and machine-learning models [52].

Due to the best effectiveness of the random forest (RF) [53] model, as already demonstrated in similar studies [17,54,55], this classifier was chosen as a model for the subsequent LCc analysis.

The RF classifier is an ensemble model that uses many classification trees to form a forest. Each tree uses a different set of variables and observations for training using the "bagging" technique (*Bootstrap + Aggregation*), reducing the risk of overfitting [56]. In the GEE environment, it is possible to set several parameters, such as the number of decision trees forming the forest, the minimum leaf population, the proportion of training data used to create other trees, the randomization seed, and the number of variables used for node splitting.

About one-third of the samples are not used by the decision tree in the Bootstrap process. This sample is referred to as the *Out-Of-Bag Error* (OOBE), and its purpose is not to train the model, but to evaluate its performance. The OOBE decreases as the number of trees increases, until it stabilizes [57]. Another factor to be considered in the RF model is the importance of the variables, in which the mean decrease can express in the *Gini Coefficient* (MDGI) [58]. The MDGI indicates the importance of each variable in determining the node root split. The greater this value, the greater the variable's importance in the model.

The RF model was trained on 70% of the training data, with the number of decision trees set to $n = 100$ and the *mtry* parameter (number of variables randomly sampled as candidates at each split) set by default as the root square of the number of variables. These thresholds have already been successfully used in other satellite image classification works [42,59]. To evaluate the performance of the classifier and the suitability of the parameters, the OOBE was analyzed.

The following methodological step was assessing the importance of the variables in the classification process to understand the importance of each parameter in this classification task. In addition, the NDVI standard deviation data were analyzed by extracting the values of each pixel for each sample point collected. The data were exported to R, where the data distribution and any significant differences between the averages of the single class were analyzed through ANOVA and Tukey HDS (honestly significant difference) tests. ANOVA is a statistical test used to analyze mean differences between three or more data groups, while the Tukey HDS test compares all pairs of means. Our study used them to evaluate statistically significant mean differences between pairs of values.

Finally, the classifications were exported to the QGIS environment for further processing (sieving, graphical elaboration, etc.).

2.4. Accuracy Assessment

An accuracy analysis has been performed using 30% of the sample points as validation points to compare the goodness of the obtained results of each classification algorithm analyzed. The overall accuracy (OA), the user's accuracy (UA), the producer's accuracy (PA), and the kappa coefficient (K) have been adopted as accuracy indicators.

The OA, given as a percentage, is obtained by the ratio between the correctly classified pixels and the total pixel number [60]. The UA and PA refer to the classification accuracy calculated for each class. The UA is given by the ratio between the correctly classified pixel and the total pixels of a determinate class. The PA is calculated as the ratio between the

correctly classified pixels and the number of pixels selected as validation for a determinate class [61].

Another metric used to assess the accuracy of different models is Cohen's kappa coefficient (K) [62]. This index commonly assesses the agreement between predicted and observed categorical variables, such as land cover classes. This index ranges between -1 and $+1$, representing, respectively, the minimum and maximum agreement limits between the observed and predicted variables by the model. K values above 0.8 indicate a very good agreement between the predicted and observed variables.

As suggested by Momeni et al. [63], to compare the accuracies between different models and datasets using the same ground truth points, the McNemar test can be used. The McNemar test [64] is a non-parametric test that does not assume the normality of the data. It provides a chi-square test statistic (z) and its *p*-value, which allows the determination of a significant difference between the proportions of correctly and incorrectly classified pixels between the two models [65]. This study compared several datasets (D1-D2-D3) with the same ground truth for each sampling year. To apply the McNemar test (Equation (3)), the confusion matrices of each classifier were decomposed in correctly and incorrectly classified pixels; this results in comparing the two classifiers from a 2×2 contingency matrix, as follows:

$$z = \frac{b - c}{\sqrt{b - c}} \tag{3}$$

where b is the number of samples correctly classified by model 1, but misclassified by model 2, while c represents the number of samples correctly classified by model 2, but misclassified by model 1.

The accuracy assessment has been performed in the GEE environment using the already implemented accuracy function.

2.5. Land Cover Change (LCc) Analysis

The land cover change analysis was processed in the GEE. The classification maps produced in aster format for the overall time interval considered (1985–2020) were compared using logical function. As for urban expansion, logical functions were operated on each image based on the property of temporal irreversibility, evaluating changes in the urban areas that remain stable over time, referring to the last year considered in the diachronic analysis (2020).

3. Results

3.1. Model Accuracies

Tables 3–5 shows the OA, UA, PA, and K values for each sampling year and dataset for Landsat-5 (Table 3), Landsat-7 (Table 4), and Landsat-8 (Table 5).

Comparing the different accuracies of the datasets that were obtained from Landsat-5, D3 always shows a higher accuracy in terms of the OA and K; however, when comparing D1 and D2, both accuracies and K are variable. Regarding the PA and UA, we can see that D3 shows improvements, mainly in the classification of Permeable Lands and Urban classifications.

The accuracy differences are significant for D3 compared to D2 and D1 for 1985 and 1990, while the differences between the datasets in 1995 are not significant.

The classifiers show a high accuracy of results in general, with the OA ranging from 0.89 to 0.95.

Concerning the datasets that were obtained via Landsat-7, D3 shows a better performance than D1 and D2, in terms of both the OA and the K. In particular, when comparing D3 with D1 and D2, important improvements in terms of both the PA and the UA can be observed for the Urban and Permeable Lands classifications.

These differences between the datasets are significant, except for that between D2 and D3 for 2000 and 2005, and between D1 and D3 in 2005. Among the datasets for 2010, there are no statistical differences in accuracy.

The results generally show high accuracy, with the OA ranging from 0.88 to 0.95. Comparing the results of the different datasets that were obtained from Landsat-8, D3 always produces higher OA and K. In particular, it shows a higher accuracy in the Urban and Permeable Land classifications. The best performance of the D3 compared to the D1 and D2 datasets is statistically relevant (Table 6) and results in a very high OA (96%–97%).

Table 3. Accuracy metrics of the different datasets derived from Landsat-5.

	1985		Urban	Woodland	Permeable L.	Water B.	OA	K
D1		UA	0.86	0.96	0.85	1	0.89	0.84
		PA	0.87	0.95	0.85	1		
D2		UA	0.88	0.94	0.85	1	0.89	0.84
		PA	0.85	0.98	0.85	1		
D3		UA	0.96	0.99	0.9	1	0.95	0.92
		PA	0.91	0.98	0.95	1		
	1990		Urban	Woodland	Permeable L.	Water B.	OA	K
D1		UA	0.95	0.93	0.75	1	0.87	0.8
		PA	0.87	0.81	0.91	1		
D2		UA	0.94	0.92	0.86	1	0.91	0.87
		PA	0.95	0.92	0.86	0.8		
D3		UA	0.99	0.94	0.92	1	0.95	0.92
		PA	0.99	0.94	0.92	1		
	1995		Urban	Woodland	Permeable L.	Water B.	OA	K
D1		UA	0.93	0.88	0.86	1	0.89	0.84
		PA	0.97	0.9	0.8	0.83		
D2		UA	0.93	0.88	0.78	0.88	0.86	0.8
		PA	0.9	0.86	0.83	0.8		
D3		UA	0.97	0.81	0.9	1	0.9	0.85
		PA	0.95	0.95	0.85	0.8		

Table 4. Accuracy metrics of the different datasets derived from Landsat-7.

	2000		Urban	Woodland	Permeable L.	Water B.	OA	K
D1		UA	0.94	0.91	0.9	1	0.92	0.87
		PA	0.91	0.99	0.85	1		
D2		UA	0.97	0.91	0.8	1	0.9	0.87
		PA	0.85	0.91	0.89	1		
D3		UA	0.99	0.95	0.91	1	0.95	0.92
		PA	0.95	0.96	0.94	1		
	2005		Urban	Woodland	Permeable L.	Water B.	OA	K
D1		UA	0.92	0.95	0.77	1	0.88	0.85
		PA	0.85	0.89	0.89	1		
D2		UA	0.96	0.96	0.83	1	0.91	0.88
		PA	0.88	0.9	0.94	1		
D3		UA	0.98	0.99	0.9	1	0.95	0.93
		PA	0.95	0.92	0.98	1		
	2010		Urban	Woodland	Permeable L.	Water B.	OA	K
D1		UA	0.95	0.95	0.83	1	0.91	0.88
		PA	0.91	0.87	0.92	1		
D2		UA	0.96	0.94	0.87	0.91	0.92	0.89
		PA	0.91	0.93	0.91	1		
D3		UA	0.99	0.94	0.85	0.84	0.93	0.92
		PA	0.98	0.89	0.9	1		

Table 5. Accuracy metrics of the different datasets derived from Landsat-8.

2015		Urban	Woodland	Permeable L.	Water B.	OA	K
D1	UA	0.95	0.94	0.85	1	0.92	0.88
	PA	0.9	0.96	0.86	1		
D2	UA	0.92	0.92	0.83	0.8	0.9	0.85
	PA	0.9	0.94	0.82	1		
D3	UA	1	0.94	0.95	1	0.96	0.94
	PA	0.99	0.96	0.93	1		
2020		Urban	Woodland	Permeable L.	Water B.	OA	K
D1	UA	0.95	0.93	0.89	1	0.92	0.9
	PA	0.93	0.93	0.9	1		
D2	UA	0.9	0.98	0.88	1	0.92	0.9
	PA	0.93	0.95	0.9	0.87		
D3	UA	0.96	0.96	0.97	1	0.97	0.96
	PA	1	0.98	0.95	0.85		

Table 6. McNemar test (only *p*-value) on differences in accuracy between pairs of classifiers based on the different datasets for each year. The values in bold indicate statistically significant differences with $p < 0.05$.

	1985	1990	1995	2000	2005	2010	2015	2020
D1-D2	0.96	0.19	0.36	0.43	0.21	0.77	0.5	1
D1-D3	**<0.05**	**<0.05**	089	0.2	**<0.05**	0.76	**<0.05**	**<0.05**
D2-D3	**<0.05**	**<0.05**	0.2	**<0.05**	0.05	0.76	**<0.05**	**<0.05**

After evaluating the accuracies of each dataset for each sampling year, it was decided to use the models with the highest OA for the final classifications.

3.2. RF Performance

Based on other similar works [51,66], the RF classifier was set with $n = 100$ trees and the default *mtry* as the square root of the number of variables.

Table 7 shows the OOBE value for each model. A low error can be observed for D3 in each sampling year, with the maximum error in 1995.

Table 7. The OOBE of RF models in each sampling year.

Year	1985	1990	1995	2000	2005	2010	2015	2020
D1	0.091	0.098	0.118	0.094	0.082	0.110	0.098	0.082
D2	0.085	0.096	0.116	0.083	0.082	0.078	0.078	0.064
D3	0.066	0.066	0.081	0.054	0.054	0.06	0.045	0.048

Figure 3 shows that the most significant variable within the RF model for all of the sensors and sampling years is the STDDEV variable, which represents the standard deviation of the NDVI index. Concerning the NDVI and the SwiRed indices, their importance within the RF models, relative to the individual reflectance bands, is slightly higher in the models that were derived from Landsat-5 and Landsat-8. In contrast, no discernible differences are apparent in the models that were derived from Landsat-7. Generally, the importance of the individual reflectance bands within the models is heterogeneous, but never exceeds that of the STDDEV.
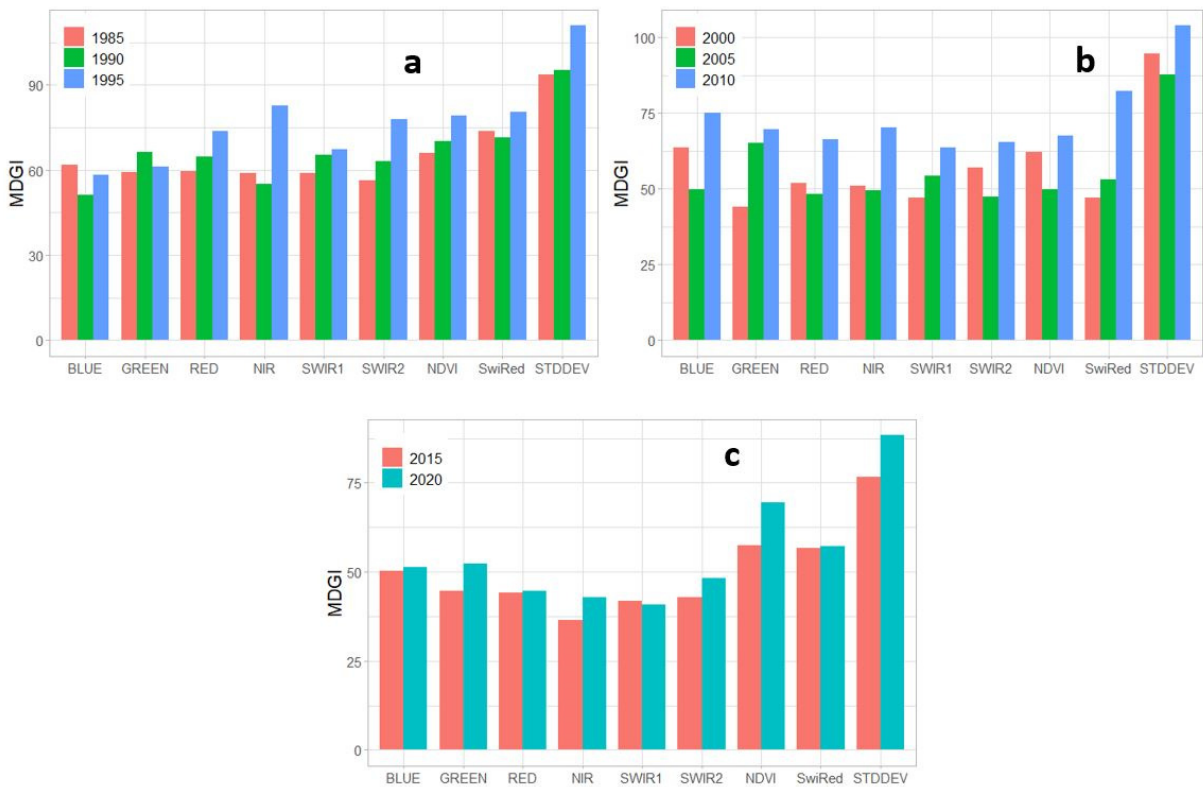
Figure 3. Median importance of the variable (MDGI) in the RF models, based on the D3 dataset, for Landsat-5 (**a**), Landsat-7 (**b**), and Landsat-8 (**c**).

3.3. NDVI Standard Deviation

Analyzing the values of NDVI standard deviation at all of the sampling points for each sampling year, sensor, and class, it is evident that the data averages differ significantly between the classes. The maximum values are observed in the Permeable Lands (PL) class, as shown in Figure 4 and Table 8.
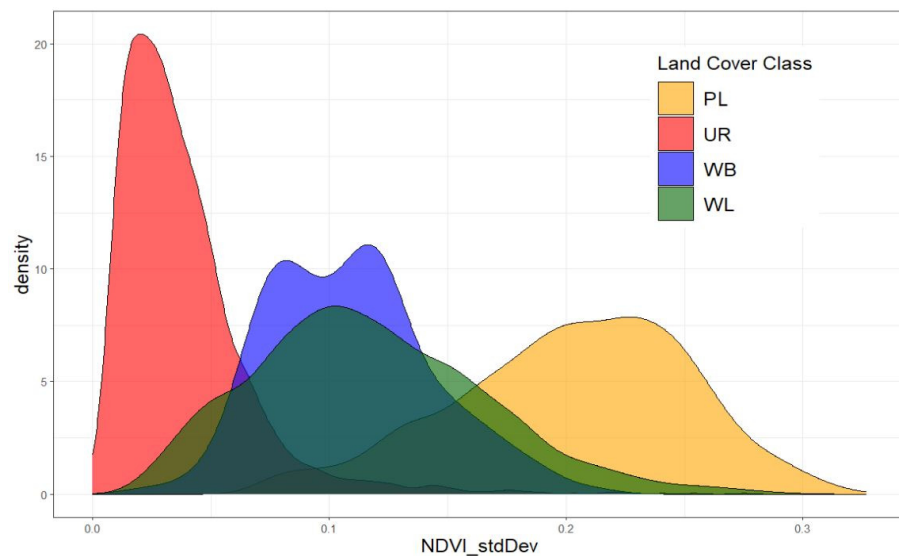

Figure 4. Density plot of the NDVI standard deviation parameter, the data referring to all sampling points in each sampling year.

Table 8. Tukey HDS on NDVI standard deviation of all sampling pixels in each year. The table shows the difference (*Diff.*) between the mean values of the classes, *lwr* and *upr* show the lower and upper limits of the confidence intervals, respectively, and *p*-value shows the adjusted *p*-value for the tests. Values in bold show *p*-value < 0.05.

	Diff.	*Lwr*	*Upr*	*p*-Value
UR-PL	−0.164	−0.166	−0.161	9.67×10^{-12}
WB-PL	−0.093	−0.100	−0.086	9.67×10^{-12}
WL-PL	−0.085	−0.087	−0.082	9.67×10^{-12}
WB-UR	0.070	0.063	0.077	9.67×10^{-12}
WL-UR	0.079	0.076	0.081	9.67×10^{-12}
WL-WB	0.008	0.001	0.015	0.078

The most substantial average difference is between the Urban (UR) and the PL classes. Although the differences in this parameter between the Woodland (WL) and the Water Bodies (WB) classes are small and statistically insignificant (Table 8), both of them exhibit significant differences with the PL and UR classes, displaying average values that fall between the UR and the PL class averages.

The analysis of the data distribution (Figure 4) reveals a degree of overlap between the Permeable Lands (PL) and the Urban (UR) classes and the Water Bodies (WB) and the Woodland (WL) classes. This finding is supported by Figure 5, which demonstrates a slight deviation from the normal distribution of the error, implying the presence of a subset with pronounced values of the NDVI standard deviation.
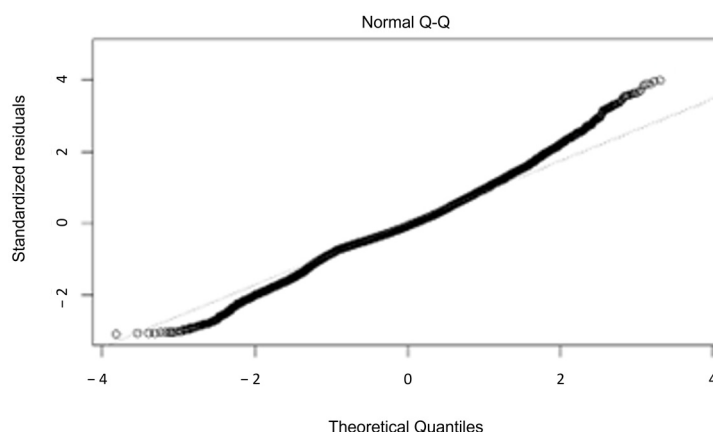
Figure 5. Quantile Plot of the residuals for the ANOVA model. The plot shows the distribution of the residuals of the ANOVA model compared to the normal distribution.

3.4. Classification

The land cover classifications that were obtained through the RF algorithm on the D3 datasets (Figures 6 and 7) show the evolution of the landscape over the examined time span.

The presence of a widespread wooded area in the inner hilly margins can be observed, growing over the years. Important expansions are also appreciable in the "Urban" classification class, with consolidation and expansion of historical centers being more accentuated in the areas that are close to the coast and generally in the proximity of the major roads and rail axis. This can be deduced from a constant reduction in "Permeable Land" areas, mainly within the municipalities of Cerveteri, Ladispoli, and Fiumicino. No water bodies of a significant extent were detected in the surveyed areas.
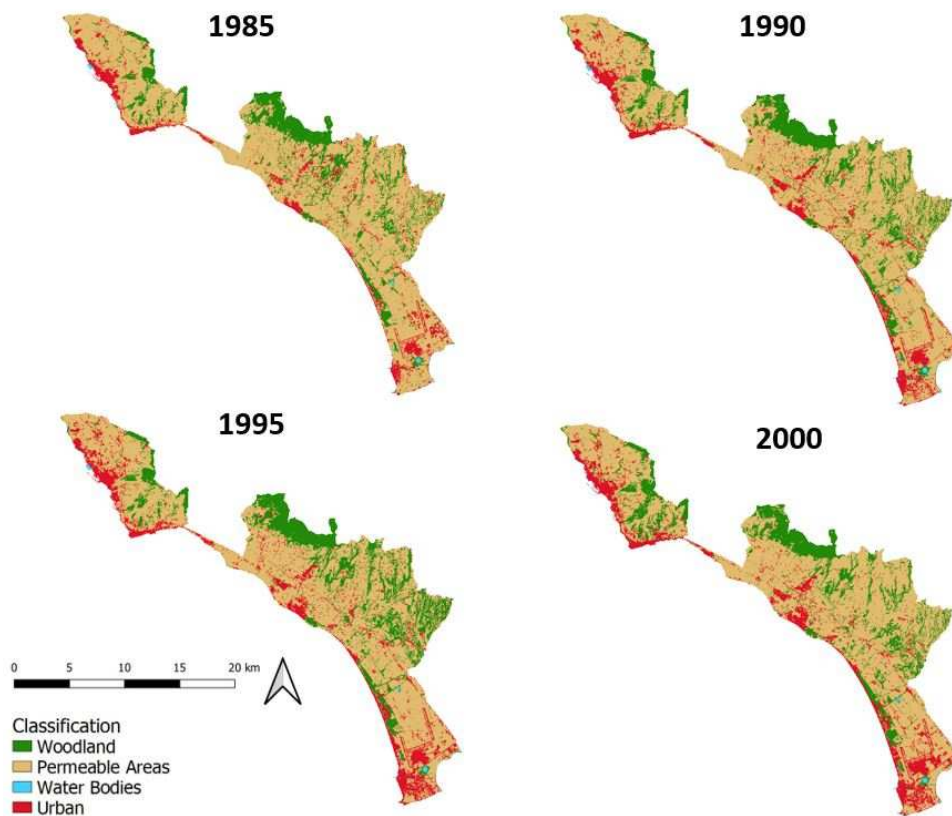
Figure 6. Land cover classification by RF algorithm. Years: 1985, 1990, 1995, and 2000. LC classifications were sieved in QGIS, threshold = 5 pixels, 4-bit connectivity.
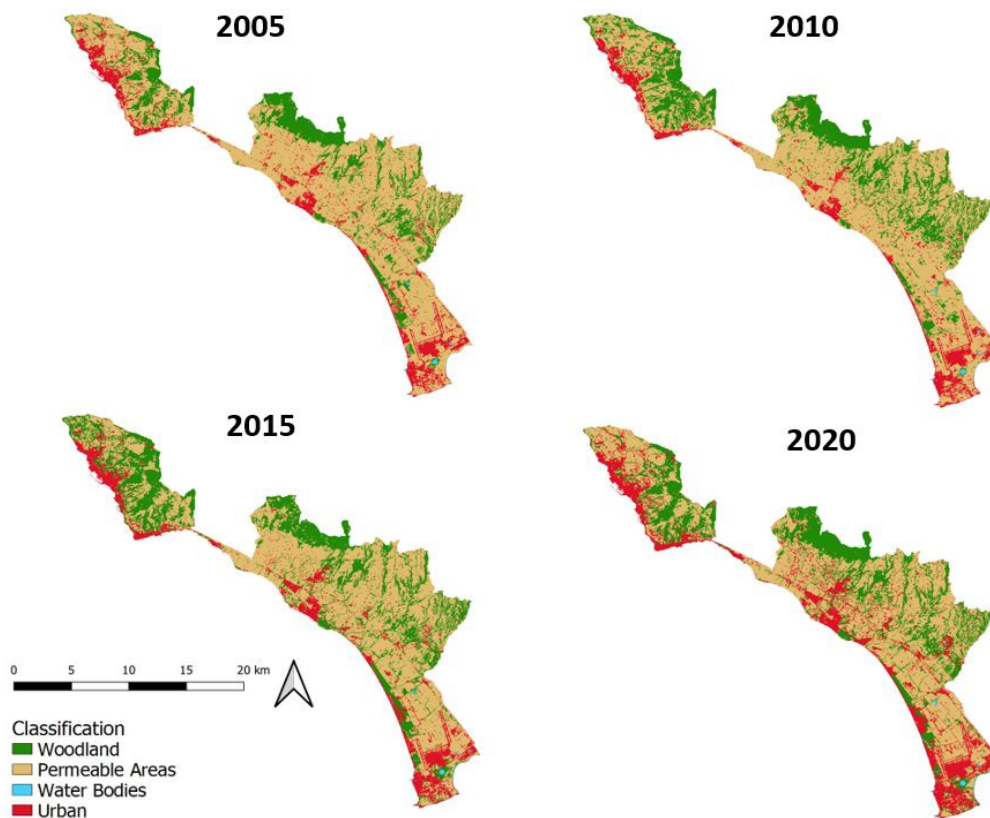

Figure 7. Land cover classification by RF algorithm. Years: 2005, 2010, 2015, and 2020. LC classifications were sieved in QGIS, threshold = 5 pixels, 4-bit connectivity.

3.5. Land Cover Change 1985 ÷ 2020

Based on the classifications that have been described in the previous section, and evaluating the changes in all of the land use classes, maps of the LCc were produced (Figures 8–11), with particular reference to urban expansion dynamics (Figures 12–14), for each sampling year. From 1985, the LC changes to the "Urban" class were considered to be stable in all of the subsequent classifications, until 2020, once they were detected. Other changes that were detected for the "Urban" class (i.e., from "Urban" to any of the other LC class) were considered to be a misclassification. Considering the dynamic of urban expansion and the geometric resolution of the Landsat images, a reverse trend from "Urban" to other classes has to be considered unlikely.
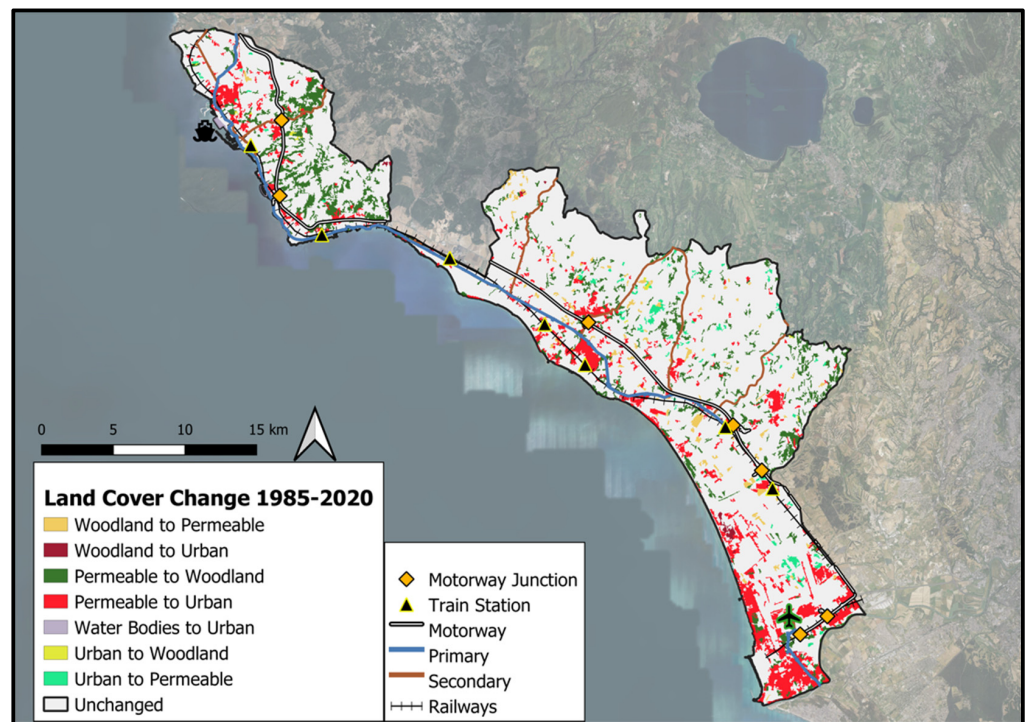
Figure 8. Variation and dynamics of LC within the study area for the total time span (1985 ÷ 2020), sieved image at a threshold of 5 pixels, 4-bit connectivity.

The LCc analysis (Table 9) shows a significant increase in the study area of the "Urban" and "Woodland" classes over the 35 years that were considered, at the expense of the "Permeable Areas" class, followed by the less marked LCc transitions from "Woodland" to "Permeable Areas" and to "Urban", and from "Urban" to "Permeable Areas." Changes related to the "Water Bodies" class were less significant, due to their small extension. In this case, the LCc detected was mainly related the construction of the Civitavecchia tourist harbor and expanding the main harbor dock.

Table 9. Land cover transition matrix during the analyzed time span (1985 ÷ 2020). Areas in km^2.

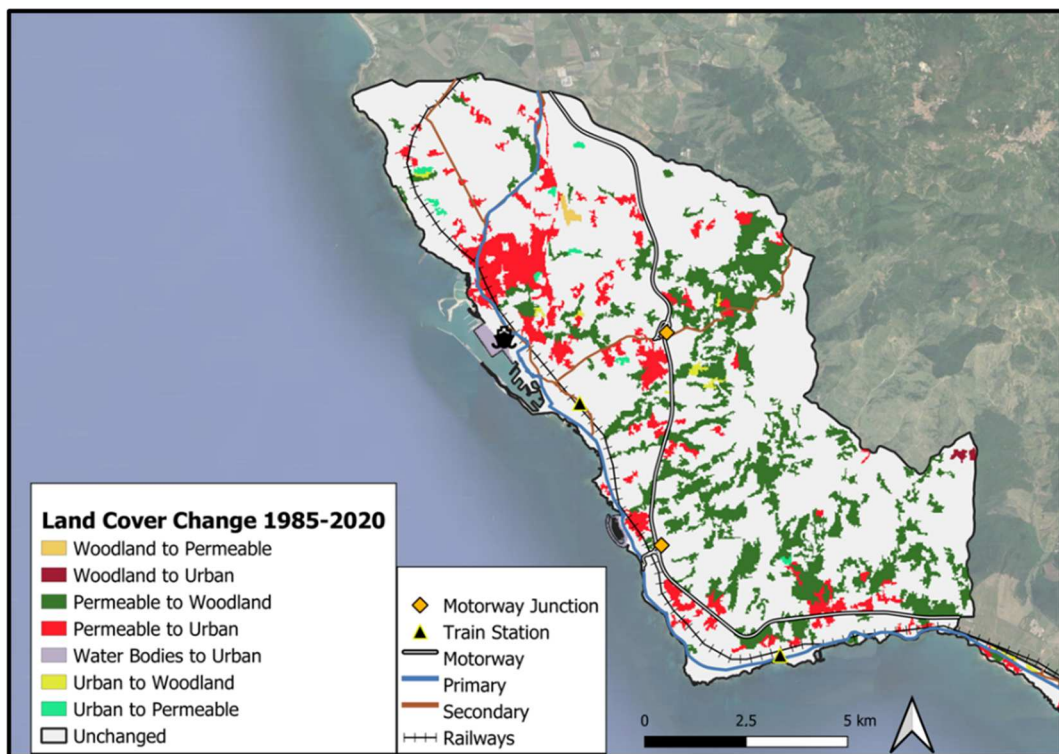| LC Transitions 1985–2020	Urban	Permeable Areas	Woodland	Water Bodies
Urban	-	11.5	1.2	<0.5
Permeable Areas	67.4	-	70.4	
Woodland	2	11.9	-	<0.5
Water Bodies	0.5	<0.5	<0.5	-

Figure 9. Variation and dynamics of LC within the Civitavecchia and Santa Marinella municipalities for the total time span (1985 ÷ 2020), sieved image at a threshold of 10 pixels, 4-bit connectivity.
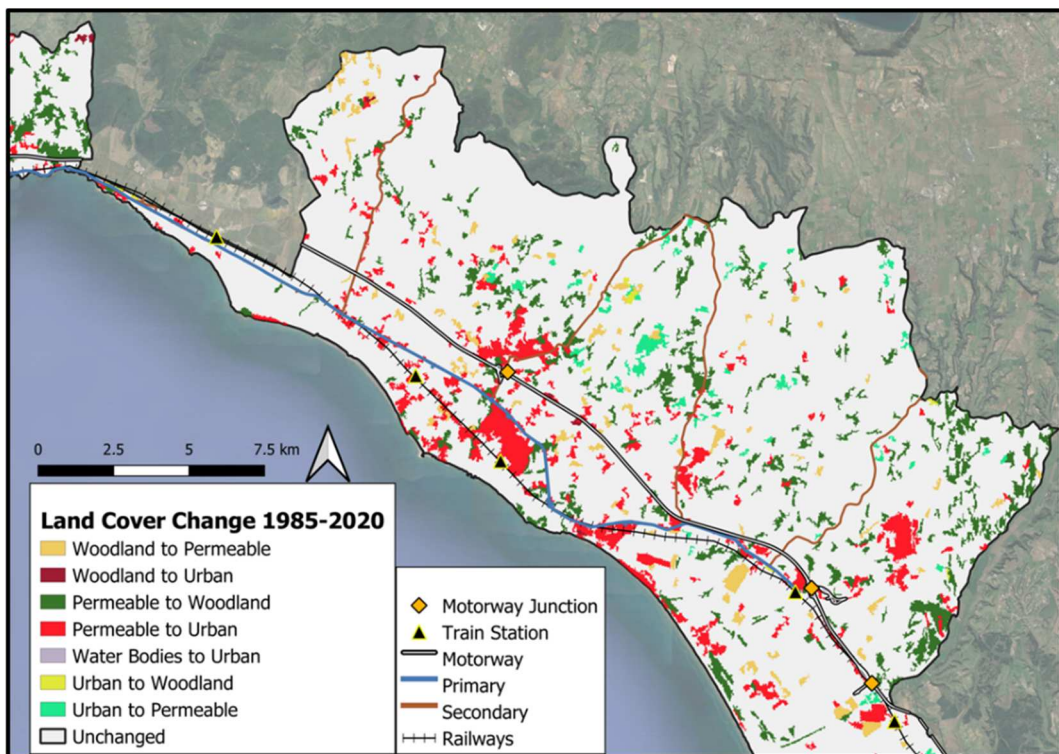

Figure 10. Variation and dynamics of LC within the Ladispoli and Cerveteri municipalities for the total time span (1985 ÷ 2020), sieved image at a threshold of 10 pixels, 4-bit connectivity.
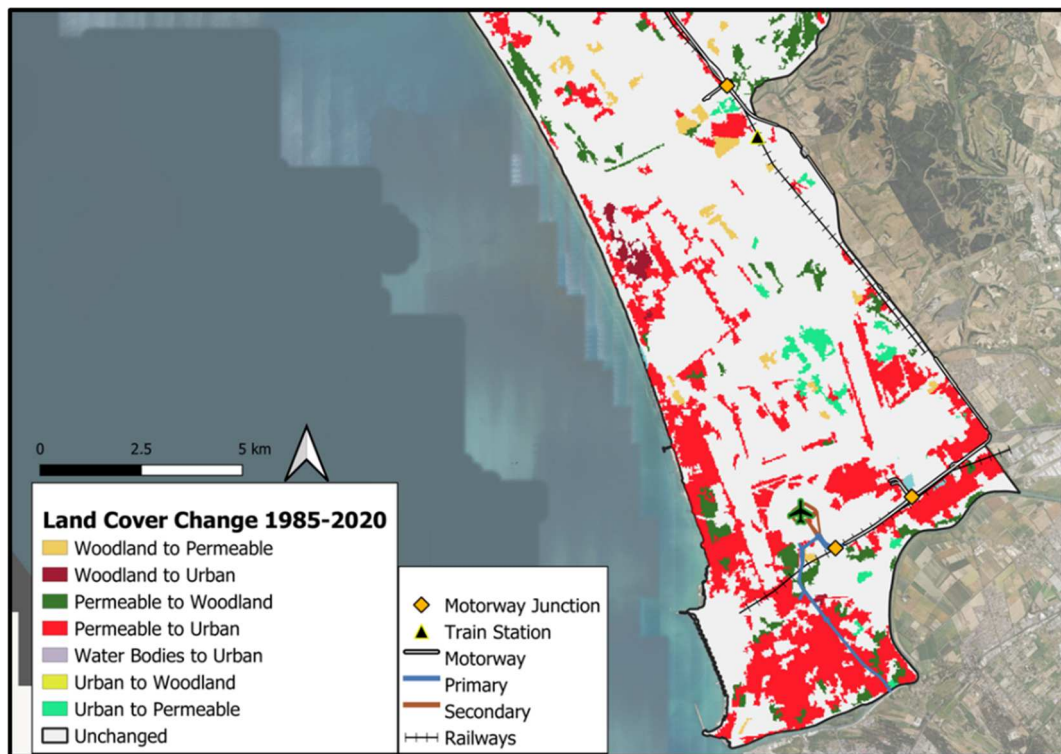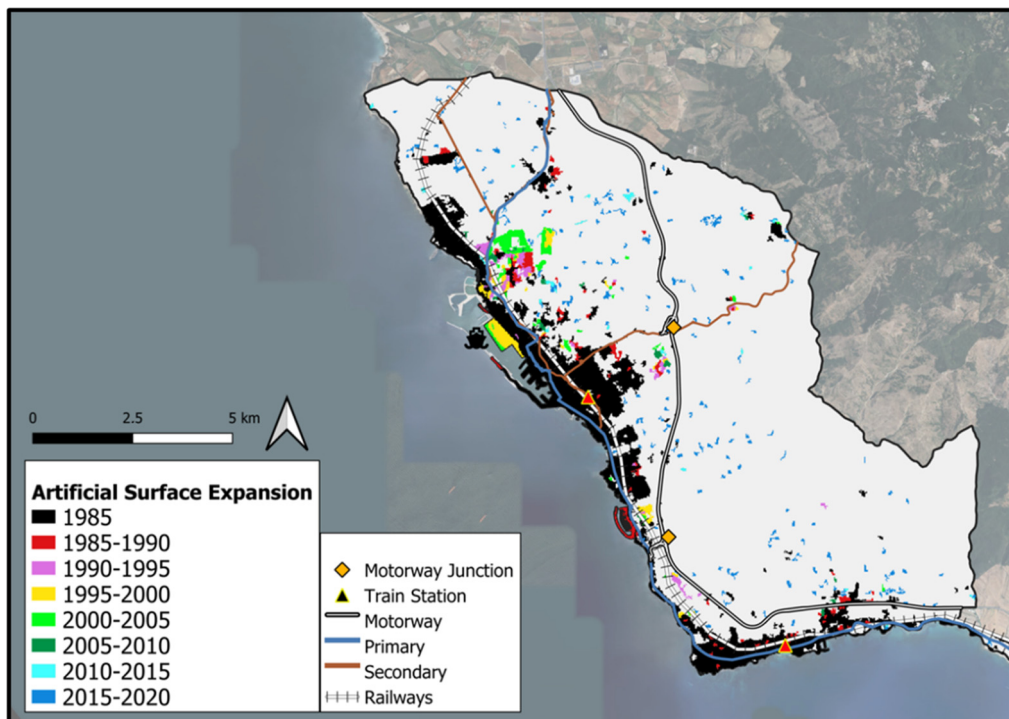
Figure 11. Variation and dynamics of LC within the Fiumicino municipality for the total time span (1985 ÷ 2020), sieved image at a threshold of 5 pixels, 4-bit connectivity.

Figure 12. Artificial surface evolution in the municipalities of Civitavecchia and Santa Marinella through the years. Image sieved at a threshold of 10 pixels, 8-bit connectivity.
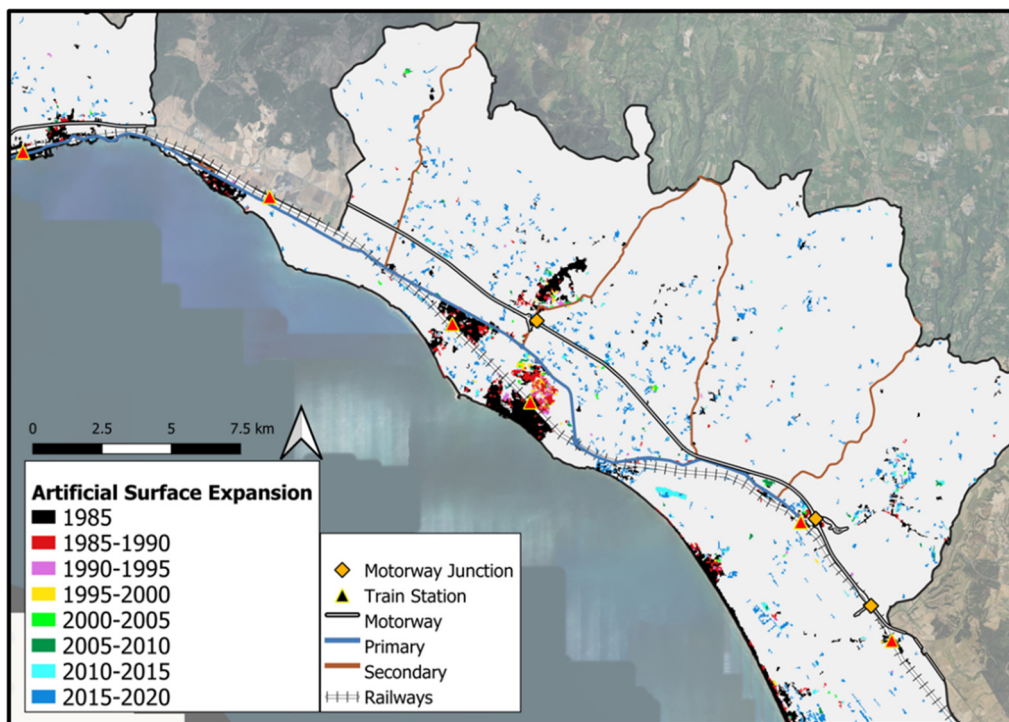
Figure 13. Artificial surface evolution in the municipalities of Cerveteri and Ladispoli through the years. Image sieved at a threshold of 10 pixels, 8-bit connectivity.
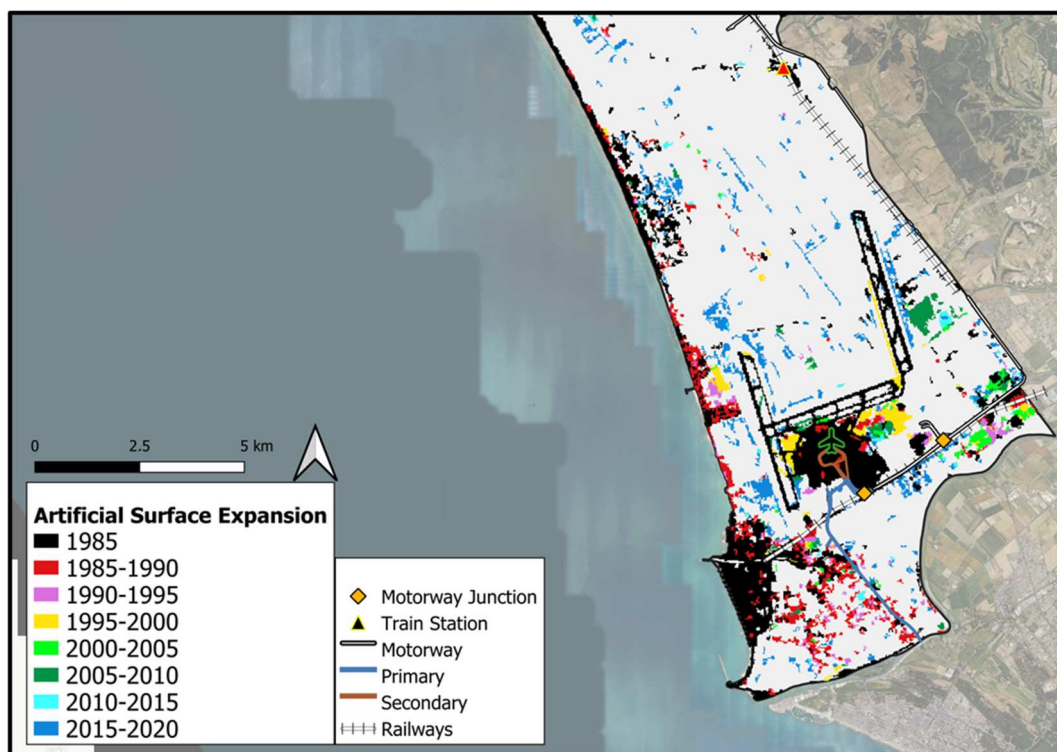

Figure 14. Artificial surface evolution in the municipality of Fiumicino through the 35 years analyzed. Image sieved at a threshold of 10 pixels, 8-bit connectivity.

The LCc map shown in Figure 8 shows an evident change in the areas close to the historical centers of the municipalities that were considered, with an expansion of dense urban areas near the transportation network. Overall, such a change to "Urban" is more

marked in the flat areas and those closer to the coastline. The urban sprawl phenomena can be observed in the inland areas, especially in the municipalities of Cerveteri and Fiumicino.

Significant reforestation processes can be observed in the inland hilly areas, especially in the northernmost areas of the municipalities of Civitavecchia and Santa Marinella, and more widely in the hinterland of the municipalities of Cerveteri and Fiumicino. There are a few changes from the "Urban" to the "Permeable Areas" class, especially in the plain between Cerveteri and Fiumicino.

Focusing on Civitavecchia and Santa Marinella (Figure 9) municipalities, a significant urbanization process has been detected in these areas at the expense of the "Permeable Areas" class, especially nearby to the coastline and along the "Aurelia" (SS1) consular road. A significant development of the built environment in the northernmost part of the Civitavecchia urban center can be noticed, resulting in a dense industrial agglomeration. The urbanization process has also characterized the municipality's inner part and the suburbs, with the construction of a hospital and new residential neighborhoods, respectively. On the other hand, in the hinterland's hilly areas, widespread reforestation processes can be noticed, with a significant development of the "Woodland" class at the expense of the "Permeable Areas" class, especially in the high-slope areas. The sprawling phenomena in the hilly hinterland are mainly related to isolated residential settlements.

In Cerveteri and Ladispoli (Figure 10), an important expansion of urban centers beside the main transportation lines (road and railway) can be observed, with a dense expansion of a mixed urban–industrial zone in the proximity of the Ladispoli railway station. An important urban expansion can be noticed in the minor villages of the two municipalities that have been mentioned above (Marina di Cerveteri and Marina di San Nicola), with newly built-up areas of an exclusively residential type. An intense, sprawling phenomenon can be observed in the hinterland zones and far from the inhabited centers, involving flat areas with an agricultural vocation. At the same time, an increase in the "Woodland" areas can be appreciated, involving the internal areas in a widespread manner. There are no major land use changes in the marginal areas, with large areas being covered by vegetation.

In the Fiumicino area (Figure 11), significant urbanization processes can be noticed at the expense of the "Permeable Areas" class, especially in the coastal residential zones (Fregene and Focene) and in the areas close to "Leonardo Da Vinci" Airport. Here, significant urbanization phenomena are linked to the development of both industrial and residential areas. Moreover, meaningful sprawling processes can be observed in the areas behind the historic center of Fiumicino. However, the area still keeps an extended zone with an agricultural vocation, falling in the village of Maccarese, with vast agricultural areas that are devoted to arable crops.

3.6. Spatio-Temporal Urban Expansion

Focusing on the spatio-temporal urban expansion, the following figures show the dynamic that was observed and mapped in the municipalities of Civitavecchia and Santa Marinella (Figure 12), Cerveteri and Ladispoli (Figure 13), and the Fiumicino area (Figure 14).

4. Discussions

The results showed excellent classifier accuracies in identifying the macro categories that were under investigation, with overall accuracies ranging from 0.86 to 0.97.

Compared to classifications using only reflectance bands, integrating indices such as NDVI and SwiRed led to a slight increase in accuracy, which was not statistically significant. However, the performance of the classifiers was significantly improved by adding the NDVI standard deviation as a variable, calculated on a separate collection of images based on the variability of this index for each pixel throughout the years.

This auxiliary variable leads to higher OA and K and is more important within the RF model, simultaneously reducing the OOB. In particular, this index seems to accurately differentiate between stable urban surfaces and soils that exhibit more dynamic spectral behavior during the year. Notably, this parameter shows a certain degree of similarity in

the pixels that were composed of soil, tall vegetation, and water. It is reasonable to assume that some of the pixels that were composed of deciduous forests may exhibit a specific variation in NDVI throughout the year, and the shallow water basins in the area may be affected by eutrophication phenomena. However, other variables, such as the median NDVI and SwiRed during the summer, can counterbalance this similarity in the RF model.

Despite the excellent accuracies that have been obtained, there are some limitations to consider in our methodology. The first concerns the use of specific indices, i.e., many indices in the literature for distinguishing urban areas were not evaluated and could lead to even higher accuracies. Another limitation is related to image composition. Although the GEE allows for masking and reducing images by mitigating the bias caused by cloud cover, some mainly cloudy periods could lead to a loss of information at specific phenological stages. For example, weather variability can lead to different vegetation responses in sparse areas and farmland. In addition, the choice of time windows for all of the datasets is in the function of the AOI, where snow cover is sporadic, vegetation cover is homogeneous in terms of diversity, and there are no particularly rocky or shaded areas. Thus, it is essential in future developments to establish specific time windows in order to evaluate changes in specific indices, depending on the spatial and radiometric resolution of the sensors and, obviously, on the purposes of the work.

Concerning the LCc analysis, in the study area that was characterized by a formerly agricultural vocation, an important expansion of artificial surfaces can be observed in the 1985–2020 time span, mainly affecting the flat areas, alongside the main roads and the railway, with a greater intensity in correspondence in the road junctions and the railway stations (Figure 8). The urbanization process was already in progress in the early years that were analyzed here, with significant densification between 1985 and 1995 involving all of the historic urban centers along the coastline. Here, an extensive "littoralization" process has happened. Initially, it was mainly linked to the pressure from the development of beach tourism, as in many other Italian coastal areas [67,68]. Then, it was replaced by another typology of use of the same areas, represented by the stable residence of owners or lodgers.

Considering the urban expansion, minor changes can be observed within Civitavecchia's urban center, which existed and consolidated before the surveyed period (Figure 12). However, a progressive expansion over time of industrial and commercial hubs has occurred in the area behind the harbor; moreover, near the highway junction, the construction of a hospital and major residential neighborhoods can be noticed. Such expansion can be observed along the coastline, which is related to the construction of a tourist port between 1985 and 1990 and the enlargement of the main harbor. The municipality of Santa Marinella has been affected by a phenomenon of urban consolidation, in which it is possible to notice a compaction of the urban texture during the 1985–1990 time interval. After this period, it has kept an unchanged urban layout in the subsequent years (Figure 12).

Important built-up expansions of the residential type can also be noticed in the municipalities of Ladispoli ("Cerreto" district) and Cerveteri ("Tyrsenia" district), developing beyond the historical aggregate and especially along the following major commute routes (Figure 13): the Rome–Civitavecchia railway line (FL5), Via Aurelia (SS1), and the A-12 motorway, which guarantee a fast connection with Rome and de facto make the four municipalities "residential" extensions of the capital [4]. The municipality of Ladispoli and the village of Marina di Cerveteri show the phenomena of urban consolidation in the 1985–1990 time interval, while the only relevant urban built-up area that was detected after 1990 is represented by the expansion of the "Cerreto" district of Ladispoli between 1990 and 1995. Then, in the area closest to the highway junction, the construction of a hospital and an industrial zone, with new man-made structures and warehouses, can be noticed. Overall, the municipality of Cerveteri shows a slight expansion outside of the urban center, mainly in the direction along the A-12 highway, represented by the "Tyrsenia" neighborhood. Recent sprawling phenomena can be observed in these two municipalities' peripheral areas.

In particular, urban evolution can be observed in typical seaside districts, such as Marina di Cerveteri and Santa Severa, which modified their functionality. In the 1970s and 1980s, they were characterized by a floating population (temporary rents for summer vacations), whilst from the late 1990s, the commuters, benefiting from transportation connections, started to rent the dwellings there, as has typically happened in many metropolitan areas [69].

In the Fiumicino area (Figure 14), a major expansion of artificial surfaces in the innermost areas and along the "Roma—Fiumicino" highway can be noticed. Such expansion is a mix of industrial and commercial areas, aggregates in many cases, and large residential settlements far from the main urban center, which remains unchanged, except for a few scattered residential zones that built up in the early 1990s. Finally, the expansion of the airport area can be observed through the years, with a gradual growth of the surrounding artificial surfaces (Figure 14).

Thus, the northern (Civitavecchia) and the southern (Fiumicino) parts of the investigated area can be associated with a similar dynamic. The largest expansions of artificial surfaces after 1995 are mainly related to the construction of extended industrial and commercial areas. This trend can be linked to the development of the main infrastructures (such as the "Leonardo Da Vinci" Airport and the Civitavecchia Harbor) and the ex novo construction of large malls ("Parco Leonardo" and "Da Vinci Market" near Fiumicino and "La Scaglia" near Civitavecchia) surrounded by a mixed zone with industrial activities (e.g., warehouses) and new residential neighborhoods close to major roads.

Urban sprawl has occurred mainly at the expense of permeable areas, represented by all of the zones with an agricultural vocation or, a potential one, which can negatively impact ecosystem services. Moreover, significant reforestation processes have been identified at the expense of the agricultural areas in the more inaccessible hilly zones. This is probably due to the reduced pressure on the wooded areas, which were once intensively exploited for livestock farming, agriculture, and silviculture. This trend aligns with the findings of other studies that have been conducted in the same area [70,71]. Those activities generally result in a clear degradation of areas that were once mainly devoted to the fruitful orchard- and vineyard-based cropping systems, and now are affected by strong fragmentation processes that are to be ascribed to the intense, sprawling phenomena. However, extensive coverage of agricultural land in the "Maccarese plain" is still present, and it is possible to observe the widespread phenomena of "Land Concentration" [72,73], with large parcel systems devoted to arable monoculture.

As reported in Table 9, a significant extension of areas that were classified as "Urban" in 1985 were subsequently classified as "Permeable Areas" in 2020. This finding can be explained by some rural modifications and dynamics related to specific agricultural activities. In particular, a progressive dismantlement of greenhouses and covers for crop protection (i.e., tunnels) has occurred in the study area over the years. Those areas were originally (1985) classified as "Urban" (scilicet as artificial surfaces). Due to the transformations that have been mentioned above of rural activities, those areas were later (2020) classified as "Permeable." However, misclassifications may have influenced the LCc results, since the data represent a direct comparison between the two classifications.

In addition, considering the seasonal conditions and the phenological phases of vegetation and crops, mixed areas (cultivated/shrublands) increased in the peri-urban zones, which, at the beginning of the 1990s, were mainly covered with ploughed soil for arable crops. These areas were progressively transformed because of urban sprawling. These border zones show an LC mixture, where the gardens and the residual farmland areas are not clearly distinguishable, similarly to the locally fragmented pastures, which are often abandoned [74]. In this latter case, observing a contained re-naturalization process is also possible, especially in the most inaccessible areas. This trend, also observed in other studies in the same or similar areas [75,76], can lead to the disappearance of the rural landscape heritage.

Hydraulic works and infrastructures have also contributed to the land transformation of the study area. Clay-rich soils characterize the plain of Cerveteri and Ladispoli, which was once covered by broadleaf forests and marshes. During the 1950s, the reclamation of swamped areas began and, consequently, many areas were converted to agricultural uses. Then, the expansion of the cultivated areas was mainly related to the realization of the "Consorzio di Bonifica Litorale Nord" aqueduct (www.consorziobonificalitoralenord.it—last accessed: 13 May 2022), which has served the countryside areas of Cerveteri and Ladispoli since 2002, both for domestic (garden irrigation) and agricultural uses. This has favored the introduction of water-demanding crops to exploit the arable land, such as maize and cucurbitaceous. At the same time, such water availability has guaranteed an adequate supply to orchards and vineyards. Some small oases (such as Macchiatonda, Palo, and Torre Flavia) are the residuals of those pre-existing natural areas, currently classified as special protection areas according to regional legislation [77].

Finally, as described in the previous section, an increase in woodland has been detected. The woods are in the innermost and hilly zones of the study area, in which the morphological and pedological conditions have drastically limited the agricultural exploitation or, more in general, the land use modifications. Moreover, other environmental or archaeological constraints have indirectly preserved the wooded areas. Illustrative of this is the forest expansion near the archaeological site of the Etruscan Necropolis of Banditaccia [78], especially after 2004, when it was inscribed in the UNESCO World Heritage List (WHL).

As an ending point, it is worth highlighting how the free clod-platform environment allows worldwide dissemination of the results, favoring the participation of researchers, decision makers, and local communities [79].

5. Conclusions

Regarding the methodology, it has been shown that, for the classification of land use into macro categories, the standard deviation of the NDVI variable, when tailored to the specific context, can significantly improve model accuracies, notably in discriminating urban surfaces from dynamic surfaces such as agricultural soils and, generally, all pervious surfaces.

The authors propose evaluating the use of variables related to indices variation in specific time intervals. This can be carried out quickly and effectively using cloud platforms that manage vast amounts of geospatial data. Although this work focuses on urban areas expansion, it is interesting to note a possible extension of this methodology to more detailed classifications, using variations in specific temporal ranges of multiple indices. The high accuracy of the results has made it possible to highlight significant changes in land use, following well-defined patterns.

Compared with other available datasets, such as the Copernicus Urban Atlas (https://land.copernicus.eu/local/urban-atlas, accessed on 18 March 2023), the peculiarity of the approach that has been described in the present article lies in better spatial and temporal resolutions. Specifically, it has been possible to process and analyze a well-defined temporal range of 40 years, while the availability of the Urban Atlas data is more limited in terms of the temporal intervals that are available (2006, 2012, and 2018). Considering the spatial resolution, the minimum mapping unit (MMU) of the Urban Atlas is 0.25 ha for urban classes and 1 ha for rural classes. On the contrary, the Landsat spatial resolution of 0.09 ha has allowed for a more effective analysis of the urban sprawl phenomenon within the study area. Nevertheless, it is worth considering the availability of CORINE data as potential reference for validating the classifications.

The LCc were mapped and analyzed in the area that was located in the coastal area between Civitavecchia and Rome. We focused on the medium-sized Civitavecchia and four other small-sized municipalities (Santa Marinella, Cerveteri, Ladispoli, and Fiumicino) that, in a few decades, have shown a relevant urban expansion and have largely increased their population (by doubling their inhabitants).

In the case study that has been discussed in the present paper, it has been possible to confirm the presence of a landscape dynamic that is common to many Italian suburban realities; areas that were once devoted to typically rural activities, respectful of the natural cycles and times, are now transformed de facto into suburbs of large cities that are suitably connected each other, thus incentivizing a residential use related to commuting.

Such an urbanization phenomenon seems to have peaked before the 2000s, with a consolidation of urban cores. At the same time, in the subsequent years, the development of artificial areas was mainly related to services, with the expansion of the main infrastructures and the construction of facilities (e.g., hospitals and large shopping malls) to meet the needs of the growing population.

Within this framework, a loss of the purely rural functionality of these areas is evident, along with non-negligible reforestation in areas that are less suitable for urban development and phenomena of concentrated urbanization and sprawling in the surrounding countryside.

The land degradation neutrality goals (SDGs 15) for 2030 seem far from being achieved; however, better knowledge and comprehension of the soil consumption dynamic could help local and regional administration to deal with this. New tools and data availability become powerful instruments to effectively monitor policy goals on a smaller temporal scale and increase the awareness of how important it is to protect soil from imperviousness and abandonment processes. Soil and environmental protection policies that were focused on natural parks or specially protected areas failed to manage the zones outside of the protected ones and avoided the soil consumption processes.

In future developments, novel approaches and techniques will be investigated and implemented in order to improve the classification of mixed areas (especially in those zones where sparse urbanization permeates agricultural/natural areas). This will allow us to overcome the limitations that have been mentioned above, which were observed during the study described here. To this end, high spatial/spectral resolution images, jointly with spectral mixture techniques and texture analysis, can adequately support the abovementioned goals. Moreover, integrating different data, such as SAR and night-time light sensors, can lead to better classification performances. This could result in a more detailed classification, for example, between deciduous and evergreen forests or to distinguish certain types of artificial coverage.

References

1. Pandey, P.C.; Koutsias, N.; Petropoulos, G.P.; Srivastava, P.K.; Ben Dor, E. Land use/land cover in view of earth observation: Data sources, input dimensions, and classifiers—A review of the state of the art. *Geocarto Int.* **2021**, *36*, 957–988. [CrossRef]
2. Modica, G.; Praticò, S.; Di Fazio, S. Abandonment of traditional terraced landscape: A change detection approach (a case study in Costa Viola, Calabria, Italy). *Land Degrad. Dev.* **2017**, *28*, 2608–2622. [CrossRef]

3. Modica, G.; Vizzari, M.; Pollino, M.; Fichera, C.R.; Zoccali, P.; Di Fazio, S. Spatio-temporal analysis of the urban–rural gradient structure: An application in a Mediterranean mountainous landscape (Serra San Bruno, Italy). *Earth Syst. Dyn.* **2012**, *3*, 263–279. [CrossRef]

4. Pollino, M.; Lodato, F.; Colonna, N. Spatio-Temporal Dynamics of Urban and Natural Areas in the Northern Littoral Zone of Rome. In *Computational Science and Its Applications—ICCSA 2020*; ICCSA 2020 Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2020; Volume 12253, pp. 567–575.

5. Calzolari, C.; Tarocco, P.; Lombardo, N.; Marchi, N.; Ungaro, F. Assessing soil ecosystem services in urban and peri-urban areas: From urban soils survey to providing support tool for urban planning. *Land Use Policy* **2020**, *99*, 105037. [CrossRef]

6. Available online: https://www.un.org/development/desa/publications/2018-revision-of-world-urbanization-prospects.html (accessed on 18 March 2023).

7. Reba, M.; Seto, K.C. A systematic review and assessment of algorithms to detect, characterize, and monitor urban land change. *Remote Sens. Environ.* **2020**, *242*, 111739. [CrossRef]

8. Fichera, C.R.; Modica, G.; Pollino, M. Land Cover classification and change-detection analysis using multi-temporal remote sensed imagery and landscape metrics. *Eur. J. Remote Sens.* **2012**, *45*, 1–18. [CrossRef]

9. Gómez, C.; White, J.C.; Wulder, M.A. Optical remotely sensed time series data for land cover classification: A review. *ISPRS J. Photogramm. Remote Sens.* **2016**, *116*, 55–72. [CrossRef]

10. De Luca, G.; Silva, J.M.N.; Di Fazio, S.; Modica, G. Integrated use of Sentinel-1 and Sentinel-2 data and open-source machine learning algorithms for land cover mapping in a Mediterranean region. *Eur. J. Remote Sens.* **2022**, *55*, 52–70. [CrossRef]

11. Sexton, J.O.; Song, X.-P.; Huang, C.; Channan, S.; Baker, M.E.; Townshend, J.R. Urban growth of the Washington, D.C.–Baltimore, MD metropolitan region from 1984 to 2010 by annual, Landsat-based estimates of impervious cover. *Remote Sens. Environ.* **2013**, *129*, 42–53. [CrossRef]

12. Hu, H.; Ban, Y. Unsupervised Change Detection in Multitemporal SAR Images Over Large Urban Areas. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3248–3261. [CrossRef]

13. Appiah, D.; Schröder, D.; Forkuo, E.; Bugri, J. Application of Geo-Information Techniques in Land Use and Land Cover Change Analysis in a Peri-Urban District of Ghana. *ISPRS Int. J. Geo-Inf.* **2015**, *4*, 1265–1289. [CrossRef]

14. Christensen, P.; McCord, G.C. Geographic determinants of China's urbanization. *Reg. Sci. Urban Econ.* **2016**, *59*, 90–102. [CrossRef]

15. Mertes, C.M.; Schneider, A.; Sulla-Menashe, D.; Tatem, A.J.; Tan, B. Detecting change in urban areas at continental scales with MODIS data. *Remote Sens. Environ.* **2015**, *158*, 331–347. [CrossRef]

16. Zhou, Y.; Smith, S.J.; Zhao, K.; Imhoff, M.; Thomson, A.; Bond-Lamberty, B.; Asrar, G.R.; Zhang, X.; He, C.; Elvidge, C.D. A global map of urban extent from nightlights. *Environ. Res Lett.* **2015**, *10*, 054011. [CrossRef]

17. Mao, W.; Lu, D.; Hou, L.; Liu, X.; Yue, W. Comparison of Machine-Learning Methods for Urban Land-Use Mapping in Hangzhou City, China. *Remote Sens.* **2020**, *12*, 2817. [CrossRef]

18. Huang, B.; Zhao, B.; Song, Y. Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. *Remote Sens. Environ.* **2018**, *214*, 73–86. [CrossRef]

19. Gbanie, S.; Griffin, A.; Thornton, A. Impacts on the Urban Environment: Land Cover Change Trajectories and Landscape Fragmentation in Post-War Western Area, Sierra Leone. *Remote Sens.* **2018**, *10*, 129. [CrossRef]

20. Reynolds, R.; Liang, L.; Li, X.; Dennis, J. Monitoring Annual Urban Changes in a Rapidly Growing Portion of Northwest Arkansas with a 20-Year Landsat Record. *Remote Sens.* **2017**, *9*, 71. [CrossRef]

21. Samal, D.R.; Gedam, S.S. Monitoring land use changes associated with urbanization: An object based image analysis approach. *Eur. J. Remote Sens.* **2015**, *48*, 85–99. [CrossRef]

22. Pandey, B.; Zhang, Q.; Seto, K.C. Time series analysis of satellite data to characterize multiple land use transitions: A case study of urban growth and agricultural land loss in India. *J. Land Use Sci.* **2018**, *13*, 221–237. [CrossRef]

23. Mahmoud, H.; Divigalpitiya, P. Spatiotemporal variation analysis of urban land expansion in the establishment of new communities in Upper Egypt: A case study of New Asyut city. *Egypt J. Remote Sens. Sp. Sci.* **2019**, *22*, 59–66. [CrossRef]

24. Chai, B.; Seto, K.C. Conceptualizing and characterizing micro-urbanization: A new perspective applied to Africa. *Landsc. Urban Plan.* **2019**, *190*, 103595. [CrossRef]

25. Ma, Y.; Wu, H.; Wang, L.; Huang, B.; Ranjan, R.; Zomaya, A.; Jie, W. Remote sensing big data computing: Challenges and opportunities. *Future Gener. Comput. Syst.* **2015**, *51*, 47–60. [CrossRef]

26. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]

27. Schmitt, M.; Hughes, L.H.; Qiu, C.; Zhu, X.X. Aggregating Cloud-Free Sentinel-2 Images with Google Earth Engine. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *IV-2/W7*, 145–152. [CrossRef]

28. Belcore, E.; Piras, M.; Wozniak, E. Specific Alpine Environment Land Cover Classification Methodology: Google Earth Engine Processing for Sentinel-2 Data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *XLIII-B3-2*, 663–670. [CrossRef]

29. Kumar, L.; Mutanga, O. Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. *Remote Sens.* **2018**, *10*, 1509. [CrossRef]

30. Tamiminia, H.; Salehi, B.; Mahdianpari, M.; Quackenbush, L.; Adeli, S.; Brisco, B. Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. *ISPRS J. Photogramm. Remote Sens.* **2020**, *164*, 152–170. [CrossRef]

31. Yang, L.; Driscol, J.; Sarigai, S.; Wu, Q.; Chen, H.; Lippitt, C.D. Google Earth Engine and Artificial Intelligence (AI): A Comprehensive Review. *Remote Sens.* **2022**, *14*, 3253. [CrossRef]

32. Fattore, C.; Abate, N.; Faridani, F.; Masini, N.; Lasaponara, R. Google Earth Engine as Multi-Sensor Open-Source Tool for Supporting the Preservation of Archaeological Areas: The Case Study of Flood and Fire Mapping in Metaponto, Italy. *Sensors* **2021**, *21*, 1791. [CrossRef] [PubMed]

33. Lasaponara, R.; Abate, N.; Masini, N. On the Use of Google Earth Engine and Sentinel Data to Detect "Lost" Sections of Ancient Roads. The Case of Via Appia. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 3001605. [CrossRef]

34. Danese, M.; Gioia, D.; Biscione, M. Integrated Methods for Cultural Heritage Risk Assessment: Google Earth Engine, Spatial Analysis, Machine Learning. In Proceedings of the Computational Science and Its Applications–ICCSA 2021: 21st International Conference, Cagliari, Italy, 13–16 September 2021; pp. 605–619.

35. Clemente, J.P.; Fontanelli, G.; Ovando, G.G.; Roa, Y.L.B.; Lapini, A.; Santi, E. Google Earth Engine: Application of Algorithms for Remote Sensing of Crops in Tuscany (Italy). *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *XLII-3/W12*, 291–296. [CrossRef]

36. Balestra, M.; Chiappini, S.; Malinverni, E.S.; Galli, A.; Marcheggiani, E. A Machine Learning Approach for Mapping Forest Categories: An Application of Google Earth Engine for the Case Study of Monte Sant'Angelo, Central Italy. In Proceedings of the Computational Science and Its Applications–ICCSA 2021: 21st International Conference, Cagliari, Italy, 13–16 September 2021; pp. 155–168.

37. Pérez-Cutillas, P.; Pérez-Navarro, A.; Conesa-García, C.; Zema, D.A.; Amado-Álvarez, J.P. What is going on within google earth engine? A systematic review and meta-analysis. *Remote Sens. Appl. Soc. Environ.* **2023**, *29*, 100907. [CrossRef]

38. Goldblatt, R.; You, W.; Hanson, G.; Khandelwal, A. Detecting the Boundaries of Urban Areas in India: A Dataset for Pixel-Based Image Classification in Google Earth Engine. *Remote Sens.* **2016**, *8*, 634. [CrossRef]

39. Sengupta, D.; Chen, R.; Meadows, M.E.; Choi, Y.R.; Banerjee, A.; Zilong, X. Mapping Trajectories of Coastal Land Reclamation in Nine Deltaic Megacities using Google Earth Engine. *Remote Sens.* **2019**, *11*, 2621. [CrossRef]

40. Zhang, Z.; Wei, M.; Pu, D.; He, G.; Wang, G.; Long, T. Assessment of Annual Composite Images Obtained by Google Earth Engine for Urban Areas Mapping Using Random Forest. *Remote Sens.* **2021**, *13*, 748. [CrossRef]

41. Carneiro, E.; Lopes, W.; Espindola, G. Urban Land Mapping Based on Remote Sensing Time Series in the Google Earth Engine Platform: A Case Study of the Teresina-Timon Conurbation Area in Brazil. *Remote Sens.* **2021**, *13*, 1338. [CrossRef]

42. Phan, T.N.; Kuch, V.; Lehnert, L.W. Land Cover Classification using Google Earth Engine and Random Forest Classifier—The Role of Image Composition. *Remote Sens.* **2020**, *12*, 2411. [CrossRef]

43. Solano, F.; Praticò, S.; Piovesan, G.; Chiarucci, A.; Argentieri, A.; Modica, G. Characterising historical transformation trajectories of the forest landscape in Rome's Metropolitan area (Italy) for effective planning of sustainability goals. *Land Degrad. Dev.* **2021**, *32*, 4708–4726. [CrossRef]

44. Bozzano, F.; Esposito, C.; Mazzanti, P.; Patti, M.; Scancella, S. Imaging Multi-Age Construction Settlement Behaviour by Advanced SAR Interferometry. *Remote Sens.* **2018**, *10*, 1137. [CrossRef]

45. Della Ventura, G.; Patané, A. Le miniere dei Monti della Tolfa-Allumiere (Roma). In *I siti della memoria geologica nel territorio del Lazio. Memorie Descrittive della Carta Geologica d'Italia*; Pantaloni, M., Console, F., Argentieri, A., Mantero, D., Eds.; Servizio Geologico d'Italia—ISPRA: Rome, Italy, 2020; Volume 106, 328p, pp. 23–32. ISBN 978-88-9311-0815.

46. Pietro, R.; Azzella, M.; Facioni, L. The Forest Vegetation of the Tolfa-Ceriti Mountains (Northern Latium—Central Italy). *Hacquetia* **2010**, *9*, 91–150. [CrossRef]

47. Rouse, J.W.; Hass, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the great plains with ERTS. *Third Earth Resour. Technol. Satell. Symp.* **1973**, *1*, 309–317.

48. Capolupo, A.; Monterisi, C.; Tarantino, E. Landsat Images Classification Algorithm (LICA) to automatically extract land cover information in Google Earth Engine environment. *Remote Sens.* **2020**, *12*, 1201. [CrossRef]

49. Jones, H.G.; Vaughan, R.A. *Remote Sensing of Vegetation: Principles, Techniques, and Applications*, 1st ed.; Oxford University Press: Oxford, UK, 2010; p. 353.

50. Javed, A.; Cheng, Q.; Peng, H.; Altan, O.; Li, Y.; Ara, I.; Huq, E.; Ali, Y.; Saleem, N. Review of Spectral Indices for Urban Remote Sensing. *Photogramm. Eng. Remote Sens.* **2021**, *87*, 513–524. [CrossRef]

51. Tassi, A.; Gigante, D.; Modica, G.; Di Martino, L.; Vizzari, M. Pixel- vs. Object-Based Landsat 8 Data Classification in Google Earth Engine Using Random Forest: The Case Study of Maiella National Park. *Remote Sens.* **2021**, *13*, 2299. [CrossRef]

52. Nguyen, Q.H.; Ly, H.-B.; Ho, L.S.; Al-Ansari, N.; Le, H.V.; Tran, V.Q.; Prakash, I.; Pham, B.T. Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil. *Math. Probl. Eng.* **2021**, *2021*, 1–15. [CrossRef]

53. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

54. Praticò, S.; Solano, F.; Di Fazio, S.; Modica, G. Machine Learning Classification of Mediterranean Forest Habitats in Google Earth Engine Based on Seasonal Sentinel-2 Time-Series and Input Image Composition Optimisation. *Remote Sens.* **2021**, *13*, 586. [CrossRef]

55. Li, X.; Chen, W.; Cheng, X.; Wang, L. A Comparison of Machine Learning Algorithms for Mapping of Complex Surface-Mined and Agricultural Landscapes Using ZiYuan-3 Stereo Satellite Imagery. *Remote Sens.* **2016**, *8*, 514. [CrossRef]

56. Cutler, D.R.; Edwards, T.C.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random Forests for Classification in Ecology. *Ecology* **2007**, *88*, 2783–2792. [CrossRef] [PubMed]

57. Li, Z.; Xin, X.; Tang, H.; Yang, F.; Chen, B.; Zhang, B. Estimating grassland LAI using the Random Forests approach and Landsat imagery in the meadow steppe of Hulunber, China. *J. Integr. Agric.* **2017**, *16*, 286–297. [CrossRef]

58. Han, H.; Guo, X.; Yu, H. Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest. In Proceedings of the 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 26–28 August 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 219–224.

59. Mellor, A.; Haywood, A.; Stone, C.; Jones, S. The Performance of Random Forests in an Operational Setting for Large Area Sclerophyll Forest Classification. *Remote Sens.* **2013**, *5*, 2838–2856. [CrossRef]

60. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data [Internet]*; CRC Press: Boca Raton, FL, USA, 2019.

61. Story, M.; Congalton, R.G. Remote Sensing Brief Accuracy Assessment: A User's Perspective. *Photogramm. Eng. Remote Sens.* **1986**, *52*, 397–399.

62. Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46. [CrossRef]

63. Momeni, R.; Aplin, P.; Boyd, D. Mapping Complex Urban Land Cover from Spaceborne Imagery: The Influence of Spatial Resolution, Spectral Band Set and Classification Approach. *Remote Sens.* **2016**, *8*, 88. [CrossRef]

64. McNemar, Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* **1947**, *12*, 153–157. [CrossRef]

65. Modica, G.; De Luca, G.; Messina, G.; Praticò, S. Comparison and assessment of different object-based classifications using machine learning algorithms and UAVs multispectral imagery: A case study in a citrus orchard and an onion crop. *Eur. J. Remote Sens.* **2021**, *54*, 431–460. [CrossRef]

66. Matarira, D.; Mutanga, O.; Naidu, M. Google Earth Engine for Informal Settlement Mapping: A Random Forest Classification Using Spectral and Textural Information. *Remote Sens.* **2022**, *14*, 5130. [CrossRef]

67. Cipriani, L.E.; Ferri, S.; Lami, G.; Pranzini, E. Human Impact on Shoreline Evolution Along the Follonica Gulf (Southern Tuscany): How Tourism May Kill the Goose that Lays the Golden Eggs. *J. Coast. Res.* **2011**, *61*, 290–294. [CrossRef]

68. Fuerst-Bjeliš, B.; Durbešić, A. Littoralization and behind: Environmental Change in Mediterranean Croatia. In *The Overarching Issues of the European Space/Strategies for Spatial (Re)Planning Based on Innovation, Sustainability and Change*; Pina, H., Martins, F., Ferreira, C., Eds.; Fundação Universidade Do Porto—Faculdade de Letras da Universidade do Porto: Porto, Portugal, 2013; pp. 136–147. [CrossRef]

69. Manganelli, B.; Murgante, B. The Dynamics of Urban Land Rent in Italian Regional Capital Cities. *Land* **2017**, *6*, 54. [CrossRef]

70. Salvati, L.; Sabbi, A. Exploring long-term land cover changes in an urban region of southern Europe. *Int. J. Sustain. Dev. World Ecol.* **2011**, *18*, 273–282. [CrossRef]

71. Biasi, R.; Colantoni, A.; Ferrara, C.; Ranalli, F.; Salvati, L. In-between sprawl and fires: Long-term forest expansion and settlement dynamics at the wildland–urban interface in Rome, Italy. *Int. J. Sustain. Dev. World Ecol.* **2015**, *22*, 467–475. [CrossRef]

72. De Felice, P.; Grillotti Di Giacomo, M.G. *Land Grabbing e Land Concentration. I Predatori Della Terra Tra Neocolonialismo e Crisi Migratorie*; Franco Angeli: Milan, Italy, 2018.

73. Lelo, K. Agro Romano: Un territorio in trasformazione. In *Roma Moderna e Contemporanea*; Università Toma tre—CROMA: Roma, Italy, 2016; pp. 9–48. [CrossRef]

74. De Felice, P.; Lodato, F. La tenuta di Zambra nell'agro romano oltre la Convenzione Europea del Paesaggio. Un'analisi geografica a scala locale. In *Oltre la Convenzione Pensare, Studiare, Costruire il Paesaggio Vent'anni Dopo*; Società di Studi Geografici: Firenze, Italy, 2020; pp. 702–721.

75. Salvati, L. "A Chronicle of a Death Foretold": Urban Expansion and Land Consumption in Rome, Italy. *Eur. Plan. Stud.* **2013**, *21*, 1176–1188. [CrossRef]

76. Praticò, S.; Solano, F.; Di Fazio, S.; Modica, G. A Multitemporal Fragmentation-Based Approach for a Dynamics Analysis of Agricultural Terraced Systems: The Case Study of Costa Viola Landscape (Southern Italy). *Land* **2022**, *11*, 482. [CrossRef]

77. Modica, G.; Merlino, A.; Solano, F.; Mercurio, R. An index for the assessment of degraded Mediterranean forest ecosystems. *For. Syst.* **2015**, *24*, e037. [CrossRef]

78. Caneva, G.; Benelli, F.; Bartoli, F.; Cicinelli, E. Safeguarding natural and cultural heritage on Etruscan tombs (La Banditaccia, Cerveteri, Italy). *Rend. Lincei Sci. Fis. Nat.* **2018**, *29*, 891–907. [CrossRef]

79. Pollino, M.; Modica, G. Free Web Mapping Tools to Characterise Landscape Dynamics and to Favour e-Participation. In *Computational Science and Its Applications—ICCSA 2013. Lecture Notes In Computer Science*; Murgante, B., Misra, S., Carlini, M., Torre, C.M., Nguyen, H.-Q., Taniar, D., Apduhan, B.O., Gervasi, O., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; Volume 7973, pp. 566–581. [CrossRef]