University of Tuscia - Viterbo

Department for Innovation in Biological, Agrofood and Forest systems (DIBAF)

PhD course in

Science, Technology, and Biotechnology for Sustainability
XXXV cycle

LAND COVER, FORESTRY AND FOREST ECOSYSTEM SERVICES: METHODS FOR THE PRODUCTION OF NEW NATIONAL MAPS

AGR/05

PhD Student:
Dott.ssa Alice Cavalli

Alice Cavalli

PhD Coordinator
Prof. Andrea Vannini

Tutor
Prof. Giuseppe Scarascia-Mugnozza

Co-tutore
Ing. Michele Munafò

A.Y. 2021/22

Short Abstract

Land cover, land use and their changes are important to understand the evolution of the territory which is crucial for sustainable resources management, environmental protection and studying climate change impact. In particular, forest and afforestation - the change from non forest to forest cover - monitoring is important because of the positive effects linked to forest, as the mitigation of climate change effects, soil protection, carbon sequestration, but also because forest change can be a symptom on the ongoing climate change, as in the case of movement to higher elevations or latitudes of the tree line. Remote sensed data can be a valuable tool to monitoring land cover, land use and their changes, allowing to study these phenomena at different spatial and temporal resolution.

Given the importance of land cover and land use monitoring, the aims of this thesis are: (i) to develop a land cover and land cover change map of Italy using remotely sensed data, (ii) to develop and verify a new methodology to map afforestation occurred in the last 35 years using remotely sensed data, (iii) to construct a map of afforestation occurred between 1985 and 2019 in Italy, (iv) to assess the afforestation areas using statistically rigorous procedures, (v) to estimate the potential carbon sequestration in afforestation areas, (vi) to define the type of tree vegetation most affected by afforestation between 1985 and 2019.

The land cover map of Italy classifies the national surface in eight land cover classes and four land cover change classes, with an Overall Accuracy of 83%, and of 86% for woody vegetation class. The effectiveness of remotely sensed data in afforestation mapping was assessed with an Overall Accuracy of 87% and the national afforestation map was constructed using two classes, afforestation and non afforestation, with an Overall Accuracy of 87%. The afforestation estimates illustrate that afforestation between 1985 and 2019 occurred approximately on 2.8 ± 0.2 million hectares, with different intensity along administrative regions and elevation classes.

As regard the potential carbon sequestration, results illustrate that 202 ± 13 million t of potential carbon will be sequestrated when the trees in the afforestation areas mature. Finally, the integration between the national land cover map and the afforestation estimation dataset allowed to assess that broadleaves have been most affected by afforestation between 1985 and 2019, with 2.5±0.2 million hectares, the 89% of the total afforestation.

La copertura, l'uso del suolo e i loro cambiamenti sono importanti per comprendere l'evoluzione del territorio, fondamentale per la gestione sostenibile delle risorse, la protezione dell'ambiente e lo studio dell'impatto dei cambiamenti climatici. In particolare, il monitoraggio delle foreste e dell'espansione forestale - il passaggio da una copertura non forestale a una copertura forestale - è importante per gli effetti positivi legati alle foreste, come la mitigazione degli effetti del cambiamento climatico, la protezione del suolo, il sequestro di carbonio, ma anche perché il cambiamento delle foreste può essere un indicatore del cambiamento climatico in atto, come nel caso dello spostamento verso quote o latitudini più elevate della linea degli alberi. I dati telerilevati possono essere uno strumento prezioso per monitorare la copertura, l'uso del suolo e i cambiamenti, consentendo di studiare questi fenomeni a diverse risoluzioni spaziali e temporali.

Data l'importanza del monitoraggio della copertura e dell'uso del suolo, gli obiettivi di questa tesi sono i seguenti: (i) sviluppare una mappa della copertura e del cambiamento di copertura del suolo in Italia utilizzando dati telerilevati, (ii) sviluppare e verificare un nuovo metodo per mappare l'espansione forestale avvenuta negli ultimi 35 anni utilizzando dati telerilevati, (iii) costruire una mappa dell' espansione forestale avvenuta tra il 1985 e il 2019 in Italia, (iv) valutare le aree di espansione forestale utilizzando procedure statisticamente rigorose, (v) stimare il sequestro di carbonio potenziale nelle aree di espansione forestale, (vi) definire il tipo di vegetazione arborea maggiormente interessata dall' espansione forestale tra il 1985 e il 2019.

La carta di copertura del suolo dell'Italia classifica la superficie nazionale in otto classi di copertura del suolo e quattro classi di cambiamento di copertura del suolo, con un'accuratezza complessiva dell'83%, che raggiunge l'86% per la classe di vegetazione legnosa. L'efficacia dei dati telerilevati nella mappatura dell'espansione forestale è stata valutata con un'accuratezza complessiva dell'87% e la carta nazionale dell'espansione forestale è stata costruita utilizzando due classi, espansione forestale e non espansione forestale, con un'accuratezza complessiva dell'87%. Le stime di espansione forestale indicano che essa tra il 1985 e il 2019 si è verificata approssimativamente su 2,8 ± 0,2 milioni di ettari, con un'intensità diversa a seconda delle regioni amministrative e delle classi altimetriche.

Per quanto riguarda il sequestro di carbonio potenziale, i risultati indicano che 202 ± 13 milioni di tonnellate di carbonio saranno sequestrate quando gli alberi nelle aree di espansione forestale avranno raggiunto la maturità. Infine, l'integrazione tra la carta nazionale della copertura del suolo e il dataset di stima dell'espansione forestale ha permesso di valutare che le latifoglie sono state maggiormente interessate dall'espansione forestale tra il 1985 e il 2019, con 2,5±0,2 milioni di ettari, l'89% dell'espansione forestale totale.

Keywords (5-7)

Land cover monitoring, decision rules, random forests, afforestation, Landsat, Sentinel

Extended abstract

Land cover is defined as the biophysical cover of Earth's surface, as artificial surfaces, agricultural areas, forests and semi-natural areas, water and wetlands by the Directive 2007/02 of the European Commission. Different land cover types provide various ecosystem services and support habitats; land cover is also considered a climate variable, an element for the description of the environment. Human activities depend on land cover and its evolution, and the use and functions of a territory and its description of land in terms of its socio-economic and ecological purpose define the land use (2007/02/EC). Therefore, the comprehension of the processes that lead to the transformation of the territory are crucial for sustainable resource management, environmental protection, modelling of climate change scenarios, and for the assessment of the impacts on the earth system directly linked to human activities. Remote sensing is important to the study of Earth phenomena, which can be monitored and assessed at different spatial and temporal resolutions (Agrillo et al., 2021). There are many algorithms for investigating land cover and land use starting from the classification of satellite images. Threshold methods are used for their simplicity and because they require few parameters. The identification and use of thresholds related to the spectral characteristics and their variation over time are useful to assess both physical and biological characteristics of land cover, to classify cover change, to define crop phenology for determining start and end of growing season, and to classify tree species or burned areas.

Land cover changes among forest and other land use classes include the gain and the loss of tree cover. Afforestation constitutes the gain of arboreal cover, and it consists in the conversion from other land uses to forest, while deforestation is the conversion of forest to other land uses, representing the loss of tree cover (FAO, 2020). Afforestation can be natural, via spontaneous colonization of soil by vegetation, or anthropogenic through trees planting or direct sowing.

Afforestation processes may increase the terrestrial carbon sink, create new ecological corridors, protect soil (Spadoni et al., 2020) and water, contribute to combating global warming, and improve biodiversity (IPCC 2015). Monitoring afforestation processes is therefore important for evaluating these positive effects. On the other hand, proper monitoring is also needed to assess the negative impacts that mismanaged afforestation processes may impose on the territory such as loss of grazing habitats and reduced landscape values (Buscardo et al., 2008; IPCC 2015; Veldman et al., 2015). Afforestation may also be a symptom of the effect of climate change, as in the case of

migration to higher elevations or latitudes of the geographic range of tree species and of the tree line, which is the elevation or latitude above which trees give way to shrubs (Chersich et al., 2015). Finally, because of its important contribution to land use dynamics, afforestation processes have a central role in the implementation of policies that aim to preserve forest resources and provide ecosystem services. Afforestation processes, as land cover, can be analyzed using remotely sensed data. The afforestation detection model can be constructed by analysing observations of change in both the response variable and the remotely sensed predictor variables between two dates (McRoberts et al., 2015). This method is based on the analysis of trends in variables or indices linked to afforestation (such as vegetation indices as Normalized Difference Vegetation Index - NDVI) during the observation period; these variables or indices are then used to identify and classify the afforestation spectral response (Bollandsås et al., 2013).

The aims of this thesis are: (i) to develop a land cover and land cover change map of Italy using remotely sensed data, (ii) to develop and verify a new methodology to map afforestation occurred in the last 35 years using remotely sensed data, (iii) to construct a map of afforestation occurred between 1985 and 2019 in Italy, (iv) to assess the afforestation areas using statistically rigorous procedures, (v) to estimate the potential carbon sequestration in afforestation areas, (vi) to define the type of tree vegetation most affected by afforestation between 1985 and 2019.

In the first part of the research the **Land Cover Map** has been constructed; it represents the land cover in Italy in 2018, considering also some of important land cover change, occurred between 2017 and 2018. The national area was classified using Sentinel-2 and Sentinel-1 images and the rule based classification algorithm trained on the Region Of Interest, and then applied to classify the national area. This method is based on pixel backscatter and spectral response. The national land cover map is composed by 8 land cover classes and four land cover change classes based on the EAGLE concept and has a spatial resolution of 10 m. The woody vegetation class and surface is equal to 13,614,064 ha (45% of Italy); the broad-leaved occupy the larger area (12,006,482 ha) and confers occupy 1,607,582 ha. The 43% of Italian land surface is covered by herbaceous vegetation, mostly a temporary cover of annual crops (12,657,890 ha) while the permanent grassland area is 239,834 ha. The abiotic surface[1] covers 10% of the national area, with the artificial surfaces which occupy 1,127,999 ha and the natural surfaces which occupy 953,574 ha. Water and ice surface amounts at 449,510 ha. As regards the land cover change occurred between 2017 and 2018, this surface is 97,102 ha, and is composed by land consumption, ecosystem restoration, burned areas

[1] Any unvegetated surfaces, either covered with man-made artificial structures or geologically natural material surfaces

and other disturbances. The overall accuracy of the map is 83%; regarding the single class accuracies water bodies, abiotic artificial surfaces and permanent herbaceous have the largest values (>90%), the land cover changes have the lowest accuracy (between 30% and 83%), while broadleaves and conifers accuracies are respectively 87% and 90%.

In the second part of the study the **effectiveness of remote sensed data in afforestation areas monitoring** and the analysis of the spectral difference between forest, non-forest and afforestation areas was assessed. This analysis followed the construction of national land cover map and was the preliminary analysis for the construction of the national afforestation map, illustrated in the last part of the research. To define the afforestation training area Land Use Inventory of Italy (IUTI, 1990-2016), Orthophotos WMS (1988-2012), VHR images (2012-2020) and Landsat images (1984-2020) were used; first, the training dataset based on IUTI elements (afforestation) was defined though photointerpretation. Then the analysis of the spectral differences between forest (F), non forest (NF) and afforestation (A) areas, between 1984 and 2020, was made using spectral indexes (NDVI, NBR, EVI), bands (red, green, blue, nir, swir1, swir2) and tasseled cap (brightness, wetness, greenness, angle), calculated on Landsat images, which have been properly preprocessed. All the input data previously described have been used to calculate the mean, standard deviation, R (Pearson correlation) and slope (linear regression) of every pixel value, to analyze their trend in each temporal range (NF - A - F) and verify if there is a difference between them. The random forests (Breiman, 2011) algorithm, previously calibrated with random search and tested with 5-fold cross validation, was finally iterated on these data, to verify if it is possible to recognize the different phases automatically (Bergstra and Bengio, 2012; Laurin et al., 2021). The result is composed by 100 combinations of max features and max depth and for every couple of data the overall accuracy and the standard deviation is calculated. For each result (comparison of NF - A; F - A; NF - F - A; NF - F) the best hyperparameters[2] iteration (best Overall Accuracy) is selected. The 5-fold cv result with the highest OA derives from the comparison of F and A (86,8%); this means that forest areas and afforestation areas have the highest difference in spectral response, and it confirms that remote sensing can be a valid support in the assessment of afforestation.

Finally, using the positive outcome from the analysis of the spectral difference between forest, non forest and afforestation, the **Italian afforestation map** was constructed, considering 35 years (1985-2019) and estimating the afforestation area using a stratified estimator. Finally, the potential carbon (t) that will be sequestered in areas afforested between 1985 and 2019 once the forest is

[2] Random forests hyperparameters are parameters that can be calibrated to obtain the minimum error rate (Laurin et al., 2021). In this study, we tuned the following two: max features (the number of features considered) and max depth (the maximum depth of the tree, the longest path between the root node and the leaf node).

mature was assessed. In this research we analyzed the afforestation which brought to the present forest cover. The training dataset from the preliminary study was integrated using the same procedure, resulting in a final training dataset of 1,578 polygons, divided into forest, non-forest and afforestation. The training dataset was used together with the Landsat Best Available Pixel time series (1985-2019) to calculate a set of temporal predictors that, together with random forests (following the procedure used in the preliminary study), facilitated construction of a map of afforested areas in Italy. The map was then used to guide selection of a stratified estimation sample dataset of more than 4,000 points that, after a complex photointerpretation phase, was used to estimate afforestation areas and associated confidence intervals at national, administrative Regions and elevation classes levels. The results are composed by the national afforestation map and the national, administrative Regions, elevation classes afforestation estimates. To estimate the potential carbon that will be sequestered in areas afforested between 1985 and 2019 once the forest is mature, we used the estimates of total biomass per unit of forest area over all pools (aboveground biomass, belowground biomass, litter, necromass and organic and mineral soil) from the National Inventory of Forests and Carbon Sinks (INFC, 2005) for individual administrative Regions (RAF, 2018). For each Region, we multiplied the estimated area of afforestation by the INFC estimate of biomass per unit area to obtain an estimate total biomass once the forest is mature. To estimate the potential C-sequestration (t), we multiplied the estimate of total biomass by the default carbon fraction in biomass of 0.5 (Vitullo et al., 2007). The 95% confidence interval estimates were also calculated for the carbon estimates at national and administrative Region levels.

The afforestation map predicted two classes of land cover: afforestation and non-afforestation between 1985 and 2019. The estimated map OA obtained was 87%.

Afforestation did not occur uniformly at the national level; in fact, when considering administrative Region boundaries, estimates of afforestation areas for four administrative Regions (Abruzzo, Calabria, Latium and Sardinia) were greater than 200,000 ha, with the largest estimate of 260,653 ± 58,522 ha (10.80% of the administrative Region area) for Sardinia. Estimates were less than 90,000 ha for Apulia, Friuli Venezia Giulia, Molise, Valle d'Aosta and Veneto, with the least estimate of 28,644 ± 12,114 ha (8.78% of the administrative Regional area) for Valle d'Aosta. In the remaining administrative Regions the estimated afforestation areas were between 100,000 and 200,000 ha. Abruzzo and Liguria had the greatest percentage afforestation estimates relative to the administrative Region area.

The intermediate elevation classes are the areas where the largest afforestation area between 1985 and 2019 was estimated. Below 1,200 m asl, estimates of afforested surface were greater than 200,000 ha while the largest estimate of 549,497 ± 84,979 ha (13.94%) was observed for the 400-

600 m asl elevation class. For higher altitudes, estimates were less than 100,000 ha, a finding that aligns with the decrease in forest habitats. The greatest percentage afforestation area estimate was ~20% for the 600-1200 m asl class.
At the national level, the afforestation area between 1985 and 2019 covered approximately 2.8 ± 0.2 million ha.
The Region with the largest estimate of potentially sequestered carbon (the carbon that will be sequestered in areas afforested between 1985 and 2019 once the forest is mature) was Calabria with 17,479,338 ± 4,289,144 t, while the Region with the least estimate of potential carbon sequestrated was Valle d'Aosta with 1,468,005 ± 620,842 t. In 11 of the 20 administrative Regions, estimates were greater than 10 million t of potential carbon sequestrated, and nationwide, 202,420,138 ± 13,908,807 t of potential carbon would be sequestrated when the trees in the afforestation areas mature.

To improve the analysis of afforestation one of the future developments illustrated in the last part of the research was tested. Land cover and afforestation data obtained in the two phases of the research were intersected to analyze the typology of vegetation subject to afforestation between 1985 and 2019.
The national land cover map has been reclassified considering the two tree cover classes, conifers and broadleaves, obtaining a tree cover map. Following the procedure of afforestation map construction, we processed the tree cover map and then we extracted the tree cover map values using the estimation dataset of 4,000 samples, obtaining an estimation dataset with the information on the vegetation type (conifers and broadleaves). Using the points classified as afforestation/non-afforestation weighted to afforestation/non afforestation values that we calculated with the stratified estimator, we extracted the afforestation/non-afforestation area classified into conifers and broadleaves. Results showed that broadleaves were the vegetation type most affected by afforestation between 1985 and 2019, with 2,531,856,761±180,696 ha, 289,541±20,668 ha of afforestation can be associated to conifers expansion.

Conclusions and Future Perspectives

The land cover classification methodology allowed to provide nationwide land cover map with high spatial resolution and accuracy.
The map fits well into the Mirror Copernicus initiative of the Space Economy Strategic Plan of the Italian Ministry of Economic Development, which provides for the creation of frequently updated products at a national scale, for the monitoring of the territory.

The approach based on decision rules guarantees the control of the choices adopted for the classification, and a high level of modularity. An important development of the methodology will involve the adaptation of the thresholds according to the characteristics of the territory, considering in the first place the influence of hydrometeorological variables, seasonal climatic variations on the indices and elevation models. The versatility of the algorithm and of the classification system make the methodology suitable for future integrations and for adaptation for specific purposes, while remaining consistent with the requirements proposed at the National and European level.
Three main conclusions were drawn from the afforestation study. First, thanks to remotely sensed data with fine spatial resolution and available for long time series, it is possible to predict locations where afforestation occurred by analyzing their evolution over time. Second, predicted afforestation maps can be integrated with a photointerpretation phase and auxiliary information such as administrative Regional boundaries and the DEM to facilitate analyses of territorial differences of the afforestation phenomenon, offering estimates that can be important from a management point of view. Moreover, integration of the map with sample data in combination with the statistically rigorous stratified estimators has shown that afforestation in Italy over the last 35 years has been about 2,8 million ha, associated with 202 million t of potential carbon sequestrated (the carbon that will be sequestered in areas afforested between 1985 and 2019 once the forest is mature). Finally, the integration between the national land cover map and the afforestation data allowed us to assess that afforestation, occurred between 1985 and 2019, involved broadleaves by 89% of the land surface. These results contribute to better understanding of the dynamics of tree cover which is fundamental for monitoring climate change.

Afforestation, in addition to being an extremely significant environmental phenomenon, is also a topical issue and is included in many international and national programs and legislations.
The analysis of afforestation areas can be strengthened by improving the classification to decrease omission and commission errors using additional auxiliary data related to the afforestation process such as aspect, slope, pedology, geographical area. Accurately predicting the vegetation associations, as we did in the last part of the research integrating the national land cover map and the afforestation estimates, while establishing the timing when the afforestation process could be considered as completed, could be useful also to define the time range of the afforestation process, while also considering the different growth rates for different species.
The method of afforestation classification can be the basis for studying afforestation dynamics. Future research can involve the classification method to assess an array of ecosystem services such as C-sequestration which can be estimated using both the classification method and chronosequences (Alberti et al., 2008). The classification method can be useful to predict where

afforestation is occurring in support of *ad hoc* forest management strategies. Using the classification method and the analysis of the afforestation spectral response, anthropogenic and natural afforestation can be distinguished. Finally, the classification method facilitates the assessment of the rate of afforestation within elevation classes or ecoregions, and to characterize afforestation trends.

References

Agrillo, E., Filipponi, F., Pezzarossa, A., Casella, L., Smiraglia, D., Orasi, A., Attorre, F., & Taramelli, A. (2021). Earth observation and biodiversity big data for forest habitat types classification and mapping. *Remote Sensing*, *13*(7). https://doi.org/10.3390/rs13071231

Alberti, G., Peressotti, A., Piussi, P., Zerbi, G., 2008. Forest ecosystem carbon accumulation during a secondary succession in the Eastern Prealps of Italy. Forestry, 81, 1-11.

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, *13*, 281–305.

Bollandsås, O. M., Gregoire, T. G., Næsset, E., Øyen, B. H., 2013. Detection of biomass change in a Norwegian mountain forest area using small footprint airborne laser scanner data. Statistical Methods and Applications, 22(1), 113–129.

Breiman, L. (2001), Random Forests. Machine Learning 45,5-32.

Buscardo, E., Smith, G. F., Kelly, D. L., Freitas, H., Iremonger, S., Mitchell, F. J. G., O'Donoghue, S., McKee, A. M., 2008. The early effects of afforestation on biodiversity of grasslands in Ireland. Biodiversity and Conservation, 17(5), 1057–1072**.**

Chersich, S., Rejšek, K., Vranová, V., Bordoni, M., Meisina, C., 2015. Climate change impacts on the Alpine ecosystem: An overview with focus on the soil - A review. Journal of Forest Science, 61(11), 496–514.

FAO, Food and Agriculture Organization, 2020. Global Forest Resources Assessment 2020: Main report. Rome.

IPCC - Intergovernal Panel on Climate Change, 2015. Agricolture, Forestry and Other Land Use (AFOLU) Chapter 11.

Laurin, G. V., Francini, S., Luti, T., Chirici, G., Pirotti, F., & Papale, D. (2021). Satellite open data to monitor forest damage caused by extreme climate-induced events: A case study of the Vaia storm in Northern Italy. *Forestry*, *94*(3), 407–416.

McRoberts, R. E., Næsset, E., Gobakken, T., Bollandsås, O. M., 2015. Indirect and direct estimation of forest biomass change using forest inventory and airborne laser scanning data. Remote Sensing of Environment, 164, 36–42.

Spadoni, G. L., Cavalli, A., Congedo, L., Munafò, M., (2020). Analysis of Normalized Difference Vegetation Index (NDVI) multi-temporal series for the production of forest cartography. Remote Sensing Applications: Society and Environment, 20.

Veldman, J. W., Overbeck, G. E., Negreiros, D., Mahy, G., le Stradic, S., Fernandes, G. W., Durigan, G., Buisson, E., Putz, F. E., Bond, W. J., 2015. Where Tree Planting and Forest Expansion are Bad for Biodiversity and Ecosystem Services. BioScience, 65(10), 1011–1018.

Vitullo M., Federici S., 2007. La contabilità del carbonio contenuto nelle foreste italiane [National carbon account in Italian forests]. Silvae, III (9).

Summary

Introduction

Land cover is the biophysical cover of Earth's surface, as artificial surfaces, agricultural areas, forests and semi-natural areas, water and wetlands (EC, 2007). Human activities depend on land cover and its evolution, and the use and functions of a territory and its description of land in terms of its socio-economic and ecological purpose define the land use (EC, 2007). The land cover monitoring is important to study the evolution of the territory, allowing to study the environmental impact and the consequences of the land cover and the land cover change.

Ecosystem services supply depends on land cover type, and land cover can be important to study the environment. The analysis of the land cover and land cover change are therefore crucial for sustainable resource management, environmental protection, modelling of climate change scenarios, and for the assessment of the impacts on the earth system directly linked to human activities.

The world's forest cover almost 4.06 billion of hectares of which 1.017.461 ha are in Europe. In Italy the 40% of the national area (11.8 million of hectares) are covered by forests. The Food and Agriculture Organization (FAO) agency define the forest as "*territory with arboreal coverage greater than 10% compared to an extension greater than 0.5 ha, where the trees reach a minimum height of 5 m when ripe and a minimum width of 20 meters*" (FAO, 2001). FAO define also the forest change processes: afforestation which constitute the gain of arboreal cover and the deforestation which is the loss of arboreal cover. The forest change phenomena do not occur homogeneously around the world, for example in Africa and South America deforestation trends up, while in Europe afforestation is generally a dominant process. In Italy forest cover increased constantly through the years, with an average growing annual rate of 0.3% (INFC, 2005).

Forest cover and afforestation can have many positive effects, being the main carbon sink, connecting different habitats, protecting soil and water, contributing to decrease the effect of the global warming. On the other hand, forest and afforestation can have some negative impacts as the loss of pastures (in the case of afforestation) and they can be a symptom of the effect of climate change, as in the case of movement to higher elevations or latitudes of the geographic range of tree species and of the tree line, the elevation or latitude above which trees give way to shrubs (Chersich et al., 2015).

Given the importance of land cover monitoring and assessment the aims of this thesis are: (i) to develop a land cover and land cover change map of Italy using remotely sensed data, (ii) to develop and verify a new methodology to map afforestation occurred in the last 35 years using remotely sensed data, (iii) to construct a map of afforestation occurred between 1985 and 2019 in Italy, (iv) to assess the afforestation areas using statistically rigorous procedures, (v) to estimate the potential carbon sequestration in afforestation areas, (vi) to define the type of tree vegetation most affected by afforestation between 1985 and 2019.

1. Land monitoring and ecosystem services

1.1. Environmental policies

1.1.1. International and European

The last 30 years have seen many United Nations conferences in which programs and actions have been implemented to protect the environment and mitigate the effects of climate change. In 1992, the Agenda 21 program focused on defining global plans and actions, and together with the Rio Declaration on Environment and Development, and the Statement of principles for the Sustainable Management of Forests was signed and adopted by 178 countries at the United Nations Conference on Environment and Development (UNCED) held in Rio de Janeiro. In the Brazilian city in 2012, the declaration "the Future we want" was adopted during the conference on sustainable development. The declaration sets several goals (Millennium Development Goals) relating to poverty eradication, food security, the sustainability of urban areas, cities and population health and defines sustainable development objectives to be adopted by 2014. In 2015, 193 UN countries signed up to the 2030 Agenda for Sustainable Development, which includes a commitment by signatory countries to pursue and monitor the 17 Sustainable Development Goals by 2030.

In 2016, the Paris Agreement made further commitments to curb the effects of climate change, including recognising the importance of reducing deforestation and forest habitat degradation to limit greenhouse gas emissions.

In November 2021, at the conclusion of COP 26 in Glasgow, further decisions emerged on the promotion of conversion from traditional to renewable energy sources, the commitment to halt deforestation by 2030 and to promote the conservation of ecosystems. With regard to forests, the conference published the Declaration on Forests and Land Use, which commits to conserving forests, better managing natural resources and promoting activities and financing that facilitate the transition to resilient and sustainable economic activities.

In Europe, numerous actions have been developed over time to achieve sustainable management and conservation of the environment. These projects contain guidelines for pursuing the various objectives, and also promote the approval of activities by Member States.

The Council Directive of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora, known as Habitats Directive and the Birds Directive (79/409/CEE) form the core of the EU's biodiversity conservation policy and are the legal basis on which Natura 2000 is founded.

The purpose of the Habitats Directive is "to conserve biodiversity by preserving natural habitats and wild fauna and flora in the European territory of the Member States to which the Treaty applies"

(Article 2). To achieve this objective, the Directive lays down measures to ensure the maintenance or restoration, at favorable conservation status, of the habitats and species of Community interest listed in its annexes.

The Directive is built around two pillars: the Natura 2000 ecological network, made up of sites aimed at the conservation of habitats and species listed in Annexes I and II respectively, and the protection regime for species listed in Annexes IV and V.

The Directive lays down rules for the management of Natura 2000 sites and the assessment of their impact (Art. 6), funding (Art. 8), monitoring and reporting on the implementation of the provisions of the Directive (Art. 11 and 17), and the granting of derogations (Art. 16). It also recognizes the importance of landscape elements that play an ecological connection role for wild flora and fauna (art. 10). The Directive was implemented in Italy in 1997 through the Regulation D.P.R. 8 September 1997 n. 357.

In 2006 the EU Forest Action Plan was published. The main objective of this plan was to support sustainable management and recognize the multifunctional role of forests. These objectives followed a series of principles, based on the implementation of national forest programs to implement international forest-related commitments, the importance of recognizing the cross-sectoral role of forests, the need to increase the competitiveness of the European forestry sector by maintaining sustainable resource management and respect for the principle of subsidiarity.

Alongside socio-economic objectives, such as encouraging technological research to increase competitiveness in the forestry sector, the exchange of marketing information on non-wood products or the exploitation of biomass for energy production, as well as the enhancement of the forest heritage to increase the welfare of the population, measures to protect and improve the environment were also decided upon. These include pursuing actions under the Kyoto Protocol to mitigate the effects of climate change, achieving EU biodiversity targets, implementing a European forest monitoring system and increasing the protection of forest areas.

In 2019, the European Green Deal was proposed by the European Commission, with the aim of protecting biodiversity, reducing pollution, promoting the circular economy and better waste management, and ensuring sustainable resource management. The first target is included in the European Union biodiversity strategy for 2030, published in 2020. It aims to conserve biodiversity, with the objective of embarking on a path of increasing biodiversity by 2030, to be implemented through an increase of protected areas on a European scale, an update of the soil protection strategy during 2021, the introduction of mandatory targets for the improvement of degraded ecosystems and the introduction of a European soil observatory for the period 2021-2027.

The forest strategy, published on 14 July 2021 and based on the biodiversity strategy, is one of the major initiatives of the European Green Deal. It will contribute to the achievement of several targets contained in the biodiversity strategy, such as a 55% reduction in greenhouse gas emissions by 2030 and climate neutrality by 2050. In addition, the EU Forest Strategy emphasises the multifunctionality of forests, and the fundamental role of forest management in achieving climate neutrality and preserving a good quality of life in rural areas. This strategy envisages a series of actions aimed at improving the quantity and quality of European forests, enhancing their degree of protection and resilience, to face the consequences of climate change, ensuring the maintenance of the socio-economic function of forests. In order to implement these objectives, sustainable forest exploitation activities will be promoted, both for durable products and for the production of energy from biomass; ecotourism and the development of new economic approaches based on forest sustainability will be promoted and funding will be given to forest owners for their management; the protection of the last primary forests in Europe will be guaranteed and the restoration of existing forests will be ensured in order to increase their resilience. Finally, reforestation and afforestation will be promoted, with the planting of 3 billion trees by 2030.
These actions will be implemented at the same time as defining new forest monitoring and data collection strategies and developing the forest research sector.

1.1.2. National

In Italy, in addition to the adoption of the international and European environmental goals, some important new strategies were introduced, with the aim to pursue the sustainable development and to mitigate the climate change effects. The Economic and Financial Document (DEF 2019-2020) provides for numerous investments aimed at implementing innovation projects with a view to environmental sustainability. In addition, with the National Integrated Energy and Climate Plan (PNIEC, 2021), each State sets out its contributions to the 2030 European objectives, including that of reducing greenhouse gases by 2050. With the Green New Deal and the National Recovery and Resilience Plan (PNRR, 2021), public and private investments are promoted that focus on protecting the environment by improving pollution levels, energy transition and preserving biodiversity. These programs include action to mitigate the effects of climate change, with urban regeneration measures, measures to combat hydrogeological instability, and the enactment of a law on land consumption. As regard forest laws, in Italy the national laws on forest were promulgated starting from the end of XIX century. The law n. 3917 was published in 1877 and defines the forest and subjects it to restrictions; it also authorizes reforestation operations and defines penalties and rights of one of the forest areas. The Law n.277 was promulgated in 1910 and defines the structure of forest administration and its tasks, state lands, and permitted actions in forest management (logging, reforestation, safety works). It then regulates silviculture activities and defines forestry education courses. The RDL 31/12/1923 (Serpieri law), "Riordinamento e riforma della legislazione in materia di boschi e terreni montani" is probably the most important forest law in Italy and it is still effective. It was focused on mountain areas, and specific for the soil and the mountain slopes. Some of the priority elements of the reading concerned the establishment of hydrogeological constraints, the creation of consortia and other management bodies, and the encouragement of reforestation. Between the post-war period and the 1980s, there was a succession of laws relating to forestry, both for management and commercial use; later, regional laws were passed allowing individual Regions to implement autonomous forestry plans, effectively decentralising state control. The landscape, including forestry, is one of the fundamental principles of the constitution (Article 9) and has been protected by various laws since the 1980s.

Law 431/1985 (Galasso law) subjected several categories of natural assets (including forests) to landscape restrictions. Subsequently, the promulgation of the Legge Quadro sulle aree protette (L. 394/1991) promoted sustainable management actions also within forests, protecting them through the implementation of planning tools. Legislative Decree 42/2004 protects cultural heritage, which includes both cultural and landscape assets. Article 142 transposes Law 431/1985, confirming the protection of landscape heritage, including forests; article 149 regulates the management and exploitation of forest resources. The provisions of legislative decree 227/2001 (Orientation and modernization of the forestry sector) aimed at valorisation of silviculture both in socio-economic and environmental terms, as well as the conservation and management of the Italian forest heritage, according to the Resolutions of the Interministerial Conferences on the protection of forest in Europe (held between 1990 and 1998).

On 20 April 2018, the Italian National law on forest and forest supply chain (Legislative Decree No 34, TUFF), which repealed Legislative Decree 227/2001, was published in the Official Gazette. Since 2001 there have been numerous European directives and regulations concerning climate, environment, landscape, socio-economic development, and other issues in which sustainable forest management plays a major role. Moreover, the TUFF unifies all regional laws, defining a national framework for the management and conservation of forests and their associated ecosystem services. The text of the law, in addition to providing directives on forest management and protection, gives the definition of forest (Article 3) and makes explicit the need for the elaboration of a geo-referenced national forestry map that must be aligned with European directives on cartographic products, with the purpose of ensuring greater accessibility to data (Article 15).

Finally, in 2022 the National Forest Strategy was published in the Official Gazette, which follows the directives of Legislative Decree 34/2018. The document is valid for 20 years and is an important tool to outline national forestry policies, also incorporating part of the activities foreseen by the European Biodiversity Strategy and 2030 Forestry Strategy and integrating the National Bioeconomy Strategy for the forest-wood system. The mission of the strategy is to create resilient forests for the country, rich in biodiversity and resilient to the effects of climate change, while also offering ecological benefits.

1.2. Land cover and forest ecosystem processes

1.2.1. Land cover and land use

The Directive 2007/02 of the European Commission define land cover (LC) as the biophysical cover of Earth's surface, as artificial surfaces, agricultural areas, forests and semi-natural areas, water and wetlands. Land use (LU) is the characterization of the territory according to its current and future planned functional dimension or socio-economic purpose (e.g. residential, industrial, forestry); it describe land using its socio-economic and ecological purpose. The current land use is the effective functions of a territory, while the planned land use refers to planning and management activities.

Land use and land cover are interconnected and affect each other, moreover, their changes impact natural resources and their balance (Sallustio et al., 2016). The land cover and land use characteristics are often combined to describe territories, leading to the creation of many classification systems with different aims and different level of detail. For example, the Corine Land Cover classification system is based on hybrid land use/land cover classes. It is important to solve the mismatch between different classification system to create a single, versatile and replicable classification method.

1.2.2. Forest

Forests are an essential resource of our Planet, they play a strategic role in terms of landscape, history, culture and economy (RAF, 2018). They contribute significantly to human well-being, providing a multitude of ecosystem services: wood production, carbon sink, food production, soil protection, climate control, biodiversity conservation, environment decontamination, hydrological cycle regulation (Gamfeldt *et al.*, 2013; Miura *et al.*, 2015; Spadoni *et al.*, 2020).

Thirty-one percent of world's total surface is occupied by forests (4.06 billion hectares) and Europe occupies a prominent position in this context, as a matter of fact European forests account 1.017.461ha (FAO, 2020). The Italian forest heritage, is one of the most important in Europe:

nowadays, almost 40% of the Italian land is wooded, amounting to 11.8 million (Munafò et al., 2018).
Nevertheless, forest heritage is at risk due to anthropogenic phenomena: some directly related to the action of man (for instance, deforestation), others as an indirect consequence of unsustainable human activities (for instance, climate change, drought) (Forzieri *et al.*, 2021).
Forests benefits are manifold and go beyond their boundaries; this is the reason why the Italian National Legislation (law no. 34, 2018, and law no. 42, 2004) undergo the forest asset to the joint competence of different administrations, recognizing it as "part of the natural national heritage", a public resource to protect and enhance following the sustainable development principles (TUFF, ART. 1) and including it among the areas of landscape interest protected by law (Legislative decree of the Italian Republic, 2004).
Over the years several forest definitions and parameters have been identified by institutions; the most relevant definitions are given below.

The Italian National law on forest and forest supply chain describes forest as "areas covered by forest vegetation […] with natural or artificial origin at any development and evolution, a surface not less than 2,000 m^2, average width not less than 20 meters and with a coverage of more than 20% of forest tree" (art. 3, comma 3).

According to Food and Agriculture Organization (FAO) definition, a wood is any "territory with arboreal coverage greater than 10% compared to an extension greater than 0,5 ha, where the trees reach a minimum height of 5m when ripe", including also: "forest nurseries and seed arboretums, forest roads, cut fractions, firebreaks; forests included in national parks, nature reserves and other protected areas; windbreak barriers and wooded strips of a width exceeding 20m, rubber tree plantations and cork trees" (FAO, 2001).

The Italian National Forest Inventory defined in 1985 "forest" as an area covered by trees or shrubs for at least 20%, minimum extension of 2,000 m^2 and minimum width of 20 m. Later on, in 2005, the INFC updated this definition, adapting it to the definition given by FAO, shifting so the minimum extension from 2,000m^2 to 5,000m^2 and the coverage percentage from 20% to 10% (INFC, 2005).

Changes in the forest areas over time reflect changes demanded by other land uses. Expanding the knowledge of the forest heritage is essential for managing and preserving forests. Following a collaboration principle, the activities of data observation, acquisition and analysis, are divided between institutional and research authority, in order to promote knowledge of the forest heritage.

1.2.3. Ecological succession

Vegetation has a structure that is subject to continuous changes, due to natural or anthropogenic factors, which can cause its specific composition to vary. In a plant community, the death of individuals leads to changes that are referred to as cyclical: for example, the opening up of a clearing in a forest due to a disturbance event leaves room for the establishment of new species, which will evolve back into the forest. What differs from cyclical changes are those changes that are referred to as succession, which can occur in a cleared area or constitute vegetation phases. A succession consists of phases that are repeated under constant climatic and soil conditions. This evolution of vegetation is thus made up of series of precise phytocoenoses, which over time replace one another, until they arrive at a specific composition in apparent equilibrium with the environment (called climax community). This situation is not immutable, as events of loss of the climax community can also occur, with the establishment of regressive series (Prach and Walker, 2011, Raven *et al.*, 2002).

The seral stage represents the gradual transition of a community from herbaceous, shrub and tree cover. Each stage constitutes a recognizable community with its own specific structure and composition at a given time; each stage can vary in duration, from a few to tens of years, and not all stages can always be present. For example, if a forest is formed from an abandoned agricultural area, the shrub stage may be missing or replaced by young trees. Species that develop in the early successional stages are called pioneer and are typically small species with a high dispersal capacity and a fast growth rate. Late successional species, on the other hand, are characterized by greater size and longevity and a lower growth rate than pioneer species.

When the community grows on an area where there has never been vegetation, such as rocks created by cooling lava or retreating glaciers, it is called primary succession. Gradually soil is created, and the plant begins to colonise the area, and eventually a climax community is formed. Secondary succession, on the other hand, is recognised as the establishment of new vegetation after a disturbance event. In general, primary succession consists of more species than secondary succession and is also less influenced by alien species. Furthermore, recovery rates have been shown to differ between ecosystems, highlighting the role of global climate change on community recovery rates and restoration management practices (Figure 1 Gerola *et al.*, 2006).

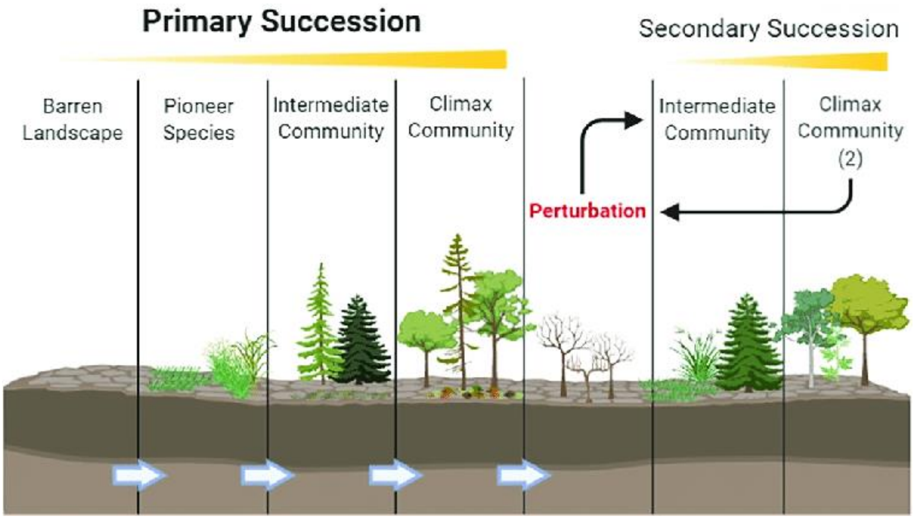

Figure 1 Ecological succession model. The primary succession (on the left) describes the colonization of barren soil, which is covered by pioneer species which are substituted by climax community at the end of the process. The climax community can be modified by a perturbation, in this case the final steady community will be different from the first climax community (secondary succession, on the right). Source (Khanolkar *et al.*, 2020).

Structural and community-specific changes are determined by many factors and can be divided into two groups: allogenic and autogenic changes. Autogenous change relates directly to the characteristics of the organisms in a community, e.g. the specific composition of a forest depends on the amount of light that penetrates the tree canopy; allogenous change relates to an environmental characteristic, e.g. temperature variation in mountainous elevations.

Autogenous change is common to all plant successions: for example, in the early stages of development the vegetation will be composed of heliophilous species, as the availability of light is greater, thus allowing the establishment of species that generally have a high rate of photosynthesis, which leads to fast growth. When the availability of light below the canopy decreases, the seedlings of heliophilic species fail to develop, leaving room for the sciaphilic species, which have high photosynthetic rates in the absence of light. They begin to germinate and, when the heliophilic plants in the upper layers die, they will be replaced by the sciaphilic species. Allogenic changes take longer than autogenic changes and generally last longer than the life of the organism itself, thus profoundly changing the appearance of the community, for example the evolution of a wetland into a forest (Smith and Smith, 2009).

1.2.4. Forest change processes

Forest change processes represent primary importance territorial transformations, and they have a central role in the definition of policies that aim to preserve forest resources and its ecosystem services provision. International organizations and other institutions at different levels have undertaken forest changes issues, giving definitions to frame this phenomenon. United Nations Food and Agriculture Organization (FAO) Forestry Department established a series of definitions, according to which, considering changes between forest and other land use classes, it is possible to distinguish different processes considering temporal and canopy cover thresholds. In particular:

"Deforestation is the conversion of forest to another land use or the long-term reduction of tree canopy cover below the 10% threshold";

"*Afforestation* is the conversion from other land uses into forest, or the increase of the canopy cover to above the 10% threshold";

"Reforestation is the re-establishment of forest formations after a temporary condition with less than 10% canopy cover due to human-induced or natural perturbations";

The expressions "long-term" and "temporary" refer respectively to more and less than 10 years. In this sense, afforestation appears to be the reverse of deforestation, while reforestation is presented as a short-term process that, if longer than 10 years, falls within the definition of afforestation (FAO, 2001, Figure 2).

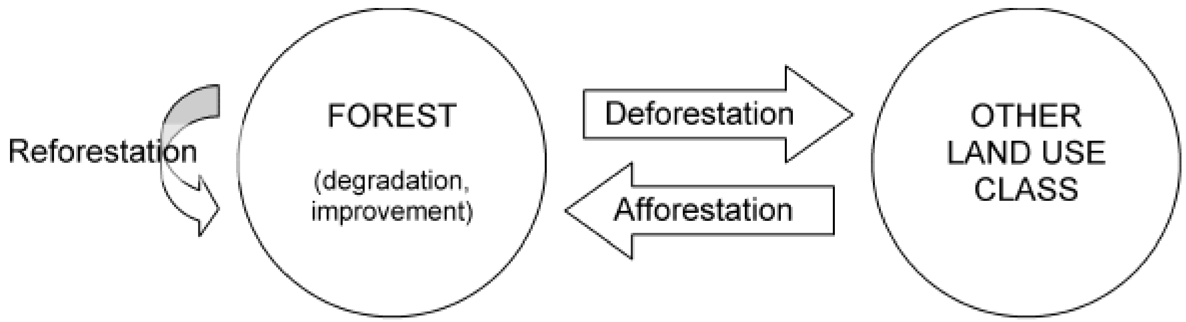

Figure 2 Deforestation, Afforestation and Reforestation definitions scheme. Source: Food and Agriculture Organization, 2000.

These three different phenomena occur heterogeneously around the globe, mainly because of different socio-economic backgrounds. While in Africa and South America deforestation trends up, in Europe afforestation, generally resulting from land-use change from agricultural use to forest use, represents broadly the dominant process (Angelucci, 2011 Gálos et al., 2013; FAO, 2001; Palmero-Iniesta et al., 2020; Figure 3).

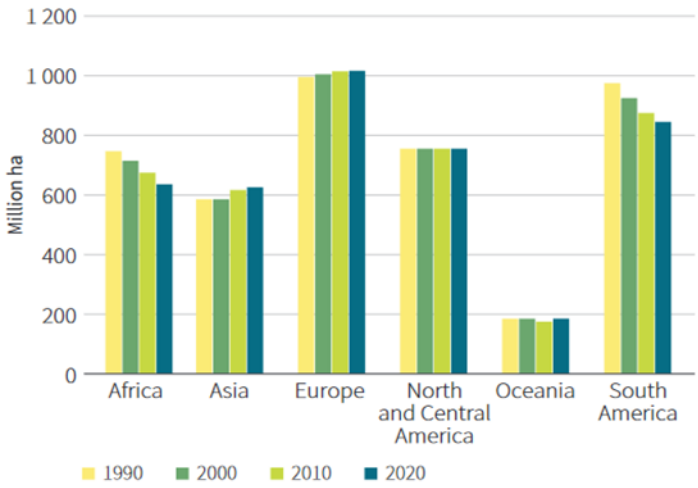

Figure 3 Changes in forest area in the World, subdivided in the decades from 1990 to 2020 (FAO, 2020) the figure illustrates a forest gain in Asia, Europe, North and Central America and Oceania, while in Africa and South America the forest decreased its surface between 1990 and 2020.

This general tendency though does not apply uniformly throughout the continent, where different trends may occur depending on the geographical area considered. It can be observed, for instance, a marked difference of forest land evolution between Eastern Europe, where there is, since the fall of the Soviet Union, prevalence of deforestation (Feranec *et al.*, 2016), and the European Mediterranean area, mainly characterized by afforestation processes (Cervera *et al.*, 2019).
For what concerns the Italian territory, forest land changes trend is in line with the other Mediterranean regions and with the overall tendency of the rest of the continent. Historical cartographic and inventory data confirms this assumption showing that forest surface has been constantly increasing through the years, with an average annual growing rate of 0,3% (Ferretti et al., 2019; Angelucci, 2011; Table 1).

Table 1 Italian forest surface evolution (Ferretti *et al.*, 2019).

Year	*Document*	*Forest surface (ha)*
1936	Forest cartography of the Italian Kingdom	6,028,301
1985	First forest inventory (IFNI85)	8,675,100
2005	Second forest inventory (INFC05)	10,467,533
2012	Corine Land Cover map CLC1	9,973,516
2016	Italian Land use inventory	11,732,622

Afforestation represents a process of increasing importance, which certainly still needs to be characterized to make the most of its potential and to correctly drive its development in line with forest sustainable management principles. As on one hand, in fact, afforestation certainly involves all the benefits connected to forest ecosystem services, on the other hand, if not well managed, it can lead to land fragmentation and other related issues (Czimczik et al., 2005; UNFCCC, 2013). In other words, most of the issues and challenges associated to afforestation depend on its location (FAO, 2001).

1.3. Remote sensing and land monitoring

Remote sensing (RS) is the set of techniques useful for acquiring information on objects placed at a distance from the observer, which is based on the properties of the interaction between electromagnetic energy from an energy source (emitted, reflected or transmitted) and radiation measurements. By observing the interactions between the earth's surface and the energy source, it is then possible to analyse the characteristics of the elements. An initial classification of remote sensing is based on the functionality of the sensor used and distinguishes between *passive remote sensing*, in which the sensor receives the electromagnetic radiation emitted or reflected by the object, and *active remote sensing*, in which the sensor emits the electromagnetic radiation and also detects the fraction that is reflected by objects on the surface.

1.4. Satellites

Among the different methodologies to assess and to characterize land cover, remote sensing certainly represents a high potential instrument (Agrillo *et al.*, 2021; Filipponi *et al.*, 2018; Koch, 1999). It can allow in fact to analyze and classify large extension areas in very reduced time enabling to obtain high resolution and easily updatable results (Agrillo *et al.*, 2021; Koch, 1999).
RS from satellite platforms stands out for its great versatility and for the large amount data produced (Laurin *et al.*, 2021). *Landsat* and the more recent Copernicus European program *Sentinel* satellite images are by far the most used, being open source data and covering many years of observations (Jönsson *et al.*, 2018).

In particular, *Landsat* and *Sentinel* images find different applications considering their intrinsic characteristics. As regard the multispectral data *Landsat* are in fact available since 1970s, providing so very long time data series useful to analyze land cover trends and evolutions. *Sentinel-2* instead, available since 2015, have, as their strength point, the high spatial resolution (10 meters against the 30 meters of *Landsat*) and their short revisiting time (2-3 days against the 16 days of *Landsat*), that make it a tool of great utility for applications where a high precision is needed. *Sentinel-1* is a

Synthetic Aperture Radar (SAR), which operates in band C and, since it is an active system, it can operate both day and night and in every weather conditions, in-fact it is not influence either by light or by cloud cover. *Sentinel-1* exploit the different signal polarization of the structural properties of elements, depending on the roughness and the moisture content.
These products provide, both multispectral and radar data, supplying thus various information about monitored land cover relatively at different bands of the electromagnetic spectrum.

The Landsat satellites are part of a NASA mission that began in the 1960s and consist of a fleet of eight satellites starting in 1972 with the launch of Landsat 1 (ended in 1978), followed by Landsat 2, 3, and 4 between 1975 and 1978 and dismissed in 1983 and 1993, Landsat 5 in 1984 (dismissed in 2013), Landsat 6 in 1993 (due to mission failure this satellite is not included in the list), and the more recent Landsat 7 (1999), 8 (2013), and 9 (2019) satellites, which are still active.
There are two reference grids, called the WorldWide Reference System (WRS): WRS-1 descending had a total of 251 paths and 119 rows where each intersection of a path and a row denotes the nominal center of a Landsat scene. WRS-2 descending has 233 paths and 124 rows for a sixteen-day ground coverage cycle. Landsat 1 through 3 used WRS-1, and Landsat 4 through 9 uses WRS-2. Each day, Landsat 8 and 9 together capture 28 swaths, or paths, cumulating to a full image of Earth every 8 days.

Landsat 1 and 2 used a Return Beam Vidicon Sensor (RBV) and a Multispectral Scanner System (MSS): the RBV sensor is composed of three bands (blue-green, orange-red, red-near infrared) with wavelength included between 0.5 and 0.8 µm and 80 m spatial resolution. The MSS sensor is composed of seven bands (green, red, and two near infrared), with wavelength between 0.5 and 1.1 µm and 80 m spatial resolution.

Landsat 3 was equipped with an RBV and an MSS sensor, the latter similar to the previous satellites, while the RBS sensor operated in a wide spectral band, from green to near infrared (0.505-0.750 µm).

Landsat 4 and 5 had an MSS sensor, similar to the previous satellites, and a Thematic Mapper (TM) sensor, composed of seven bands (blue, green, red, two near infrared, thermal infrared and mid infrared); the wavelength is between 0.45 and 2.35 µm and 30 m spatial resolution, except from the thermal infrared band which is 120 m spatial resolution.

Landsat 6 mounted an Enhanced Thematic Mapper (ETM) sensor, acquiring in eight bands, the first seven similar to the TM sensor of Landsat 4 and 5, to which was added the Panchromatic band (0.52-0.90 µm), at 15 m.

Landsat 7 mounts an Enhanced Thematic Mapper Plus (ETM+) sensor, acquiring in 8 bands (blue, green, red, two near infrared, thermal infrared, mid infrared and panchromatic), with a spatial resolution of 30 m, 60 m for thermal infrared and 15 m for panchromatic:

Landsat 8 e 9 are equipped with two sensors: Operational Land Imager (OLI), with nine spectral bands (coastal aerosol, blue, green, red, near infrared, SWIR-1, SWIR-2, panchromatic and cirrus) at 30m spatial resolution (15 for panchromatic), and a Thermal Infrared Sensor (TIRS), with two spectral bands at 100 m spatial resolution.

Sentinel-2

The Copernicus Sentinel-2 mission (started in 2015) is composed by two polar orbiting satellites (Sentinel-2a and Sentinel-2b) which have the same sun-synchronous orbit and are phased at 180° to each other. The swath width is 290 km and the revisit time is 2-3 day with two satellites at mid-latitudes.

Sentinel-2 is equipped with a MSI sensor with 13 spectral bands, four at 10 m spatial resolution (blue, green red and NIR), six at 20 m spatial resolution (four vegetation red edge and two SWIR) and three at 60 m spatial resolution (coastal aerosol, water vapour and SWIR-cirrus).

Sentinel-1

As the Sentinel-2, the Sentinel-1 mission (started in 2014) is composed by two polar-orbiting satellites, which operates day and night regardless weather conditions. The swat is 250 km and the revisit time is 2-3 days at mid-latitudes.
The radiation emitted or reflected by each material on the Earth's surface varies as a function of the wavelength: in other words, every material is characterized by a specific and measurable spectral signature.

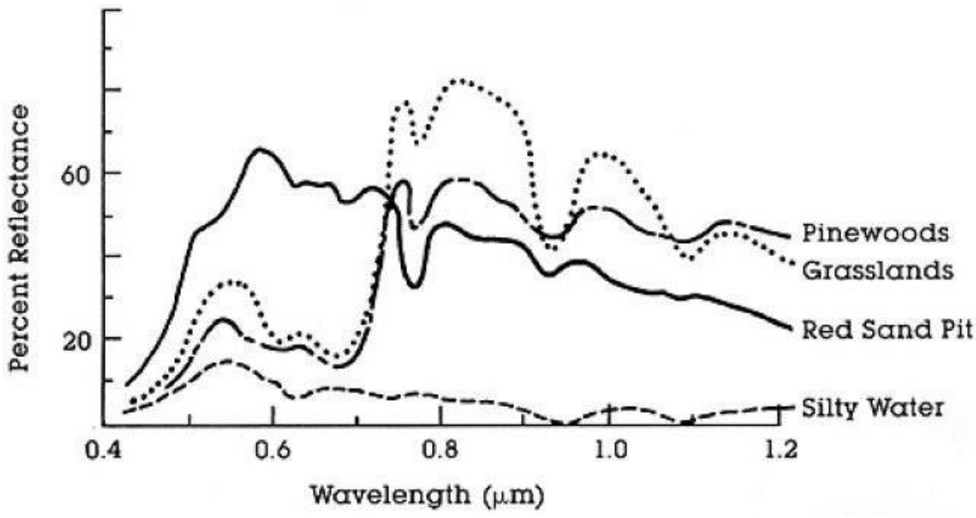

Figure 4 Spectral signatures. Source NASA, 2013. Landsat 7 Science Data User's Handbook.

To simplify the process, often the single bands are summarized in single values called Spectral Indices or, when they are related to vegetated land cover, Spectral Vegetation Indices (SVI) (Spisni *et al.*, 2012, Oberti, 2009). The most used SVI in literature are Normalized Difference Vegetation Index (NDVI) (Jönsson *et al.*, 2018), Enhanced Vegetation Index (EVI) (Bajocco *et al.*, 2019), Normalized Burn Ratio (NBR) (Roy *et al.*, 2006) and Plant Phenology Index (PPI) (Karkauskaite *et al.*, 2017).

Besides the SVI, other indices, known as Tasseled cap, represent a valuable tool for this kind of analysis. These indices, obtained as linear combinations of bands, resume multispectral data in direct information about three dimensions of the land cover considered: Brightness (B), Wetness (W), Greenness (G) and Angle (A). Even though they might appear similar to SVI, Tasseled cap do not constitute redundant data, adding instead information that are not measured by SVI: Brightness reflects information about albedo (bright-dark ratio of a given surface); Wetness is influenced by vegetation and soil moisture content; Greenness is referred to the presence of vegetation (Zhang et al., 2002); Angle refers to the proportion of vegetation/non vegetation and it is the angle created by B and G (Gómez *et al.*, 2011).

SVI and Tasseled cap acquire even more value when referred to different time periods, allowing to build temporal trends which represent important additional information about the phenology of a given territory (Persson *et al.*, 2018). Hence, to elaborate a land cover classification and to produce forest cartography, an efficient method consists in analyzing multitemporal series of SVI or Tasseled cap, so to dispose of very complete information in a synthetic way. In doing so, it is possible to establish classification rules by imposing thresholds to discriminate different land cover classes and recognize the ones of interest (Luti et al., 2021; Smiraglia et al., 2020; Spadoni et al., 2020).

Among the different classification methods, supervised classification, which works automatically but is previously “trained” by an operator, is very effective since it allows to elaborate a large amount of data and return accurate results (De Fioravante et al., 2021). For what concerns the automatic data processing stage, among the most largely used algorithms, machine learning approaches can elaborate many data reaching a very high accuracy in a short time. Rule based approach and random forests, which belongs to the family of the decision tree algorithms, are examples of these approaches and currently one of the most used for forest classification (Agrillo *et al.*, 2021).

1.4.1. Copernicus Program

Monitoring activities widely exploit the potentiality of the European Earth Observation Program Copernicus, which, according to EU Regulation n. 377/2014 "provides information on the state of the atmosphere, oceans, land, in support of climate change adaptation and mitigation policies and emergency management and civil security". The Copernicus Program is, therefore, a complex set of systems that collects information from multiple sources, namely satellites and ground, sea and airborne sensors. Copernicus integrates and processes all these resources providing to institutional, research and industrial users, reliable and up-to-date information through a series of services related to environment, territory and security.

The Program is divided into two main structures: the services and the space component. The services are divided into thematic areas (Core Services): monitoring of land, sea and oceans, atmosphere, climate change, emergency management, security and in-situ. The space component, consisting of satellites, associated ground infrastructures and data acquisition from third party providers, is managed and developed by ESA with the help of the European Organization for the Exploitation of Meteorological Satellites (Eumetsat).

Within the Copernicus program, all service data and images acquired by the Sentinel satellite constellation are made available completely openly and free of charge. These satellites provide high-resolution radar and optical images of our planet, allowing the monitoring of the territory (land cover, soil and water, etc.), the sea (temperature, sea surface trend, etc.) and the atmosphere. For the purposes of land monitoring by SNPA, are now widely used multispectral images Sentinel-2 (13 bands) characterized by a high revisit time (3-5 days) and a resolution between 10m and 60m, and the radar images of Sentinel-1 missions (Synthetic Aperture Radar - SAR - C-band).

In Italy, the Space Economy Strategic Plan, which stems from the work of the Cabina di Regia Spazio (Space Steering Committee), an initiative promoted by the Presidency of the Council of Ministers for the definition of national policy in the space sector, together with the provisions of the National Recovery and Resilience Plan (PNRR), aim to use the space sector as one of the driving forces of the new growth of the country. The system, through the realization of innovative enabling infrastructures/systems and the creation of national operational structures, is based on Big Data platforms for the storage, processing and integration of space data with other observational and forecast data from models, able to break down the barriers of access to information useful for the provision of information services for institutional and private users. There are many applications that can be enabled by these platforms, such as the National Operational Infrastructure for environmental monitoring in support of SNPA, with operational services related to monitoring of the territory, the main environmental resources and the built environment, maritime surveillance, nowcasting (short-term marine weather forecasting), precision agriculture and its impacts on land and soil (Munafò, 2021).

1.4.2. Land Monitoring Service

The Land Monitoring Core Service (LMCS) provides geographical information on land cover and several variables related to vegetation status and water cycle and is composed of 3 main components: the Global component coordinated by the JRC, which produces global scale data, the Pan-European and the Local component, both coordinated by the European Environment Agency. In addition to the 3 components, the LMCS service supports Reference Data related to in-situ data needed for Copernicus services.

1.4.3. Global component

The Global component of LMCS (https://land.copernicus.eu/global), also referred as the Copernicus Global Land Service (CGLS), is composed of four elements:

1. a global scale monitoring service, that provide bio-geographical information at global scale, at medium spatial resolution and furnish consistent time series of data;
2. a Hot Spot mapping and validation service, which is available upon request, and it allows to analyze limited geographical areas with high resolution satellite data;
3. an element dealing with Ground-based Observation for Validation (GBOV), that is used to validate products based on satellite observation;
4. support to the Global Agriculture Monitoring program of the Group of Earth Observation (Copernicus4GEOGLAM); this activity is on-request.

1.4.4. Pan-European component

The European Environment Agency (EEA) coordinates the pan-European component (https://land.copernicus.eu/pan-european), among which different Biophysical parameters are available. These parameters are CORINE Land Cover datasets, High Resolution Layers and High Resolution Snow and Ice products. In 2021-2020 other products are expected to be published, as High Resolution Phenology and Productivity, CLC+ and European Ground Motion Service.

The CORINE Land Cover (CLC) was developed in 1985 and published in 1990; it was updated in 2000, 2006, 2012 and 2018. It is composed by 44 land cover and land use classes (which identify both artificial and natural areas) classified in three hierarchical levels; the Minimum Mapping Unit (MMU) is 25 hectares (ha) and the Minimum Width of 100 m. There are also change layers, focused on land cover changes, with a MMU of 5 ha.

Considering the class 3 (Forest and semi-natural areas), the tree cover is classified in broadleaves, conifers and mixed forest, depending on the predominant vegetation type. The 324 class (Transitional forest-shrub), represent forest degradation, forest regeneration / recolonization or natural succession, giving information on areas representing natural development of forest formations, consisting of young plants of broad–leaved and conifers species, with herbaceous vegetation and dispersed solitary adult trees. Transitional process can be for instance natural succession on abandoned agricultural land, regeneration of forest after damages of various origin (e.g. storm, avalanche), stages of forest degeneration caused by natural or anthropogenic stress factors (e.g. drought, pollution), reforestation after clearcutting, afforestation on formerly non-forested natural or semi-natural areas.

The High Resolution Layers are available since 2006 (update 2009, 2012, 2015, 2018) and they are a complementary to other land cover/land use dataset, as CLC. The latest version (2018) is 10 m spatial resolution, while the spatial resolution of the other dataset is 20 m. The classification system includes five classes: imperviousness, tree cover density and forest type, wetness and water, grasslands and small woody features. The forest class is composed by a set of products that provide information on coverage density and distribution of conifers and broadleaved trees. There are three products referred to the forest status:

Dominant Leaf Type (DLT) is developed using Random Forest algorithm on Sentinel-2 images; it describes the dominant species (conifers or broadleaves);

Tree Cover Density (TCD) is referred to the percentage of the pixel surface covered by trees; it describes forest/non forest cover;

Forest Type identify forest following the FAO definition; it is 10 m or 100 m spatial resolution;

Tree Cover Change Mask (TCCM) and Dominant Leaf Type Change (DLTC) are referred to the forest changes between 2015 and 2018, and they are 20 m spatial resolution.

1.4.5. Local component

Together with pan-European the Local component is coordinated by the EEA, aiming to provide complementary information obtained through pan-European component. It focuses on specific areas and it is based on very high resolution images (2,5 m spatial resolution). It is composed by the following four products:

Urban Atlas provides information on specific Functional Urban Areas (FUA) for 2006,2012,2018. In Italy UA covers an area of 68,137 km^2, equal to 22.6% of the national surface. It is composed of a four level classification system with 27 classes, which identify the forest cover similarly to the CLC classification.

The Riparian Zones (RZ) is focused on areas along rivers. The classification system is composed of 56 classes organized on four levels: the forest, as in CLC, identify the vegetation type (broadleaved, conifers and mixed forest) and also the transitional areas. In Italy RZ covers 40,273 km^2, equal to 13.4% of the national surface.

Natura 2000 is referred to Birds and Habitat Directive species and habitats; it is available for 2006 and 2012 ad it is composed of 55 thematic classes. The forest class is similar to the RZ classification. In Italy N2000 covers 77,703 km^2, equal to 25.8% of the national surface.

Coastal Zones contain information on land cover and land use for areas included in 10 km from coastline. It is available for 2012 and 2018 both for the status and the changes. The classification system is composed by 71 classes and the forest class correspond to the RZ one.

Even if the local component products have and higher spatial resolution than pan-European component, they do not cover the entire territory, being available only for specific areas.

The EEA has recently developed and implemented a new set of mapping products, which refer to Corine Land Cover Plus (CLC+). These products aim at improving land use and land cover information to support land management. Two products can be distinguished: CLC+ Backbone, which is composed of an 18-class vector product and a 12-class raster product, having a spatial resolution of 10 m. These products are based on the EAGLE model definitions for land use and land cover and are complementary to the available CLMS products. The CLC+ Core maintains the hybrid land use and land cover classification and combines existing and future CLMS products by breaking them down into the components of the EAGLE matrix; is a database which will store the pan European and local component provided by EU Member States. The CLC+ Instances and CLC+ Legacy are part of the CLC+ Core database. CLC+ Instances harmonizes the CLC+ Core products according to EAGLE, while CLC+ Legacy includes updates of existing CLC products.

1.4.6. Eagle concept

To satisfy the purposes defined by the European legislation for monitoring activities, both at European and national level, a list of criteria for new data model has been defined and the EIONET Action Group on Land monitoring in Europe (EAGLE) Concept, aims at being a tool for class definition and for linking present and future nomenclatures (Arnold *et al.*, 2013). The EAGLE Concept does not represent a classification system, it is instead a tool for unifying LC and LU information. It is based on the distinction of land cover and land use, and it is structured in a matrix composed by three descriptors: Land Cover Component (LCC), Land Use Attribute (LUA) and further characteristics (LCH). These elements can be combined to create different classification systems or to identify similarities with existing classes, maintaining the descriptors independent.

More specifically, LCC refers to the land cover definition of the INSPIRE directive 2007/02 CE. Every LCC is unique, so they can be combined to deeply describe different land cover.

The LUA is based on the Hierarchical INSPIRE Land Use Classification System (HILUCS) and it has been adapted to the EAGLE Concept. Finally, LCH give further information on the Land Cover Component.

EAGLE allow to better understand the characteristics of different classes, also considering the conversion between different classification system. The presence of such a classification concept makes it possible to integrate national and European land monitoring tools and to produce new data that can be easily integrated at different classification levels.

1.4.7. National land cover maps

The study of land use and land cover is fundamental to understand the causes and effects of human activity on the territory. As explained before, in the European context, the Copernicus Land Monitoring Service (CLMS) offers a wide range of spatial data to support monitoring activities, but with limitations in terms of geometric, thematic or temporal characteristics. The CORINE Land Cover is a reference for the analysis of land use and land cover at national and European scale, but the low spatial resolution (Minimun Mapping Unit equal to 25 hectares) and the low frequency of updating (6 years) limit its applicability. More recently, the introduction of raster data High Resolution Layers (HRL) has allowed to have information with a high spatial detail (pixels of 20 meters in the versions of previous years and 10 meters for the most recent update to 2018) for a limited number of land cover classes. Within the local component of CLMS there are several data with high spatial and thematic resolution, but which provide information related to limited and specific spatial areas, such as Urban Atlas, Riparian Zones, Natura 2000 and the most recent Coastal Zones.

To overcome these issues, ISPRA has developed some preliminary products of land cover, aiming at the realization of cartographic products able to allow analysis on the entire national territory.

Two different cartographic data were created, one using the data of the Local component of LMCS, other using a semi-automatic classification of Sentinel-1 and Sentinel-2 images, which is part of this thesis.

As regard the first product, the main data of the Local component of CLMS have been analyzed, reclassified, converted into raster and finally merged with CORINE Land Cover, HRL and the national map of land consumption. For the LMCS data, 2018 was considered as reference year. The same process was applied to the 2012 data, resulting in two land cover and land use maps that are useful for analyzing land cover changes and assessing the change in associated phenomena, such as ecosystem services.

A further important synthesis effort involved the definition of a classification system capable of harmonizing and preserving the information offered by the input data, while maintaining consistency with the specifications of the EAGLE group regarding the separation between land use and land cover. The map has a spatial resolution of 10 meters and is based on a hierarchical classification system with four categories, defined from the structure of the EAGLE matrix.

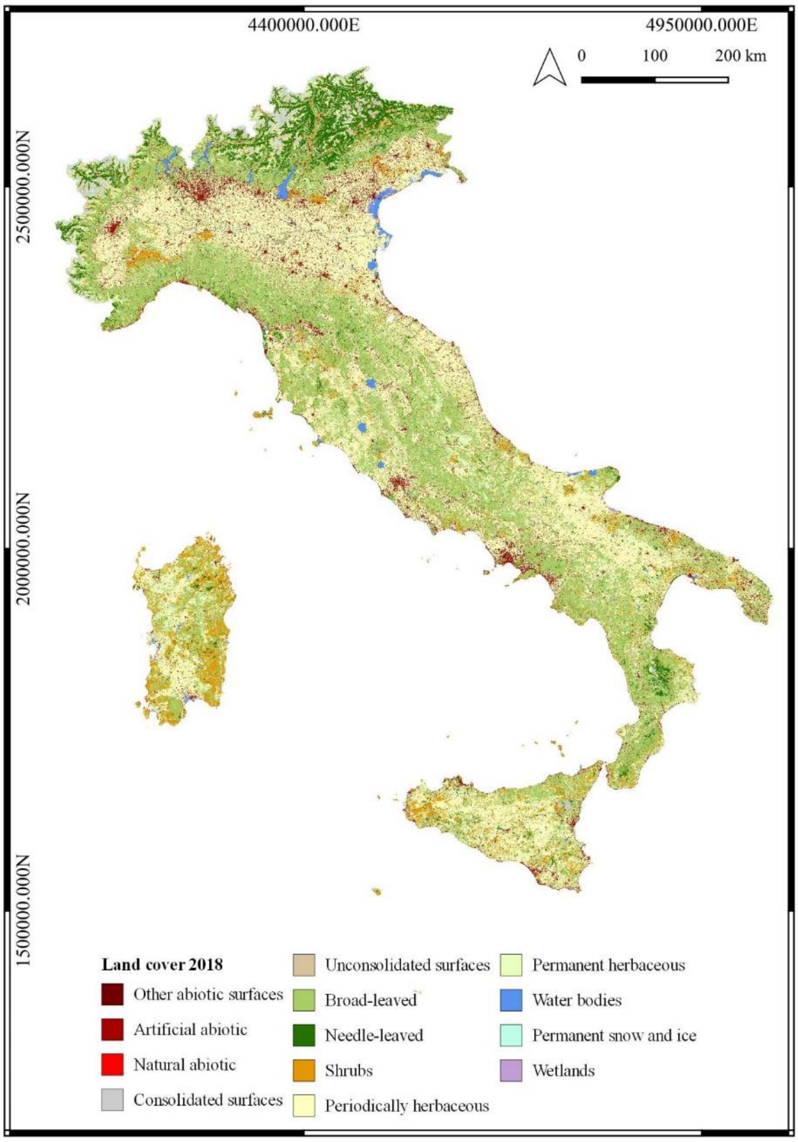

Figure 5 National land cover map and classification system. Source ISPRA-SNPA, 2022

The land use national map, recently published by ISPRA, refers to 2018 and provides mapping for the entire national territory. The map is obtained from the integration of regional land use maps (Valle d'Aosta, Lombardy, Veneto, Liguria, Emilia Romagna, Tuscany and Lazio), forest maps (Piedmont and Valle d'Aosta) and data from the Land Monitoring Service of the Copernicus Program, with reference to the Local (Urban Atlas, Coastal Zones, Natura 2000) and Pan- European (CORINE Land Cover) components. It is composed by eighteen land use classes, which are based on the Land Use Component of the EAGLE matrix (Figure 6).

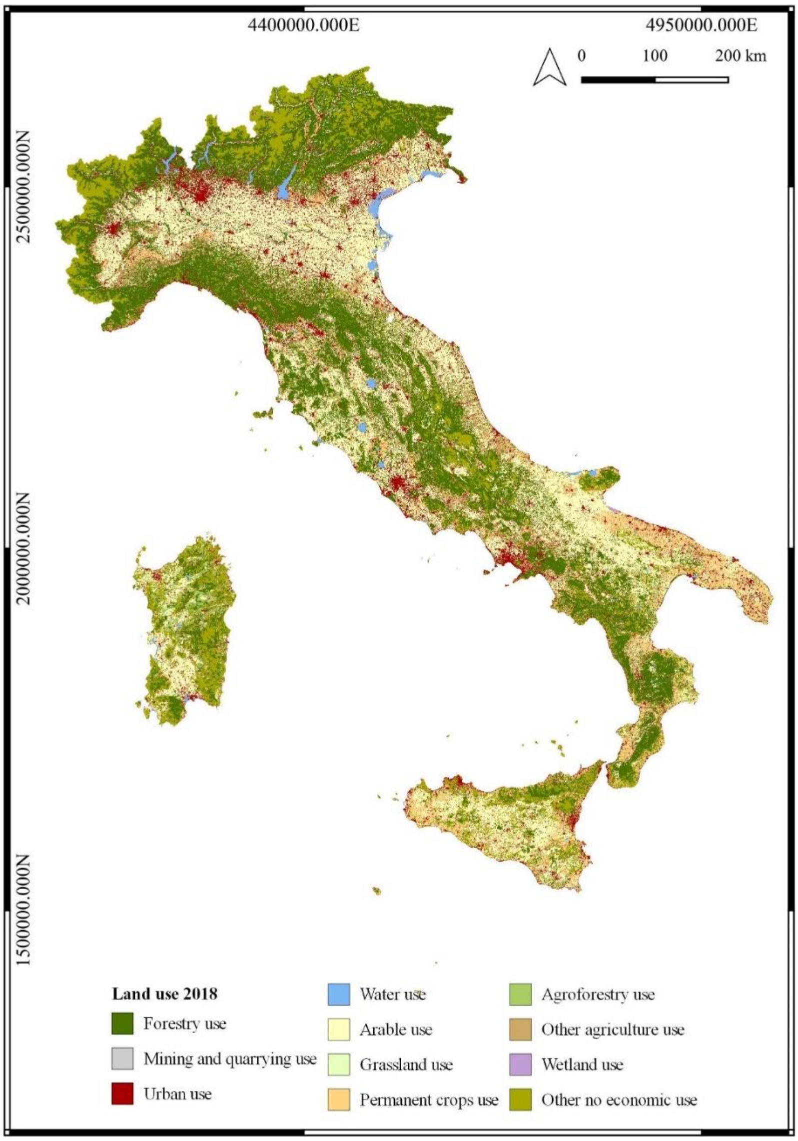

Figure 6 National Land Use map and classification system. Source ISPRA-SNPA, 2022

1.4.8. Inventory data

In addition to national land cover cartographic data there are inventory data, which can be integrated to cartographic data (Sallustio *et al.*, 2016). One of the most important is the Land Use Inventory of Italy (IUTI, Figure 7), an inventory supporting the National Register of Agroforestry Carbon Sinks of agroforestry carbon sinks developed in 2010 by the Ministry of the Environment, Land and Sea. Protection of Land and Sea and updated in recent years by ISPRA and the University of Molise. The hierarchical classification system used in IUTI is based on the 6 categories of land use defined by Gpg-Lulucf (Good practice guidance for land use, land-use change and forestry), which are: forest land, cropland, grassland, wetlands, settlements and other land. Forest land, cropland and grassland are integrated with second and third level subcategories, for a total of 9 classes (Marchetti *et al.*, 2012), subsequently augmented for specific thematic insights with respect, for example, to the agricultural sector. To identify forest from other wooded land or grassland, the FAO definition (2001) is applied. From the original sampling system that involved the photointerpretation of 1.2 million points over the whole national territory at 1990, 2000 and 2008, a subsampling system was subsequently developed that a sub-sampling system that has allowed, in the face of a reduction of about 100 times the sampling effort, to obtain estimates that are still accurate and reliable at the national scale. Although IUTI is a valuable source of information, it is seldom updated (1990-2000-2008-2013-2016) and provides information on the predominant land cover in the area of analysis, having with a minimum mapping unit of 5,000 m^2 and a minimum width of 20 m.

Figure 7 IUTI dataset.

The National Forest Inventory (INFC) represents the main source of information, at the national level information on the size and characteristics of the country's forests. It is integrable with IUTI, allowing a thorough exchange of information (Marchetti *et al.*, 2012). The survey is repeated at regular intervals of about 10 years, to allow the updating of forestry statistics on a national and regional basis. In this way it is possible to monitor over time the changes that affect the forest area and the characteristics of Italian forests, as well as providing the necessary information to implement effective policies effective policies for the protection of natural resources and for the protection of the services produced by forest ecosystems. The first Italian national forest inventory was carried out in the period 1983-85, followed by the first update in 2005.

During 2017, the activities of ground survey of the third inventory Italian national forest inventory - INFC2015, whose conclusion is expected by the end of 2019, have been started. The first phase of the survey was carried out in 2013 and led to the classification land use and land cover of more than the 301,000 INFC sampling points, through the observation of color orthophotos. The results of this first phase are represented by preliminary estimates of the forest surface forest area in the different Italian administrative Regions. The forest statistics produced by the INFC are disaggregated for the 21 administrative districts and for the different levels of the classification system: the inventory macro-categories (forest and other wooded land), the inventory categories (high forests, plantations of arboriculture and areas temporarily deprived of stand for forest), five categories for other wooded land and the 23 forest categories, which are differentiated by the tree species or group of prevailing tree species (spruce forests, beech woods forests, beech forests, oak forests, etc., RAF, 2018).

1.5. Ecosystem services

1.5.1. Classification and forest ecosystem services

Ecosystem services are defined as the benefits human populations derive, directly or indirectly, from ecosystem functions (Costanza *et al.*, 1998). There are several classifications of ecosystem services, the main ones derived from three European and international projects, namely the Millennium Ecosystem Assessment (MEA, 2005), The Economics of Ecosystems and Biodiversity (TEEB, 2010) and Common International Classification of Ecosystem Services (CICES -(Haines-Young and Potschin, 2018). MEA and TEEB classify four groups of ecosystem services:

- Supply, they provide the products such as food, raw materials
- Regulation, are the benefits derived from the processes that take place in ecosystems, such as carbon storage
- Cultural, are the non-material benefits that people can obtain from ecosystems (well-being)

- Support services are indirect services that support the production of other services.

Forests are vital to the lives of most terrestrial species, but they are also a resource for humans, providing food, water and raw materials, but also providing aesthetic and recreational benefits.
For these reasons, forests provide numerous ecosystem services, such as biodiversity conservation, climate regulation, soil conservation and prevention of soil degradation and desertification, regulation of the hydrological regime and recreational services. One of the most important roles of forests is their role in the carbon cycle, removing carbon dioxide from the atmosphere and converting it to nutrients during growth and releasing it during senescence. Forests therefore act as a source or sink of carbon, having the capacity to sequester and store it in the soil. This potential makes forests and their expansion a key tool in climate change mitigation, and the actions taken at the international level aiming to end deforestation by 2030 (COP 26) and reforestation, including in urban areas, are therefore crucial (UNFF, 2018).

2. Objectives

This study is part of land cover and land use monitoring initiated by ISPRA and to the monitoring of forests and afforestation, in collaboration with ISPRA, the University of Molise and the University of Florence. In the last phase of the research Prof. Ronald E. McRoberts of the University of Minnesota was involved. The objectives of this research are the following: (i) to develop a land cover and land cover change map of Italy using remotely sensed data, (ii) to develop and verify a new methodology to map afforestation occurred in the last 35 years using remotely sensed data, (iii) to construct a map of afforestation occurred between 1985 and 2019 in Italy, (iv) to assess the afforestation areas using statistically rigorous procedures, (v) to estimate the potential carbon sequestration in afforestation areas, (vi) to define the type of tree vegetation most affected by afforestation between 1985 and 2019.

The research is composed of four main parts. First (section 3), the national land cover map was constructed using Sentinel-1 GRD and Sentinel-2 images. It represents the land cover in Italy in 2018, considering also some of important land cover change, occurred between 2017 and 2018 and aims to provide land cover and land cover change information for operational purposes. This product is completely based on free Sentinel-1 and Sentinel-2 data, it allows annual updates and it can be compared to European and national products, since it is constructed considering the spatial resolution of many of the existing maps (10 m) and the classification is based on the EAGLE concept, which allows to unify different classification systems. The map was developed using a decision rules classification algorithm, to produce a pixel-based land cover map. The classification system is aligned to the EAGLE concept, distinguishing different land cover and land cover change classes, in which the forest class in divided in broadleaves and conifers cover. The forest cover is related also to forest disturbances analysis, in particular forest fires and other type of disturbances. The land cover is referred to 2018, while the land cover changes are referred to 2017-2018. The overall accuracy of the map is 83% and almost 90% for broad-leaved and conifers cover.

This research was published in 2021: De Fioravante P., Luti T., Cavalli A., Giuliani C., Dichicco P., Marchetti M., Chirici G., Congedo L., Munafò M., (2021) Multispectral Sentinel-2 and SAR Sentinel-1 Integration for Automatic Land Cover Classification. Land. Vol.10 pp. 1-35.

Second, the possibility of classifying areas of afforestation, non-forest, and forest using remotely sensed data was tested (section 4). This analysis follows the construction of national land cover map and is the preliminary analysis for the construction of the national afforestation map (section 5). For this, a database of 61 polygons was constructed through photointerpretation, identifying areas that have been subject to afforestation over the past 35 years. Landsat images (1985-2020) were then acquired and preprocessed. Using the training dataset and the Landsat images, 234 variables (derived from the association of different statistical parameters and different Landsat bands) were analyzed, allowing us to distinguish afforestation from non-forest and forest with high accuracy. Through this preliminary analysis, it was possible to confirm the usefulness of the remotely sensed data in classifying afforestation.

This research was published in 2022: Cavalli A., Francini S., Cecili G., Cocozza C., Congedo L., Falanga V., Spadoni G.L., Maesano M., Munafò M., Chirici G., Scarascia Mugnozza G. (2022) Afforestation monitoring through automatic analysis of 36-years Landsat Best Available Composites, iForest, vol 15, pp. 220-228.

Third, following this analysis feasibility check, we proceeded with the construction of the map of afforestation areas between 1985 and 2019 in Italy (section 5). In this research we analyzed the afforestation which brought to the present forest cover. Following the approach of the first analysis, a training dataset of 1,578 polygons identifying areas of afforestation, non-forest and forest was constructed. Using Landsat imagery and the training dataset, it was possible to classify the land into afforestation and non-afforestation, resulting in a map at 30x30 m spatial resolution. This map, served as the basis for selecting a sample of about 4,000 random points that were photointerpreted and applied to a stratified estimator, to obtain estimates of the area under afforestation at the national, regional and according to elevation bands. Finally, using estimates of afforestation areas and carbon content data in the four forest pools (derived from INFC 2005), it was possible to estimate the potential carbon sequestered in afforestation areas once the forest reaches maturity.

This research was published in 2023: Cavalli, A.; Francini, S.; Mcroberts, R.E.; Falanga, V.; Congedo, L.; Fioravante, P. de; Maesano, M.; Munafò, M.; Chirici, G.; Mugnozza, G.S. (2023) Estimating Afforestation Area Using Landsat Time Series and Photointerpreted Datasets. Remote Sensing, 15,932.

In the last part of the research, by intersecting the tree cover class data from the national land cover map and the 4,000 points used to estimate afforestation, it was possible to identify which type of vegetation (broadleaves and conifers) was most affected by the afforestation phenomenon. This result is illustrated section 6.

3. Multispectral Sentinel-2 and SAR Sentinel-1 Integration for Automatic Land Cover Classification

3.1. Introduction

Land cover is defined as the biophysical cover of Earth's surface, as artificial surfaces, agricultural areas, forests and semi-natural areas, water and wetlands by the Directive 2007/02 of the European Commission (EC, 2007; Sallustio et al., 2016). Different land cover types furnish various ecosystem services and support habitats; land cover is also considered a climate variable, an element for the description of the environment (Gómez *et al.*, 2016). Human activities depend on land cover and its evolution (Steinhausen *et al.*, 2018). Therefore, the comprehension of the processes that lead to the transformation of the territory are crucial for sustainable resource management, environmental protection, modelling of climate change scenarios, and for the assessment of the impacts on the earth system directly linked to human activities (European Environment Agency., 2020). Remote sensing is important to the study of Earth phenomena, which can be monitored and assessed at different spatial and temporal resolutions (Kuenzer *et al.*, 2014; del Río-Mena *et al.*, 2020).

Several user communities such as decision-makers, non-governmental organizations, European communities, scientists and researchers require different types of land cover information for their activities and policies (Reba and Seto, 2020). For example, land cover information is used for evaluating progress towards UN Sustainable Development Goals (SDGs) targets (UNCCD, 2017), such as target 15.3, which is related to the reach of the Land Degradation Neutrality (LDN) by 2030 (Wunder *et al.*, 2018). Despite the importance of land cover data in environmental monitoring and planning, the products currently available on a national scale are limited in number and their characteristics are not always adequate. The main national land cover and land use products are part of the Copernicus program (https://land.copernicus.eu, accessed on 28 February 2021). The CORINE Land Cover (CLC) guarantees a high thematic detail and a long historical series, but is limited from the point of view of spatial detail and updating frequency (https://land.copernicus.eu/pan-european/corine-land-cover, accessed on 28 February 2021). More recently, the High Resolution Layers (HRLs) have allowed the description with high spatial detail of the main land cover classes, maintaining a multi-year updating frequency (https://land.copernicus.eu/pan-european/high-resolution-layers, accessed on 28 February 2021). There are other Copernicus products on a national scale, with high thematic and spatial detail but related to specific areas, such as Urban Atlas, Riparian Zones, Coastal Zones and Natura 2000 (https://land.copernicus.eu/local, accessed on 28 February 2021). At the national level there is the National Land Consumption Map produced and updated by ISPRA and other land use and land cover maps on a regional scale. These regional products are often derived from CLC and are not up to date, with some exceptions such as in the Lombardy region. In this context, it is crucial to exploit the tools introduced by the Copernicus program, such as Sentinel data and DIAS computing resources. In 2014 the European Space Agency (ESA) launched Sentinel-1 and Sentinel-2 satellites, part of the Space component of the Copernicus program of Earth observation (Berger and Aschbacher, 2012). Sentinel-1 performs C-band synthetic aperture radar (SAR) imaging. It operates both day and night, acquiring images regardless of the weather and supporting the absence of optical images, with a revisit period of six days (Bhattarai *et al.*, 2021; Chatziantoniou *et al.*, 2017; Torres *et al.*, 2012). Sentinel-1 interacts with the structural properties of elements in different ways and through different signal polarizations (VV, VH) in relation to their roughness and moisture content (Muro et al., 2020, Nezry, 2014). The Sentinel-2 constellation (including Sentinel-2A and Sentinel-2B) has a 5-day revisit cycle and provides 13-band multispectral images from visible to Short Wave Infrared (SWIR), associated with three spatial resolutions from 10 to 60 m. Sentinel-2 data characteristics guarantee good results in the monitoring of land use and cover (Drusch *et al.*, 2012; Spadoni *et al.*, 2020; Teodoro and Amaral, 2019). Much research has confirmed that the joint use of information provided by Sentinel-1 and Sentinel-2 lead to overcoming the limitations of using single data products in the land cover (LC) and land cover changes (LCG) classification (Malenovský *et al.*, 2012), allowing the evaluation of different object characteristics, such as pigmentation, moisture, size and density (Erinjery *et al.*, 2018; Ienco *et al.*, 2019; Joshi *et al.*, 2016). The integration of these two data sources can improve land cover detection (Clerici *et al.*, 2017; Mansaray *et al.*, 2017), for example for grassland (Dusseux *et al.*, 2014), urban areas (Iannelli and Gamba, 2018) and land cover changes (Ienco *et al.*, 2019) or hazard assessment (Colson *et al.*, 2018).

Land cover and land use (LU) are interconnected and influence each other and their changes impact on Earth’s natural resources and balances (Duong *et al.*, 2018). The description of territory through its land use and land cover characteristics has led to the creation of many classification systems, conceived of different purposes and referring to various levels of detail and territorial areas. In many cases, a mixed approach has been adopted, which takes into consideration both the characteristics of the land cover and of land use, as in the case of CORINE Land Cover. The problem of comparing, interpreting and using data referring to different classification systems is well known (Anderson, 1976). However, the increase in information sources and classification techniques makes a standardization of classification systems more necessary. To satisfy the requirements proposed by European legislation for monitoring activities, both at the European and national level, a list of criteria for a new data model has been created, and the EIONET Action

Group on Land monitoring in Europe (EAGLE) Concept (Arnold *et al.*, 2013) aims at being a tool for analytic class definitions and for linking recent and future nomenclatures. Moreover, it allows the harmonization of European land monitoring systems and can be implemented as an object-oriented base for mapping and monitoring research. It is not a classification system, but a tool for unifying LC and LU information both in top-down and bottom-up approaches. The EAGLE concept is based on a clear distinction between land cover and land use and is summarized in a matrix composed by three separate descriptors: land cover components (LCC), land use attributes (LUA) and further characteristics (LCH). The descriptors can be combined in order to create specific classification systems for different needs, or to identify correspondences with existing classes, keeping the three descriptors independent[3] .

There are many algorithms for investigating land cover and land use starting from the classification of satellite images. The choice of the most suitable method depends on aspects such as the type of data, distribution of classes, research purposes and interpretability of the classifier, and also the balance between objectives and the available resources. In general, the automatic classification methods are time consuming with the increase of the data dimension and volume, and sometimes data interpretation can become difficult (Chen and Gong, 2013; Gómez *et al.*, 2016). Supervised classifications can be used to analyze a large amount of data; they are based on the selection of an adequate number of training samples with known values (Zhu *et al.*, 2016), which are used to predict unknown values of testing data (Holloway *et al.*, 2019). For this reason, it is important to select examples which completely represent the variability of the characteristics of the analyzed territory. In the past few years the supervised classification technique increased in popularity (Thanh Noi and Kappas, 2017), thanks to the availability of ancillary data, which provide accurate information on ground truth (Colditz, 2015). Ground truth information is also important to define thresholds values for the classification; these values are important for isolating objects from the background, in order to improve classification precision (Liu *et al.*, 2011). Threshold methods are used for their simplicity and because they require few parameters (Huang *et al.*, 2019; Xin *et al.*, 2020). The identification and use of thresholds related to the spectral characteristics and their variation over time are useful to assess both physical and biophysical characteristics of land cover (Kobayashi *et al.*, 2020; Otukei and Blaschke, 2010), to classify cover change (Sinha and Kumar, 2013), to define crop phenology for determining start and end of growing season (Huang *et al.*, 2019), and to classify tree species (Agrillo *et al.*, 2021; Persson *et al.*, 2018) or burned areas (Ngadze *et al.*, 2020). Some supervised classification methods, for example parametric supervised

[3]https://land.copernicus.eu/eagle/content-documentation-of-the-eagle-concept/manual/eagle-explanatory-documentation-v3-1-version-2021

classifiers (maximum likelihood, minimum distance), are difficult to use for classifying large multi-temporal datasets because they have limited flexibility in decision boundaries. Since these classifiers assume a normal data distribution, they furnish excellent results when data distribution is unimodal but they can be difficult to apply for multi-modal dataset analysis (Belgiu and Drăgu, 2016). Machine learning (Lary *et al.*, 2016) approaches are largely used in land cover mapping, thanks to their capacity to model class signatures, to elaborate many input data and to produce higher accuracy compared to traditional parametric classifiers, especially for complex data with many predictor variables (Maxwell *et al.*, 2017). For example, decision trees and random forests are machine learning approaches which focus on decision rules on class boundaries which improve the classification of land cover types with unknown distribution and frequency (Foody and Mathur, 2006; Kulkarni and Lowe, 2016; Liu *et al.*, 2018). On the other hand, machine learning is characterized by higher computational intensity, unknown decision rules (black box) and the need of input parameters (Gómez *et al.*, 2016; Rodriguez-Galiano *et al.*, 2012); moreover, it is not easy to apply to large areas, because it is characterized by difficult computing, such as the manual choice of a large number of training areas (Sun *et al.*, 2019) and it is negatively influenced by noisy observations and over-fitting (Gómez *et al.*, 2016).

This research aims to develop a methodology that can provide land cover and land cover change products for operational purposes which can achieve the following project requirements:

- allow the production of a 10 m land cover map on a national scale, exploiting the potential offered by free Sentinel-1 and Sentinel-2 data;
- allow at least annual updates for the entire Italian territory;
- ensure accuracy in line with Copernicus products;
- adopt a structure that allows the updating and expansion of the methodology;
- ensure compatibility with the main European and national initiatives in the field of land monitoring, such as Copernicus program and the Mirror Copernicus initiative of the Italian Ministry of Economic Development.

The main challenge of this research regards the creation of a tool capable of meeting all the indicated requirements, and this is also demonstrated by the absence of national and European products with these characteristics. In this sense, proven classification techniques were used, obtaining a methodology based on training areas and decision rules, (Berhane et al., 2018, Houhoulis and Michener, 2000) to produce pixel-based land cover classification products (Phiri *et al.*, 2019).

The study is in continuity with other previous analysis conducted by this working group on land consumption (Luti *et al.*, 2021; Strollo *et al.*, 2020) and forest areas (Giannetti *et al.*, 2020). For the identification of the land cover classes mapped by the algorithm, an EAGLE-compliant classification system has been developed. In detail, eight land cover classes have been identified and characterized:

- unvegetated abiotic surfaces (subdivided into artificial surfaces and natural surfaces);
- vegetated biotic surfaces (distinguishing between woody vegetation, subdivided into broad-leaved and conifers, and permanent and periodically herbaceous vegetation);
- water surfaces (water bodies and permanent snow and ice).

Change detection classes regarded variations in artificial surfaces (land consumption and restoration) and forest disturbances, distinct into fires and other types of disturbances. To test the methodology, this was applied to the Sentinel-1 and Sentinel-2 data series relating to 2018 for the classification of the land cover, and to the period 2017–2018 for the detection of land cover changes.

The methodology allowed the identification of land cover classes by exploiting only Sentinel-1 and Sentinel-2 data. The only ancillary data used is the National Land Consumption Map of ISPRA for the distinction of artificial abiotic surfaces.

3.2. Materials and Methods

3.2.1. Overview

The methodology presented in this study uses Sentinel-1 GRD and Sentinel-2 images to map land cover and land cover changes in Italy, on a yearly basis (Figure 8).

Decision rules are defined to identify land cover classes at pixel size of 10 m, setting fixed threshold values on composites (e.g., median, maximum) of multitemporal data. The rules are based on experience, on deep knowledge of literature and on previous analysis of temporal trend of spectral signatures, indices and backscatter coefficients of each class. In this sense, the study is in continuity with other analysis published in previous scientific papers related with land consumption (Luti *et al.*, 2021; Strollo *et al.*, 2020) and forest area monitoring (Giannetti *et al.*, 2020; Spadoni *et al.*, 2020). Furthermore, the thresholds for vegetation classes were also defined using a collection of training areas. The output classifications were validated through the photointerpretation of very high-resolution images.

The methodology was developed with Google Earth Engine (Javascript language program) but was also adapted for QGIS Semi-Automatic Classification Plugin and on a DIAS environment[4]. This DIAS experimentation refers to the agreement between the Italian Space Agency (ASI) and ISPRA for habitat mapping.

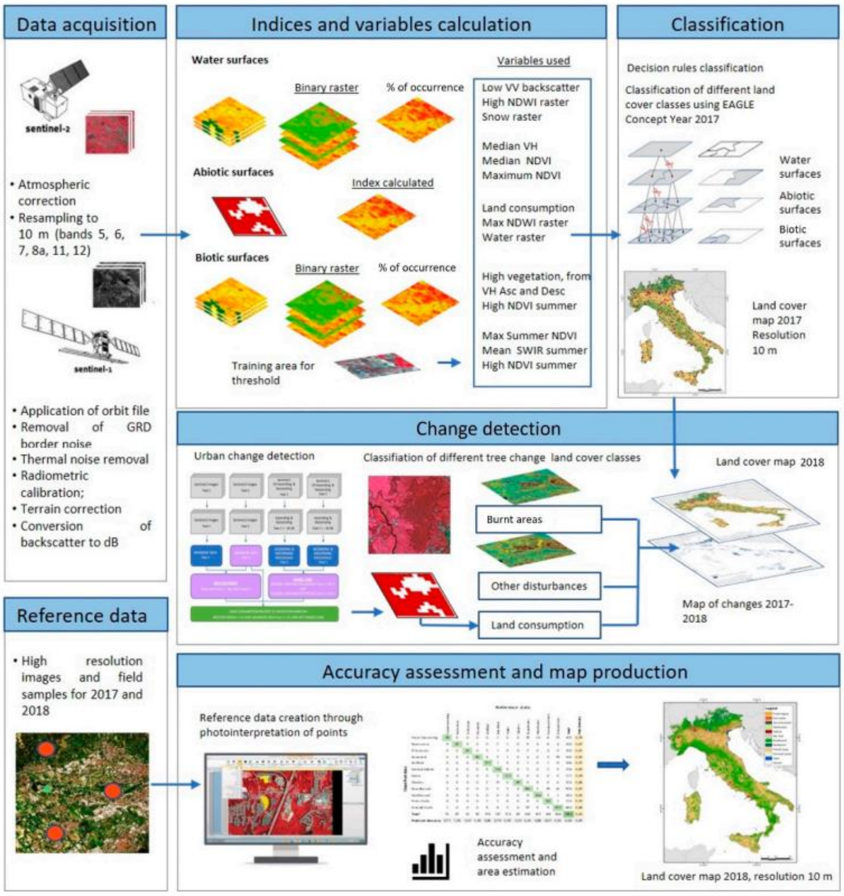

Figure 8 Workflow of the methodology.

3.2.2. Study Area

The study area is the Italian territory (Figure 9), characterized by a total surface of 301,338 km^2. Geographically Italy is divided into three parts: The Alpine arc in the North, the peninsula in the South central part and the insular area that includes Sardinia and Sicily. The climate in the Italian territory is alpine in the North, continental in the peninsula and Mediterranean along the coast, with drier summers and milder winters. The average annual rainfall is higher in the mountain areas with a maximum of 2500–3500 mm and reaches a minimum of 500 mm in Apulia and Molise.

The territory is mainly composed of mountains (35.0%) or hills (41.6%) and plains (23.2%). The mountain belts are the northern area, forming the Alps with peaks reaching over 4000 m in the

[4] ONDA https://www.onda-dias.eu, accessed on 4 March 2021

western part. The Appennines chain extends over the whole peninsula, reaching its highest peak with Gran Sasso (2912 m), located in Abruzzo. Due to these characteristics, the northern sector and the central chain are often covered by clouds or affected by perturbations that make it difficult to find cloud-free optical images for various periods of the year.

Land cover is characterized by forest, grassland, built-up areas and water (both liquid water and ice). Herbaceous vegetation is dominated by cropland and pastures and is mainly concentrated on coastal areas and, in Padana plains. The orography conditions the distribution of the settlements, which are concentrated in the plains and on the coastal areas.

Figure 9 Study area.

3.2.3. Classification System

The classification system defined for this research is consistent with the EAGLE group specifications[5] (Kleeschulte et al., 2017) and it is composed of eight land cover classes, organized into three levels (

Table 2), and four land cover change classes (Table 3).

[5]https://land.copernicus.eu/eagle/content-documentation-of-the-eagle-concept/manual/eagle-explanatory-documentation-v3-1-version-2021

Table 2 Land cover classification system.

Land Cover		
I Level	**II Level**	**III Level**
Abiotic-non vegetated	Artificial abiotic surfaces	
	Natural abiotic surfaces	
Biotic vegetated	Woody vegetation	Broad-leaved
		Conifers
	Herbaceous vegetation	Permanent
		Periodically
Water surfaces	Water bodies	
	Permanent snow and ice	

Table 3 Land cover change classification system.

Land cover change	
Land consumption	
Restoration	
Forest disturbaces	Burnt areas
	Other disturbances

The first classification level distinguishes three main land cover classes. It consists of land cover classes at the first two levels, while the "biotic-vegetated" class is then further articulated through a third classification level based on biophysical characteristics (broadleaves, conifers) and temporal parameters (permanent herbaceous, temporary herbaceous).

1. Abiotic-non vegetated: The class includes any unvegetated surfaces, either covered with man-made artificial structures or geologically natural material surfaces (with or without anthropogenic influence or impact). It includes consolidated or unconsolidated unvegetated natural surfaces, such as rocky regions and sands.

At the second classification level, the class is subdivided into artificial and natural surfaces.

- Artificial abiotic surfaces include impervious surfaces but also those of reversible land consumption (Munafò, 2021). Impervious and sealed surfaces are mainly covered with features with a certain height above ground (artificial buildings and constructions) or features without a specific height above ground (flat waterproof surfaces) such as artificial surface pavements (e.g., asphalt, concrete, tarmacadam). Reversible land consumption includes all areas where natural surface material has been replaced by man-made material or where natural material has been removed, forming a non-impermeable and undeveloped surface. This includes soil compaction, excavation or temporary impervious cover. For example: unpaved roads, construction sites or courtyards or sports fields, permanent deposits of material, photovoltaic fields, and quarries not yet restored. The EAGLE classification system includes quarries and extraction sites in the 1.2.1.1 bare rock class, but in this research they are classified as artificial abiotic surfaces since quarries and extraction sites are strongly modified due to compaction and morphology changes caused by material removal. The physical characteristics of these areas are more similar to those of the EAGLE 1.1.2 "Non- Sealed Artificial Surfaces" class, which includes any open areas where natural surface material has been replaced by artificial material or where natural material has been removed from its place of origin as result of human activity forming an unsealed and non-built-up surface.

- Natural abiotic surfaces are any kind of surface material that remains in its natural consistence or form, either with or without anthropogenic influence. Consolidated and unconsolidated surfaces include: unvegetated rocky mountainous regions, sand, and bare soil.

2. Biotic-vegetated: The class includes any terrestrial surface with spontaneous, semi- natural or artificial vegetation (e.g., crops and urban parks), with or without anthropogenic influence.

At the second classification level, biotic-vegetated surfaces are divided into woody vegetation and herbaceous vegetation.

- Woody vegetation includes perennial woody plants with a single, self-supporting main stem or trunk, containing woody tissue and branching into smaller branches and shoots.

A third classification level has been defined applying to the land cover class "trees" a characteristic related to the "leaf form", allowing for distinguishing conifers and broad-leaved trees.

- Herbaceous vegetation includes annual, biennial or perennial plants such as most ferns and grasses that do not have a persistent woody stem above ground. On the opposite to woody plants, the buds of herbaceous plants die at the end of the growing season, to regenerate from the tissues left above or below the ground (e.g., bulbs, rhizomes, tubers, seeds).

A third classification level has been defined applying to the land cover class “herbaceous vegetation”, a characteristic related to the “temporal parameter” of “duration”. In this way the class is divided on the basis of the persistence of the coverage throughout the year.

Permanent herbaceous areas are characterized by a continuous vegetation cover throughout a year. No bare soil occurs within a year. These areas are either unmanaged or extensively managed natural grasslands or permanently managed grasslands, or arable areas with a permanent vegetation cover or even set-aside land in agriculture.

Periodically herbaceous areas are characterized by at least one land cover change between bare soil and herbaceous vegetation within one year. Depending on the management intensities these areas can change between these two land cover components within a year. Normally these areas are arable areas.

3. Water surfaces class includes water in its liquid or solid state of aggregation (snow and ice), both of natural formation (water basins, rivers, streams, stagnant waters, glaciers) and of artificial origin.
    - Water bodies are water in a liquid state of aggregation regardless of location, shape, salinity and origin (natural or artificial). This includes all types of inland water surfaces without direct interference or interchange with open sea water, regardless of salinity and origin (natural or artificial). Watercourses (flowing water surfaces) come from: rivers, streams, canals, waterways. Stagnant waters (water surfaces of non-current waters, mainly lakes and ponds, or meanders of cut rivers) come from: natural lakes (both fresh and salt water), ponds, reservoirs, oxbow lakes, bottom pools, irrigation ponds, ponds for the production of artificial snow, river banks for the production of hydroelectric energy, and ponds for firefighting.
    - Permanent snow and ice class includes the snow cover that persists all year round, above the climatic snow limit and the persistent ice cover formed by the accumulation of snow.

The possible land cover change classes can be identified through a transformation matrix obtained starting from the considered land cover classes (Table 4). This paper focused on the changes that can occur with a good probability within a year, between 2017 and 2018 (code 2 in Table 4). These changes refer primarily to variations of artificial abiotic surfaces in terms of expansion (i.e., land consumption), reduction (i.e., restoration), and the loss of woody vegetation associated with forest disturbances (transformation from a woody vegetation class to natural abiotic surfaces or herbaceous vegetation). Code 1 of Table 4 indicates changes that can be identified with good approximation by exploiting a series of multi-year images. The changes associated with these land cover classes have not been analyzed in this paper, as the considered period is not wide enough to allow monitoring of transformations such as increases in woody vegetation increase or to distinguish seasonal variations from real changes in herbaceous classes. Finally, code 0 indicates transitions that are not within the objectives of the project.

A definition of the four land cover change classes analyzed in this paper is provided below.

1. Land consumption is the replacement of a non-artificial land cover to an artificial abiotic surface, both permanent and reversible. Artificial abiotic surfaces that have been changed by, or are under the influence of human activities resulting in a land consumption process can be sealed or unsealed.
2. Restoration is the replacement of an artificial abiotic surface with a semi-natural land cover where permeable land is back, such as herbaceous (permanent and periodic), woody, or urban green.
3. Forest disturbances are drastic decreases or disappearances of vegetation in areas classified as woody vegetation, not associated with land consumption. Forest disturbances are further distinguished into two classes, through the introduction of two elements relating to the status segment of the landscape characteristics of the EAGLE matrix.
    - Burnt areas: The class includes natural woody vegetation affected by recent fires. This class includes recently burnt areas of forests, moors and heathlands, sclerophyllous vegetation, transitional forest-shrub formations, and areas with sparse vegetation.
    - Other disturbances: The class “other disturbances” identifies the removal of all or most of the trees, large and small, on a surface, following a disturbance event other than a fire. Virtually all woody vegetation is removed from the site in preparation for the settlement of new wood or another class of biotic vegetation.

Table 4 Land cover transformation matrix on yearly basis. The following transformations are considered: 3—Likely (Stability), 2—Probable (change can happen during one year of observations), 1—Possible (change can happen in specific cases, but they generally need a longer time of observation to be assessed), 0—Not considered.

		2018							
		Artificial abiotic surfaces	Natural abiotic surfaces	Broad-leaved	Conifers	Permanent-herbaceous	Periodic herbaceous	Water bodies	Permanent snow and ice
2017	Artificial abiotic surfaces	3	2	2	2	2	2	0	0
	Natural abiotic surfaces	2	3	1	1	1	1	0	0
	Broad-leaved	2	2	3	0	2	2	0	0
	Conifers	2	2	0	3	2	2	0	0
	Permanent herbaceous	2	1	1	1	3	1	0	0
	Periodic herbaceous	2	1	1	1	1	3	0	0
	Water bodies	0	0	0	0	0	0	3	0
	Permanent snow and ice	0	0	0	0	0	0	0	3

3.2.4. Pre-Processing of Images and Collection of Training Areas

Before the classification process, preliminary steps are required to collect training areas and preprocess the input images, in order to obtain the backscatter values for Sentinel-1 GRDs and reflectance values for Sentinel-2 images (Figure 8).

Sentinel-2 L1C images were used in this study. The atmospheric disturbance affecting L1C pixels can increase the uncertainty of spectral signatures, therefore decreasing classification accuracy; however, Sentinel-2 L2A historical archive was not available at the time of research. The input

images were cloud masked and pixels affected by clouds excluded from calculations. The cloud mask was made through a simple algorithm based on the quality assessment band as described in Google Earth Engine website[6]. It is worth highlighting that cloud mask products provided with Sentinel-2 images generally underestimate clouds (Coluzzi *et al.*, 2018).
Sentinel-1 GRD images provided in Google Earth Engine are already converted to backscatter values, following the steps described in the website[7], which are:

- application of orbit file;
- removal of GRD border noise (low intensity and invalid data);
- thermal noise removal to reduce discontinuities between sub-swaths;
- backscatter intensity calculation using radiometric calibration;
- terrain correction (orthorectification using the SRTM 30 m DEM);
- conversion of backscatter coefficient to dB; due to the wide dynamic range of SAR data, decibels were used to amplify the small variations and make the data more readable (Flores *et al.*, 2019).

The same preprocessing can be applied in the DIAS environment.
Sentinel-1 GRD and Sentinel-2 grids are not aligned and have different spatial resolution: in detail, the Sentinel-1 GRD have ground range geometry, and the pre-processing involves the geocoding to map coordinates. Sentinel-2 images are provided in WGS 84 UTM coordinates and cover the Italian territory with about 80 granules. To use both images in the same workflow, the spatial resampling to a common coordinate reference system and spatial resolution is required. In this study, Sentinel-1 data were resampled using nearest neighbor and aligned to the Sentinel-2 grid, in the WGS 84 UTM reference system.
Another preliminary operation is the photointerpretation of 905 training areas, selected to represent the diversity of land cover classes and distributed along the entire Italy. These are used for setting the thresholds in classifying woody vegetation and for the distinction between broad-leaved and conifers vegetation.

3.2.5. Classification Process

The production of the land cover map involves the calculation of multitemporal indices and decision rules that allows for the separation of land cover classes. The classification is pixel based: a pixel is assigned to a specific class if it has spectral and geometric characteristics that satisfy all

[6] https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2, accessed on 25 March 2021
[7] https://developers.google. com/earth-engine/sentinel1, accessed on 25 March 2021

the rules and thresholds defined for that class. The classification primarily regards the assignment to each pixel of one of the three classes at the first classification level (

Table 2). In detail, the water surfaces pixels were first identified. In the areas not classified as water surfaces, the presence of abiotic-non vegetated areas was evaluated, and finally the remaining pixels were attributed to the class of biotic-vegetated surfaces. This operation provides three masks, which are then further classified with reference to the second or third classification level, according to the methodology described below. For further details, please refer to appendix A. Table 5 shows the spectral indices used for classification and calculations on the base of the Sentinel-2 images.

Table 5 Spectral indices used for land cover classification and change detection.

Satellite	***Index Name***	***Brief Index Formula***	***Reference***
Sentinel-2	Normalized Difference Vegetation Index	$NDVI = \frac{(B8 - B4)}{(B8 + B4)}$	(Rouse *et al.*, 1973)
	Normalized Burn Ratio	$NBR = \frac{(B8 - B12)}{(B8 + B12)}$	(Key *et al.*, 1999)
	Normalized Difference Water Index	$NDWI = \frac{(B2 - B8)}{(B2 + B8)}$	(McFeeters, 1996)
	Normalized Difference Snow Index	$NDSI = \frac{(B3 - B11)}{(B3 + B11)}$	(Dozier, 1989)
	Normalized Difference Coniferous Index	$NDCI = \frac{(B6 - B12)}{(B6 + B12)} \frac{(B8 - B11)}{(B8 + B11)}$	
	Burned Index	$BI = \frac{1 - (B3 + B4 + B8)}{1 + (B3 + B4 + B8)}$	

3.2.5.1. Water Surfaces

Water classes are required input for the classification of abiotic land cover classes; therefore, they must be classified before the others (Figure 21).

Permanent Snow and Ice

The classification of permanent snow and ice is based on the definition of rules starting from the NDVI index and the "snow raster". The "snow raster" derives from the methodology developed by ESA[8] to detect snow surfaces, which are identified by pixels that have band reflectance lower than threshold values reported below:

$$NDSI \geq 0.2$$

[8] https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/level-2a/algorithm

And

$$Band\ 8\ \geq 0.15$$

And

$$Band\ 2\ \geq 0.28$$

And

$$\frac{Band\ 2}{Band\ 4} \geq 0.85$$

The algorithm is based on the NDSI index, which exploits the property of snow surfaces to be very bright in the visible and dark in shortwave infrared bands (cloud and snow have similar reflectance in band 3 but band 11 reflectance for clouds is higher than snow). This index allows for the detection of snow and differentiating it from clouds, with values varying from -1 to +1. As a general rule, positive values define areas probably with a snow cover. The NDSI is flanked by three "filters" relating to band 8 (NIR), band 2 (blue) and the ratio between band 2 and 4. The thresholds are those indicated by ESA analysis. The snow raster represents the percentage of times that all the aforementioned conditions are satisfied during a specific reporting period.

The following three conditions are defined for the identification of permanent snow and ice.

1. Snow raster ≥ 10%

The first condition to indicate the presence of permanent snow and ice in the pixel refers to the Sentinel-2 for the whole year. It defines that snow raster conditions must be verified in the pixel for at least 10% of valid observations. The threshold of 10% allows the elimination of outlier pixels.

2. Snow raster summer > 1%

The second condition on the snow raster refers to Sentinel-2 images for the summer period (1 July–31 September) and a less conservative threshold of 1% has been adopted for the elimination of outliers.

3. Maximum NDVI < 0.4

For each pixel, the maximum NDVI value is calculated with respect to the annual series of valid Sentinel-2 acquisitions. A threshold of 0.4 was defined to exclude areas with the presence of vegetation.

Water Bodies

Five conditions were considered in order to identify water bodies; each of them contributes to reduce possible mistakes between the spectral signature of water and other classes. Three rules are based on indices derived from Sentinel-2 images and two used Sentinel-1 GRD data sets.

1. High NDWI raster ≥ 5%

It was observed by McFeeters et al., (1996) that water class has NDWI > 0 while non-water classes have NDWI < 0; a threshold value of 0.3 of NDWI was considered to better isolate water from the built up area and other surfaces (McFeeters, 2013). The first condition to indicate the presence of water bodies is based on NDWI values referring to Sentinel-2 for the whole year. It defines that NDWI ≥ 0.3 conditions must be verified in the pixel for at least 5% of valid observations. The threshold of 5% allows the elimination of outlier pixels.

2. Median NDVI < 0.3

From the NDVI multi-temporal series for the whole year, the median of the NDVI for each pixel was calculated. A threshold of 0.3 was then defined in order to eliminate areas with vegetation.

3. Snow raster < 20%

As for the class of perennial ice and snow, a "snow raster" was elaborated, relating to the annual Sentinel-2 series. The potential presence of water bodies was referred to pixels where snow raster conditions are verified in less than 20% of the annual valid observations.

4. Median VH polarization < -20 dB

The backscatter values used in this classification are derived from tests carried out on sample areas and from values reported in the literature (Gulácsi and Kovács, 2020; Mansaray *et al.*, 2017; Zhou *et al.*, 2017). The VH band discriminates quite well between the different classes presenting backscatter values around -20 dB, which remains constant throughout the year. Variations can be due to the wind, which makes the water surface rougher and modifies its backscatter. The potential presence of water bodies was referred to pixels with an annual median of the VH polarization < -20 dB.

5. Low VV backscatter <-15%

The VV values are generally higher than VH ones. In the case of VV backscatter the value -10 dB was determined as a good threshold for identifying water (Mansaray *et al.*, 2017; Zhou *et al.*, 2017).

A binary raster was produced starting from the monthly median (i.e., divided 12), attributing value 0 for values below -10 dB and 1 elsewhere. It was considered water class when the percentage of 1 throughout the year is less than 15%.

3.2.5.2. Abiotic Classes

The workflow for abiotic classes is based on the definition of two rules and requires only Sentinel-2 input images for the whole year. The abiotic surfaces are then distinguished between artificial and natural surfaces (Figure 22).

1. Maximum NDVI ≤ 0.35

A threshold of 0.35 is defined on the maximum annual NDVI value, in order to exclude areas that may present vegetation during the year. This calculation is in common with the processing of water classes; therefore, the same data were calculated only once.

2. Not Water

Both the abiotic surface class and the water surfaces have low maximum NDVI values. To avoid commission errors, this second rule is defined, which excludes pixels previously classified as water surfaces.

Artificial Abiotic Surfaces

The distinction of natural and artificial abiotic classes is based on the National Land Consumption Map elaborated by ISPRA. The map was produced through supervised classification of RapidEye images using maximum likelihood algorithms (Luti *et al.*, 2020) and the methodology was then revised to use of Sentinel data. The map is updated every year through photointerpretation, with the help of masks of potential changes. The masks allow to reduce the area inspected by the photo interpreters, limiting the time and costs of the updating activity (Munafò, 2021). This paper describes the results of the updating methodology, based on the use of Sentinel-1 and Sentinel-2 data; for details on the methodology for the production of the map, see Strollo et al., (2020).

Natural Abiotic Surfaces

Artificial surfaces and natural abiotic surfaces are very similar from a spectral point of view, so it is difficult to distinguish them only using satellite images. The natural abiotic surfaces class has been identified considering the pixels classified at the first level as abiotic surfaces that have not already been classified as artificial abiotic surfaces.

3.2.5.3. Biotic Classes

Biotic vegetation pixels are those that are not included in the classes of water and abiotic surfaces.

Woody vegetation was first identified and then divided into conifers and broadleaves trees. The land cover of the remaining areas was divided into periodically or permanent herbaceous, according to the EAGLE definition (Kleeschulte *et al.*, 2017).

The classification of biotic surfaces required the Sentinel-1 GRD dataset for the whole year and Sentinel-2 images for the whole year, for the summer period (1 June–31 August) and for the winter period (1 January–30 April and 1–31 December). Furthermore, 905 training areas were collected throughout the national territory, for the identification of the woody vegetation and for the distinction of conifers and broad-leaved trees.

Woody Vegetation

Woody vegetation was identified through the definition of three rules. The first two rules are related to Sentinel-2 optical images, while the third is based on Sentinel-1 GRD data (Figure 23).

1. High NDVI summer ≥ 70%

The first condition to indicate the presence of woody vegetation is based on NDVI values referring to Sentinel-2 for the summer period (between 1 June to 31 August). It defines that NDVI≥0.5 conditions must be verified in the pixel for at least 70% of the valid observations.

2. Maximum NDVI summer ≥ Threshold on minimum value of the Maximum NDVI summer

The maximum NDVI value for the summer period (between 1 June to 31 August) was calculated for each pixel. This value was compared with a threshold obtained from the training area. In detail, the maximum summer NDVI was calculated in each pixel of the training areas and the lowest of them was taken as the reference threshold. Pixels with a maximum NDVI above the training area threshold are associated with the presence of woody vegetation.

3. High vegetation

The backscatter of VH polarization for the whole year was used, as it can partially improve the detection between high vegetation (trees) and low vegetation (e.g., crops) (Holtgrave *et al.*, 2020). Since the acquisition orbit influences the angle of view, the ascending and descending orbits are considered separately. A conservative threshold of -20 dB has been defined for distinguishing low vegetation and fallow land (that generally have backscatter values lower than -11 Db (Formaggio *et al.*, 2001), from trees (that generally have higher backscatter values). Pixels with backscatter values > -20 dB in at least 2% of valid observations, both for VH ascending and VH descending, were considered as woody vegetation. The 2% threshold is used to exclude possible outliers from time series of backscatter values.

Broadleaves and conifers

Pixels classified as woody vegetation can be further distinguished in broadleaves and conifers using Sentinel-2 summer and winter images and training areas.

1. Woody vegetation raster = 1

The conditions for distinguishing broad-leaved and conifers are applied only to pixels classified as woody vegetation.

2. Mean SWIR summer ≤ Threshold Mean SWIR summer (conifers) Mean SWIR summer > Threshold Mean SWIR summer (broadleaves)

The spectral characteristics of conifers and broad-leaved are very different in their wavelengths of short wave infrared (SWIR, band 11 of Sentinel-2 images). Starting from the training area of the conifers pixels, the mean of the SWIR was calculated on the summer period (1 June–31 August) and the maximum value was set as a threshold. Conifers pixels will have mean SWIR values for summer period lower than or equal to the threshold, while broad-leaved pixels will have higher.

3. High NDCI winter > Threshold high NDCI winter (conifers)

The third condition exploits a new empirical index developed in this research and called the "Normalized Difference Coniferous Index" (NDCI). The index is calculated on the winter subset of Sentinel-2 images (1 January–30 April and 1–31 December) and is the product of two normalized difference indexes: the first is based on band 6 (red edge, 740 nm) and band 12 (short wave infrared band, 2190 nm), and the second is based on band 8 (infrared, 842 nm) and band 11 (short wave infrared band, 1610 nm).

$$\frac{(B6 - B12)}{(B6 + B12)} \frac{(B8 - B11)}{(B8 + B11)}$$

These bands show a more pronounced difference between conifers and broadleaves areas in winter than in summer. This is probably due to the drastic fall in the spectral response of broadleaves trees compared to conifers which do not drop their leaves in a senescence period (Persson *et al.*, 2018). Starting from the conifers training areas, the NDCI was calculated and a threshold of 0.3 was set. The percentage of times that NDCI > 0.3 in the winter period was then calculated for each pixel of the conifers training area. The lowest of these percentages was set as a reference threshold for the distinction of conifers and broadleaves trees. Conifers pixels will have frequencies of exceeding the threshold higher than or equal to the training area one, while broadleaves will have lower.

Herbaceous Vegetation

Herbaceous vegetation relates to pixels not previously assigned to other land cover classes (Figure 24). These areas meet the following conditions:

1. Maximum NDVI > 0.35

As seen for the identification of abiotic surfaces, the condition distinguishes the vegetated areas from the unvegetated ones.

2. Not woody vegetation

Areas with vegetation cover but which do not respect the rules for falling into woody vegetation are considered herbaceous.

Permanent and Periodically Herbaceous

The pixels of herbaceous vegetation can be divided into permanent and periodically herbaceous, by defining conditions on the NDVI values calculated on the Sentinel-2 annual series:

1. Low NDVI ≤ 5% (permanent herbaceous)
2. Low NDVI > 5% (periodically herbaceous)

Starting from the Sentinel-2 annual image series, the frequency with which NDVI < 0.35 (low NDVI) is calculated for each pixel. Pixels with NDVI values < 0.35 in less than 5% of valid observations have been classified as “permanent herbaceous”, since this condition implies a very constant presence of vegetation. Otherwise, the pixel is classified as “periodically herbaceous”, since the vegetated cover was replaced by non-vegetated cover for several periods of the year. It is worth noticing that non-vegetated cover could also be snow cover.

3.2.5.4. Land Cover Changes

The algorithm allows the user to directly identify changes, through specific procedures, for the different change classes. This approach limits the propagation of errors that would occur if changes were sought by comparing two complete land cover classifications (Lefebvre *et al.*, 2016). The identification of changes of artificial abiotic surfaces involved the creation of masks of potential changes, used for the annual update of the National Land Consumption Map. In detail, Sentinel-1 and Sentinel-2 images for 2017 and 2018 were used. For Sentinel-1 data, thresholds on ascending and descending backscatter variations were defined, while for Sentinel-2 data thresholds regard the maximum NDVI value in the second year and the NDVI difference between the two years. Areas with high backscatter variations and/or strong reductions in NDVI between the two years were included in a raster of potential changes. The raster was used as a basis for the photointerpretation activities, carried out for the updating of the National Land Consumption Map. The details of this study are described in another paper (Luti *et al.*, 2021).

A specific methodology for mapping forest disturbance was developed, which also allows distinguishing burned areas from other kinds of changes such as forest harvesting (Figure 25). The methodology is applied to areas classified as woody vegetation in the land cover map. This class acts as a mask for the delimitation of the forest area of year 1, within which the search for changes takes place. The input data used are the woody vegetation class for year 1 and the Sentinel-2 summer images (from 1 June to 31 August) for the two reference years. In this research, the changes that took place between 2017 and 2018 were analyzed. Two spectral indices were calculated for the two reference years: NDVI and NBR (Chirici *et al.*, 2020) (Table 5).
NBR is useful for identifying areas that have been burned in the reference period. Notably, healthy vegetation has very high near-infrared reflectance and low reflectance in the shortwave infrared portion of the spectrum. Recently burned areas have a low reflectance in near infrared and a high reflectance in the short-wave infrared band.
NDVI is high in vegetated areas, while it decreases if there are reductions in vegetation cover.

To identify a forest disturbance, at least one of the following two conditions must be verified.

1. NDVI Difference $\leq$ -0.2

The median of the summer NDVI is calculated for the two reference years and then the difference between the two values is considered:

$$NDVI\ difference = Median\ NDVI\ (year\ 1) - Median\ NDVI\ (year\ 2)$$

If the difference is greater than 0.2, the change in NDVI indicates the presence of a change.

2. NBR Difference ≤ -0.2

The difference between the median of the summer NBR of the two reference years is considered:

$$NBR\ difference = Median\ NBR\ (year\ 1) - Median\ NBR\ (year\ 2)$$

The threshold value 0.2 was identified empirically as optimal value for detecting tree loss and avoiding errors caused by phenological variations between years.

Distinction between “Burnt Areas” and “Other Distubances”

Starting from the pixels classified as forest disturbance, a distinction was made between “burned trees” and “other forest changes”, through an innovative burned index (BI) (Table 5). The BI is higher for dark or brown pixels, as is in case of burned areas.
The median of BI of year 2 was calculated, obtaining a raster of “Median Burned Index Year 2”. Pixels already classified as forest disturbance can be further detailed in fires if Median BI Year 2 ≥0.45.
The threshold value 0.45 was determined empirically as the optimal value for distinguishing burned areas and avoiding errors due to shadow areas where pixels tend to be dark.

3.2.6. Accuracy Assessment

The validation involved the 2018 land cover map with an eight-class classification system, integrated with the four land cover change classes.
First, a qualitative verification was carried out through a systematic visual search for macroscopic errors. An accuracy assessment on the base of Olofsson methodology was then performed (Olofsson *et al.*, 2014); this approach is widely used in the literature (FAO, 2016; Olofsson *et al.*, 2014; Stehman and Czaplewski, 1998; Stehman and Foody, 2019).
The validation was carried out by identifying, ad hoc, a random sample of points which was then photointerpreted using very fine resolution images. The points were compared with the values of the land cover map at the same locations. The sample size (n) was calculated using the equation reported by Olofsson et al., (2014) and referred to 5.25 in Cochran, (1977) .

$$\frac{(\sum WiSi)^2}{[S(\hat{O})]^2}$$

where:

Wi—is area proportion of each class derived from the map classification

Si—standard deviation of stratum i, Si =√(Ui(1 Ui)) (Cochran, 1977, Eq. 5.55)

Ui—User accuracy of class i

S(Ô)—is the standard target standard error

Si influences the sample size, and its value is related to the Ui, which is unknown. A conservative scenario was assumed, considering Ui equal to 0.6 for all twelve classes.

The target standard error for overall accuracy was assumed to be 0.01 as suggested by Olofsson et al., (2014), which corresponds to a confidence interval of 1%.

The final sample size is 2400 (Table 6).

Table 6 Calculation of sample size.

Land Cover Class	***Area (ha)***	***Wi***	***Ui***	***Si***	***Wi*Si***
Artificial Abiotic	2,127,999	0.071	0.6	0.490	*0.035*
Natural Abiotic	953,574	0.032	0.6	0.490	*0.015*
Broad-leaved	12,006,500	0.398	0.6	0.490	*0.195*
Conifers	1,607,582	0.053	0.6	0.490	*0.026*
Permanent herb.	12,657,900	0.420	0.6	0.490	*0.206*
Periodic herb.	239,834	0.008	0.6	0.490	*0.004*
Water bodies	337,737	0.011	0.6	0.490	*0.005*
Permanent snow and ice	111,773	0.004	0.6	0.490	*0.002*
Other disturbance	78,49	0.003	0.6	0.490	*0.001*
Burned areas	11,106	0.000	0.6	0.490	*0.000*
Land consumption	6600	0.000	0.6	0.490	*0.000*
Restoration	906	0.000	0.6	0.490	*0.000*
Total	30,139,974				*1*
S(Ô) overall accuracy		**0.01**			
Total number of samples		***2400***			

Sample size calculation according to Olofsson et al., (2014). The required inputs are the areas of each and the expected user's accuracy for the sample. The points were divided among the land cover classes through a stratified sampling.

The number of points attributed to each class was assumed to be equal to the average between equal and area-proportional distribution (seehttps://fromgistors.blogspot.com/2019/09/Accuracy-Assessment-of- Land-Cover-Classification.html accessed on 4 April 2021) (Table 7).

Table 7 Allocation of the sample between the considered classes.

Classes	***Allocation***		
	Equal	Proportional	***Final***
Artificial Abiotic	200	169	*185*
Natural Abiotic	200	76	*138*
Broad-leaved	200	956	*578*
Conifers	200	128	*164*
Permanent herb.	200	1008	*604*
Periodic herb.	200	19	*110*
Water bodies	200	27	*113*
Permanent snow and ice	200	9	*104*
Other disturbance	200	6	*103*
Burned areas	200	1	*100*
Land consumption	200	1	*100*
Restoration	200	0	*100*
Total	*2400*	*2400*	*2400*

The points were photointerpreted with very fine resolution images from Google Earth, considering 2018 as the reference year for the land cover classes and 2017 and 2018 for the land cover change classes. The photointerpretation followed a pixel based approach, assigning to each point the class on which it falls exactly.

3.3. Results

The land cover map was obtained from the application of the methodology (Figure 10). The map shows the eight land cover classes for 2018 and the four land cover change classes for 2017–2018 described above.

Accuracy assessment was done on the map of Figure 10 and provides the results shown in Figure 11 and in Table 8.

Table 8 Confusion matrix obtained from the accuracy assessment.

Land cover		
Class name	**User's accuracy**	**Producer's accuracy**
Artificial abiotic surfaces	0.92	0.88
Natural abiotic surfaces	0.83	0.78
Water bodies	0.98	0.90
Permanent snow and ice	0.86	1.00
Broad-leaved	0.87	0.81
Conifers	0.90	0.88
Permanent herbaceous	0.92	0.62
Periodic herbaceous	0.86	0.81
Land cover Change		
Class name	**User's accuracy**	**Producer's accuracy**
Other disturbances	0.35	0.71
Burned areas	0.67	1.00
Land consumption	0.81	1.00
Restoration	0.50	1.00

Figure 10 Land cover 2018 and land cover change 2017–2018 in Italy.

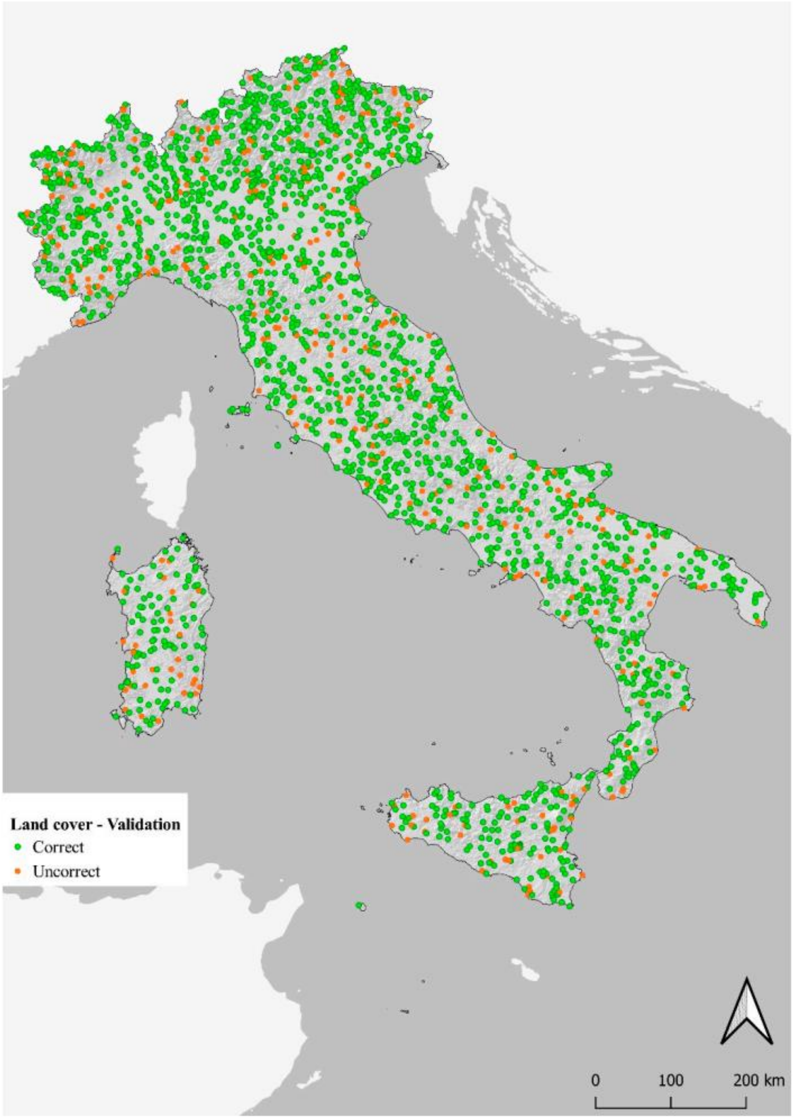

Figure 11 Result of the accuracy assessment with reference to the sample of 2400 photointerpreted points.

The map has 83% overall accuracy (Table 8). This value offers a general estimation of the accuracy, but it is influenced by the distribution of the error between the classes, especially those of land cover change. This circumstance also affects the spatial distribution of the classification errors detected by the sample points of Figure 11, since the land cover change classes do not have a specific distribution within the national territory.

Water bodies (98%), abiotic artificial surfaces (92%) and permanent herbaceous (92%) have the best user accuracy, while the land cover change classes show the lowest values, except for "land consumption" which presents a user accuracy of 81%.

As regarding the producer's accuracy, four classes have values of 100%, and 9 out of 12 greater than 81%. The lowest values regard the land cover change class of "other disturbances" (71%) and the land cover class of permanent herbaceous (62%). Analyzing the land cover data for 2018 and land cover change that occurred between 2017 and 2018 (Table 9), 10.2% of the Italian territory is represented by abiotic surfaces, of which about 70% are artificial surfaces. The rest of the territory is divided between woody vegetation (45.2%, with a prevalence of broad-leaved trees) and herbaceous vegetation (42.8%, especially periodic herbaceous). The remaining 1.49% is represented by water bodies (1.1%) and perennial ice and snow (0.4%).

Table 9 Distribution of land cover classes for 2018 and land cover change for 2017–2018 on the Italian territory.

	ha	***% Total***	***% Class***
Abiotic Surfaces	**3,081,573**	**10.22**	-
Artificial surfaces	2,127,999	7.06	*69.06*
Natural surfaces	953,574	3.16	*30.94*
Woody Vegetation	**13,614,064**	**45.17**	-
Broad-leaved	12,006,482	39.84	*88.19*
Conifers	1,607,582	5.33	*11.81*
Herbaceous Vegetation	**12,897,724**	**42.79**	-
Periodic	12,657,890	42.00	*98.14*
Permanent	239,834	0.80	*1.86*
Water and Ice	**449,510**	**1.49**	-
Water bodies	337,737	1.12	*75.13*
Permanent snow and ice	111,773	0.37	*24.87*
Land Cover Change	**97,102**	**0.32**	
Italy	*30,139,974*	*100.00*	-

The four classes of change occupy 0.3% of the national territory and about 90% are forest disturbances.

Table 10 shows that between 2017 and 2018, land consumption (4231 hectares, i.e., 64.1% of total land consumption) and restoration (636 hectares, equal to 70.2% of restorations) occurred mainly in the periodically herbaceous class. The artificial abiotic surface is the class with more changes (both land consumption and restoration) compared to the class extension.

Table 10 Distribution of land consumption between 2017 and 2018 among land cover classes.

	Land Consumption			***Restoration***		
	ha	**%**	**m²chg/ha**	**ha**	**%**	***m²chg/ha***
Abiotic Surfaces	**1,311**	**19.9**	**4.2**	**91**	**10.0**	***0.3***
Artificial surfaces	-	-	-	-	-	-
Natural surfaces	1,311	19.9	13.6	91	10.0	*0.9*
Woody Vegetation	**1,049**	**15.9**	**0.8**	**169**	**18.6**	***0.***
Broad-leaved	1,019	15.4	0.8	166	18.3	*0.1*
Conifers	30	0.5	0.2	3	0.3	*0.0*
Herbaceous Vegetation	**4,240**	**64.2**	**3.3**	**647**	**71.4**	***0.5***
Periodic	4,231	64.1	3.3	636	70.2	*0.5*
Permanent	9	0.1	0.3	11	1.2	*0.3*
Water and Ice	**0**	**0.0**	**0.0**	**0**	**0.0**	***0.0***
Water bodies	0	0.0	0.0	0	0.0	*0.0*
Permanent snow and ice	0	0.0	0.0	0	0.0	*0.0*
Italy	*6,600*	*100.0*	*2.2*	*906*	*100.0*	*0.3*

The decrease in vegetation caused by forest disturbances affects 89,596 hectares (Table 11). 12.4% (11,106 ha) of forest disturbances are burnt areas, while 87.6% (78,490 ha) had other causes.

Table 11 Distribution of forest disturbances occurring between 2017 and 2018 among woody vegetation classes.

	2017–18	***ha***	***%***
Burnt areas	**Total**	**11,106**	***12.4***
	Broad-leaved	6,610	*59.5*
	Conifers	4,496	*40.5*
Other disturbances	**Total**	**78,490**	***87.6***
	Broad-leaved	71,096	*90.6*
	Conifers	7,394	*9.4*
Total		***89,596***	***100.0***

The "other disturbances" class was further analyzed, showing that more than half of the disturbances are attributable to forest harvesting.

3.4. Discussion

The high overall accuracy of the map supports the rule-based classification method chosen, which allows the classification of the entire national territory with a relatively small number of training samples and frequent and fast updates of the map. Maxwell et al. (2018) and Thanh Noi and

Kappas, (2017) assessed that the machine learning classification method achieves a high accuracy level; however, these methods are usually applied to limited areas, at a regional or local scale (Pesaresi *et al.*, 2020; Rüetschi *et al.*, 2019), or to a few land cover classes (Agrillo *et al.*, 2021; Zimmermann *et al.*, 2007). Moreover, as illustrated by Zhang and Yang, (2020), machine learning approaches such as random forest need to be trained with many samples, which must be balanced in order to avoid misclassification errors (Millard and Richardson, 2015), making the map update process slower.

The developed methodology allowed for mapping the land cover and the changes in Italy, although with different results depending on the land cover class. The artificial abiotic surfaces and the changes related to land consumption were mapped using the National Land Consumption Map produced and updated by ISPRA. The use of ancillary data allowed for overcoming the issue related to spectral similarity between the natural abiotic surfaces and the artificial ones using Sentinel-2 images (Figure 12). The data have high accuracy, due to periodic updates and improvements through photointerpretation (Figure 13) . Hyperspectral images (such as PRISMA) could be used to improve the distinction of abiotic surfaces exploiting the bands in the SWIR region (Pepe *et al.*, 2020).

The class of permanent water bodies has an accuracy of over 90%, as the characteristics of the class can be well discriminated through the optical and radar instruments used in this research (Figure 14). The few omission errors regard boundary areas between water bodies and other land cover classes, such as river and lake banks, and some wetlands.

The class of permanent snow and ice has high accuracy (Figure 15), although it is slightly lower than permanent water bodies (86%). The accuracy of the class is affected by commission errors related to the seasonal variation of snow cover.

The classes of woody vegetation correctly distinguished between conifers and broadleaves trees in areas with homogeneous coverage, while in mixed forest pixels there are classification errors due to the spatial resolution of the data, causing mixed spectral signatures, and to the spectral similarity between the different types of trees (Figure 16) (Fassnacht *et al.*, 2016; Kuenzer *et al.*, 2014; Modzelewska *et al.*, 2020; del Río-Mena *et al.*, 2020).

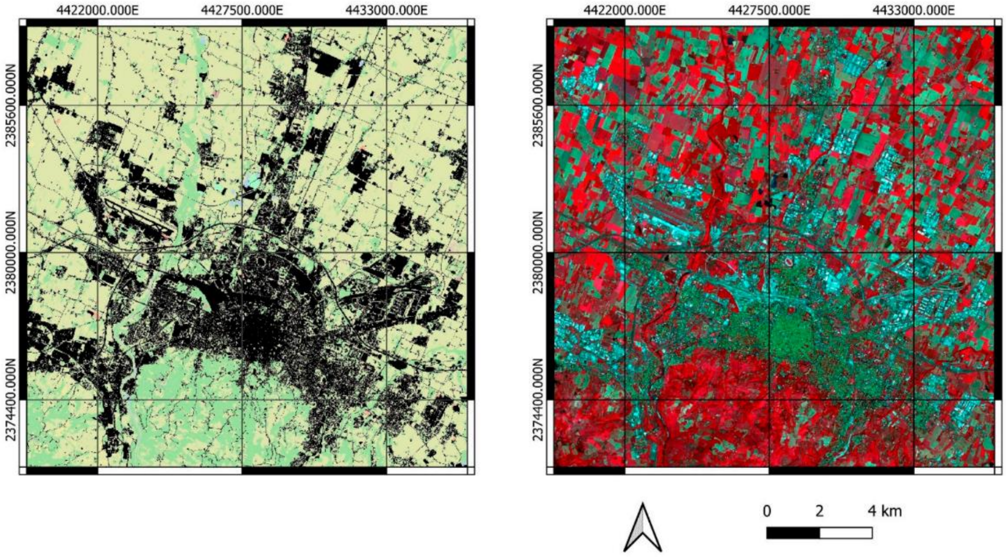

Figure 12 Artificial abiotic surfaces, in black (on the left), compared with a Sentinel-2 image (on the right).

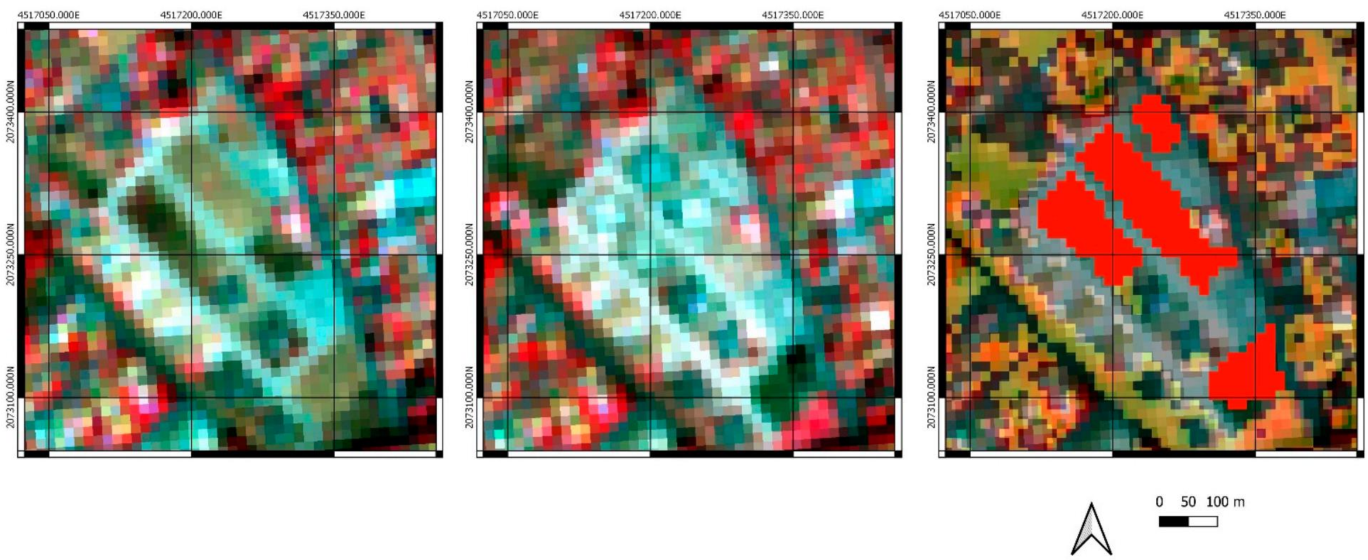

Figure 13 Land consumption, highlighted in red on the third image.

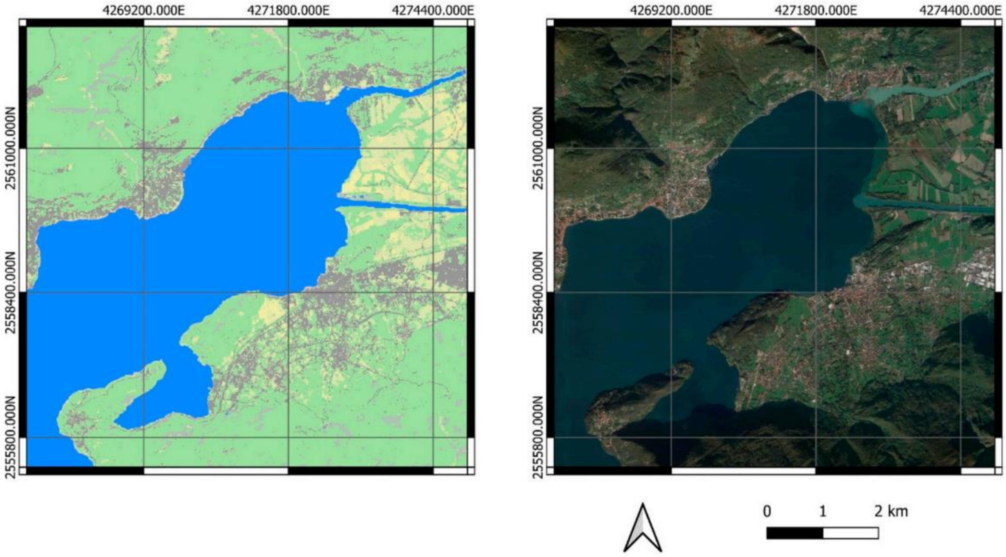

Figure 14 Water bodies.

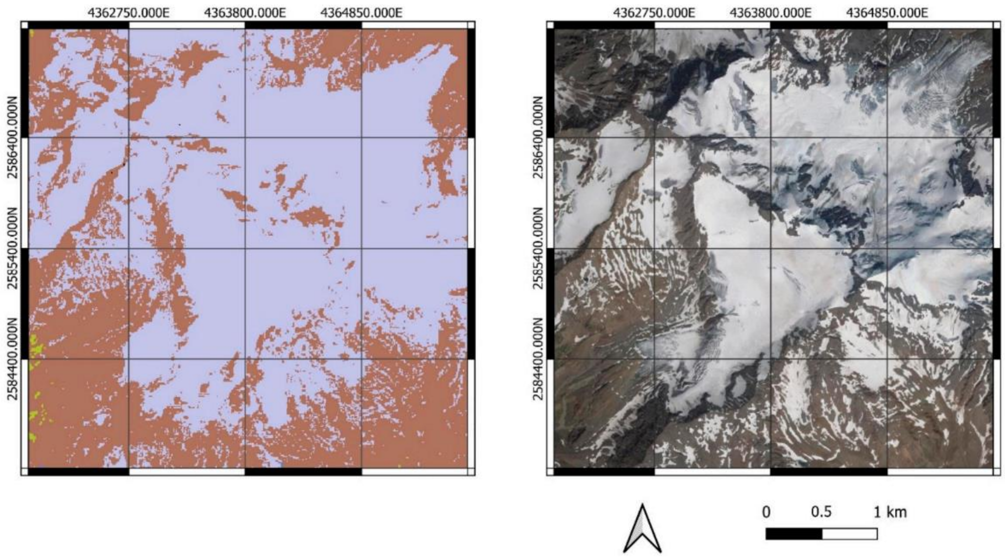

Figure 15 Permanent snow and ice.

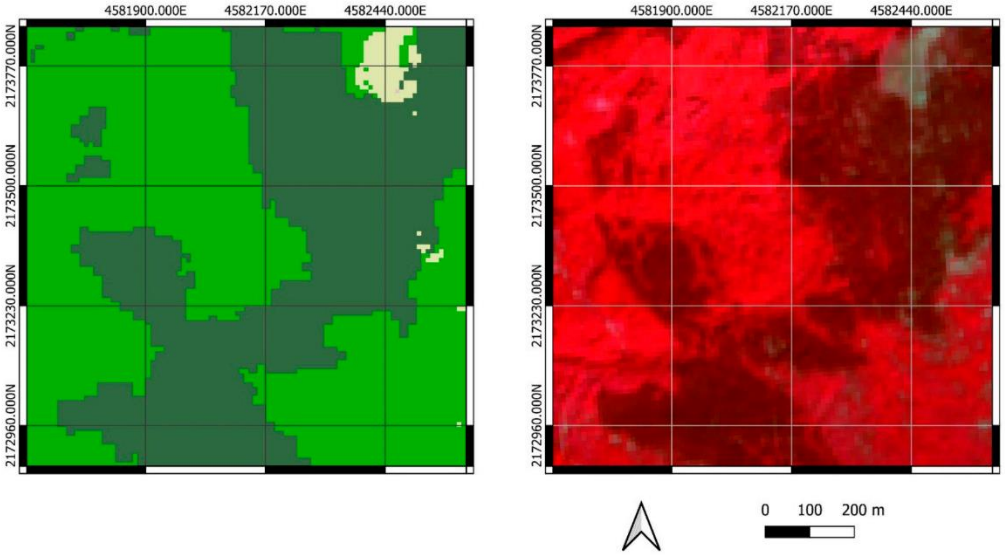

Figure 16 Woody vegetation, the distinction between broadleaves (brighter) and conifers (darker).

The results obtained for the herbaceous cover are related with the definition of the classes of periodic (Figure 17) and permanent (Figure 18) herbaceous, which consider the persistence of the herbaceous cover in the pixel throughout the year. In detail, pixels showing herbaceous cover in at least 95% of valid observations were classified as permanent herbaceous. This threshold is conservative and aligned with the definitions of CLC+ products (Kleeschulte et al., 2017), but it could be adapted to different temporal definitions.

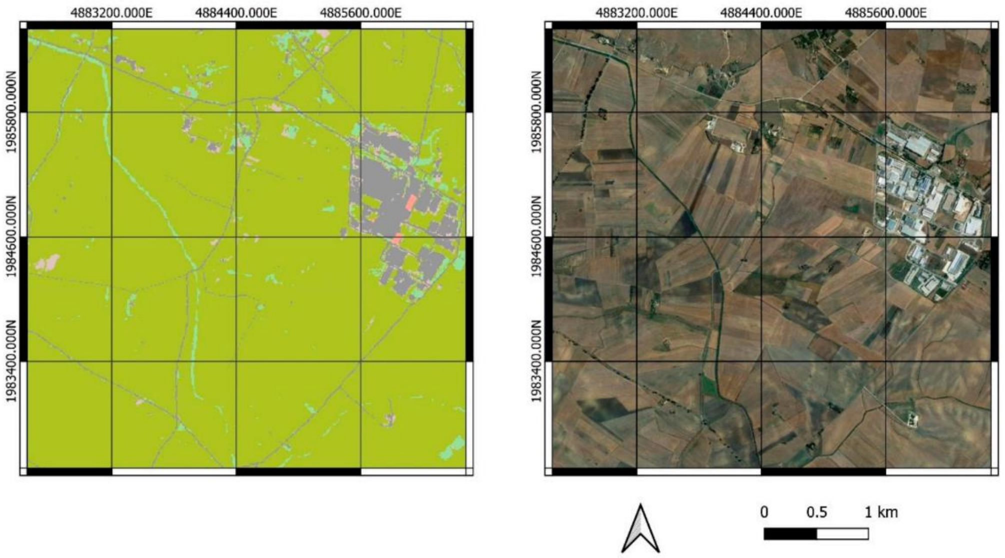

Figure 17 Periodic herbaceous.

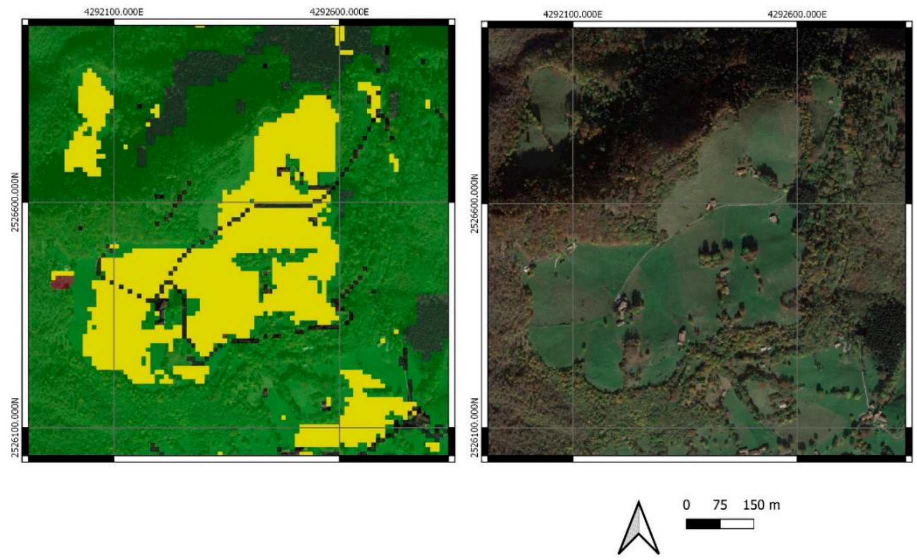

Figure 18 Permanent herbaceous.

Forest disturbances (Figure 19 and Figure 20) are mainly affected by commission errors, due to the adoption of wide thresholds for the identification of the classes. Commissions are mainly related with permanent crops, since they have a spectral behavior similar to forest disturbances during the year.

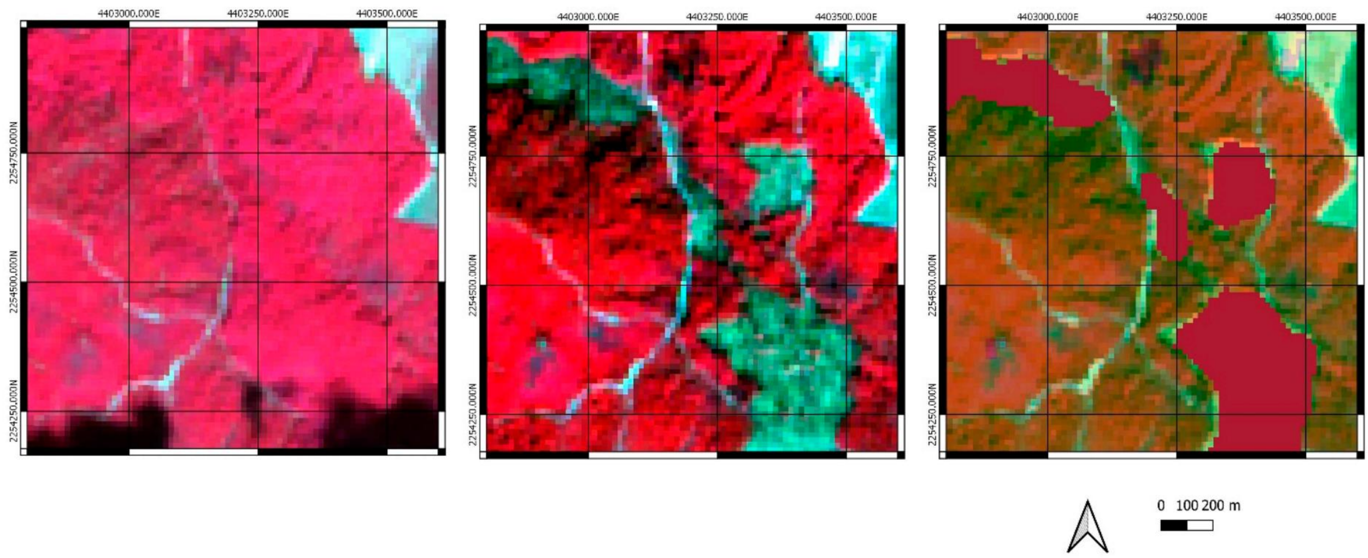

Figure 19 Other disturbances—Forest harvesting, highlighted in red on the right image.

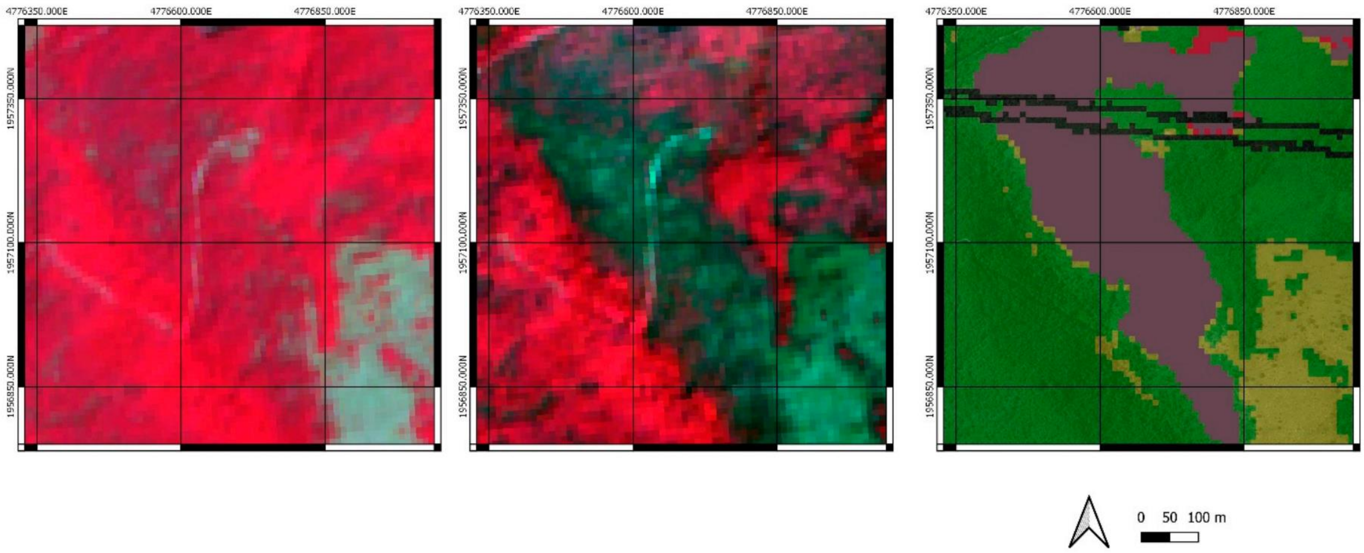

Figure 20 Burnt area, highlighted in grey on the right image.

The mapping of burned areas shows results aligned with the European Forest Fire Information System (EFFIS)[9]. The user accuracy of the class “other disturbances” is also low because it has been validated assimilating the class to “forest harvesting” only. The choice was made to evaluate the algorithm’s effectiveness in identifying this type of disturbance. Recently, several studies refer to the use of Sentinel data for monitoring forest disturbances, even if for limited areas (Puletti and Bascietto, 2019). The size of the study area conditioned the identification of changes, since cuttings are conducted with different techniques which vary according to the geographical area; therefore, in some cases it is not possible to clearly associate the spectral variations to a certain type of disturbance. The detection of forest harvesting is more effective in Central Italy, while there are classification errors where selective logging method is practiced, such as on the Alps.

The observation of the areas of potential disturbance over a longer period can allow the improvement of the data (Banskota *et al.*, 2014; Giannetti *et al.*, 2020).

The main objective of this research regards the formulation of a methodology that would allow the mapping and monitoring of land cover for the development of operational services, in compliance with the project requirements in terms of resolution, study area extension and update frequency. This paper presents a first product for 2018 that allowed the achievement of the project objectives and that actually lends itself to future refinements and improvements.

The EAGLE-based classification system allows for the introduction of new classes, such as shrubland, and can be modified for specific purposes, for example with land use attributes to identify agricultural activities in the herbaceous vegetation classes. The identification of shrubland

[9] https://effis.jrc.ec.europa.eu/, accessed on 25 February 2021

will make possible a better description of Mediterranean macchia, permanent crops such as vineyards, transition areas and forest boundaries. In this sense, it is necessary to conduct an in-depth analysis of the spectral behavior of shrubs, which are similar to scattered trees or to a mixed herbaceous and arboreal cover.

The use of Sentinel data constitutes an added value of the method, as they offer excellent results in the study of land cover at high spatial and temporal resolution, achieving high accuracy level with various classification algorithms (Phiri *et al.*, 2019). Furthermore, the free access policy also allows its use in contexts where financial resources for the acquisition of remotely sensed data are limited. Of course, additional classes would require additional methodologies, in the implementation of which the integration of hyperspectral data will be fundamental. The new HyperSpectral Precursor of the Application Mission (PRISMA) sensor, developed by the Italian Space Agency, has shown encouraging results in the distinction of forest types (Vangi, D'amico, Francini, Giannetti, Lasserre, Marchetti and Chirici, 2021). This technology will be particularly useful with the launch of the "Hyperspectral Imaging Mission for the Environment" (CHIME) (Rast *et al.*, 2021), whose data will be freely accessible in the coming years.

A characteristic of the proposed method is the definition of thresholds and a set of rules which can be modified and adapted to different conditions. The use of a decision rules classification method has made it possible to develop ad-hoc rules for each land cover class, while the definition of fixed thresholds for the entire national territory makes the method easily controllable and implementable. On the other hand, the Italian territory presents a strong heterogeneity from the morphological and climatic points of view that cannot be neglected. In this sense, an important development of the methodology will regard the adaptation of the thresholds according to the characteristics of the territory, considering in the first place the influence of hydrometeorological variables and seasonal climatic variations on the indices. This topic is addressed by the working group in the publication by Spadoni et al., (2020), which involved a study area consisting of three Sentinel-2 granules and will be studied in depth. In this way it is possible to better analyze the phenological evolution of vegetation in order to improve the distinction between conifers and broad-leaved trees and the identification of the different types of herbaceous classes and forest disturbances.

A further improvement regards the use of elevation models. At the moment this data has been used in the monitoring of changes associated with land consumption described by (Luti *et al.*, 2021) and will be able to guarantee important improvements in the land cover classification of mountain areas, especially in the shaded slopes of Alpine areas. In these areas it will be useful also to analyze the spatial correlation between pixels. This deepening can address the interpolation of missing or incomplete data (Zakeri and Mariethoz, 2021), enhancing some critical areas and allowing users to

overcome classification errors due to input data characteristics (such as image shadows, or cloud coverage).
The versatility of the algorithm and of the classification system make the methodology suitable for future integrations and for adaptation for specific purposes, while remaining consistent with the requirements proposed at the national and European level.

3.5. Conclusions

The research has allowed the creation of a land cover classification methodology capable of providing nationwide products with high spatial resolution and accuracy. The methodology shows satisfactory results, allowing the achievement of all design requirements. It was conceived with the aim of obtaining products that can be used for operational purposes, with a high update frequency. Indeed, they fit well into the Mirror Copernicus initiative of the Space Economy Strategic Plan of the Italian Ministry of Economic Development, which provides for the creation of frequently updated products at a national scale, for the monitoring of the territory.
The approach based on decision rules guarantees the control of the choices adopted for the classification, and a high level of modularity. In this sense, the rules can be refined by inserting variable thresholds according to the characteristics of the area, such as latitude and phytoclimatic zone, hydrometeorological variables or elevation, but also they can be accompanied by other rules for the identification of additional classes. In this sense, the classification system based on EAGLE has an easily expandable structure, remaining consistent with the European indications in terms of classification systems. The methodology can be implemented quickly and easily and has a low computational cost, allowing the production of land cover maps that can be updated annually and with high spatial resolution.
The maps produced by the application of the methodology can also be used as ancillary data for the development of algorithms for monitoring and characterizing specific coverage classes, such as the detailed distinction of forest types, the characterization of agricultural areas or a more detailed analysis of forest disturbances. The methodology is therefore promising, thanks to the simplicity of the approach on which it is based and the ability to be applied to large areas while maintaining high spatial resolution and accuracy levels. In this sense, it is unique in the panorama of European and national land cover monitoring products.

Appendix

A. This appendix describes the workflows of the classification methodology for each land cover class, organized in macro-class of water surfaces, abiotic surfaces and biotic surfaces (distinguished between woody vegetation and herbaceous vegetation).

Each workflow is structured in the following sections:

- "DATA" shows the input data used for the classification.
- "PRE-PROCESSING" illustrates the calculations of spectral indices and selected backscatter polarizations.
- "BINARY RASTER" reports the conditions and thresholds defined for the various inputs to create binary rasters (whose values are 0 or 1). A binary raster is calculated by attributing to each pixel the value 1 if the conditions defined by the threshold are verified and 0 elsewhere. A binary raster is calculated for each valid image acquisition (i.e., the number of valid acquisitions can be lower or equal to the total number of satellite acquisitions, depending on cloud cover).
- "VARIABLES" include the raster calculations derived from preprocessed data and binary rasters, which summarize the multi-temporal properties in a unique raster using statistics such as the median, or the percentage of occurrence calculated as:

$$P_{Class}(\frac{\sum BR}{N_{cloud}}) * 100$$

  where $\sum$ BR is the summation of binary rasters and N_{cloud} is the total number of valid acquisitions.

- The conditions for the classification of every land cover class are described in the lower boxes, which are based on the above "VARIABLES" and thresholds that are fixed or calculated from training input data.

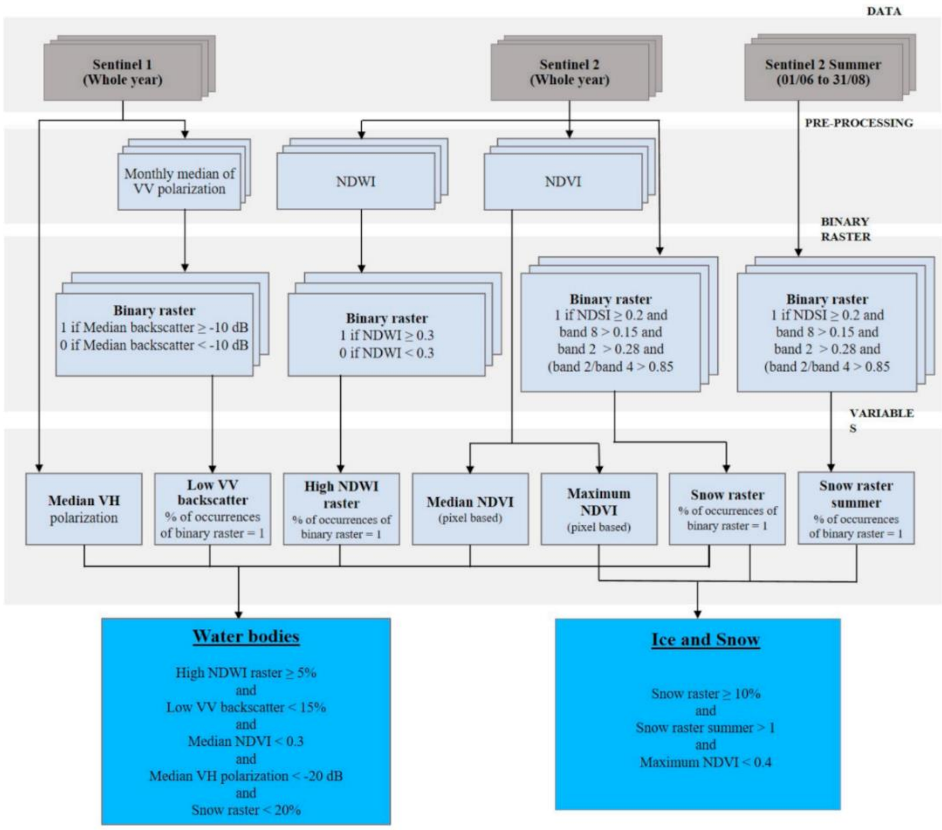

Figure 21 Workflow for water surfaces classification.

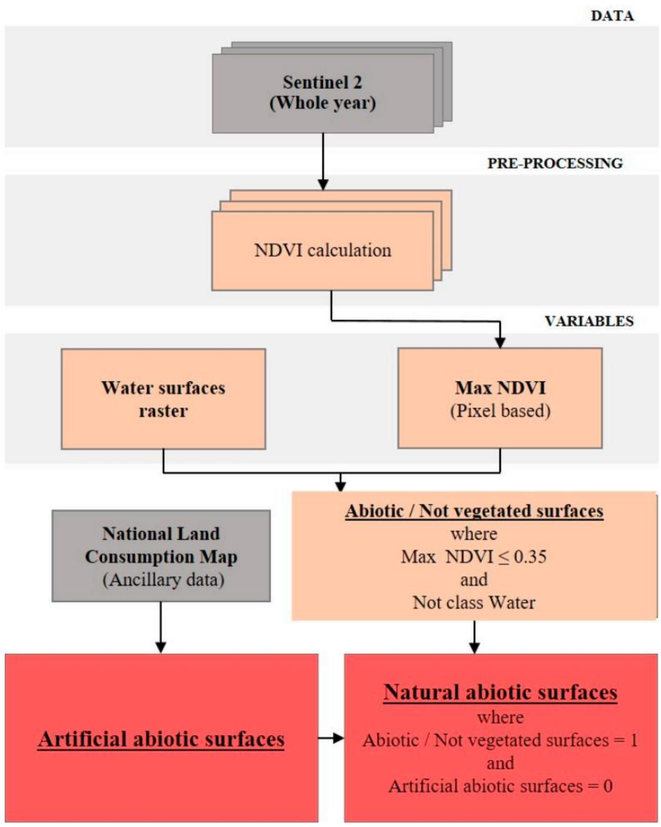

Figure 22 Workflow for abiotic surfaces classification.

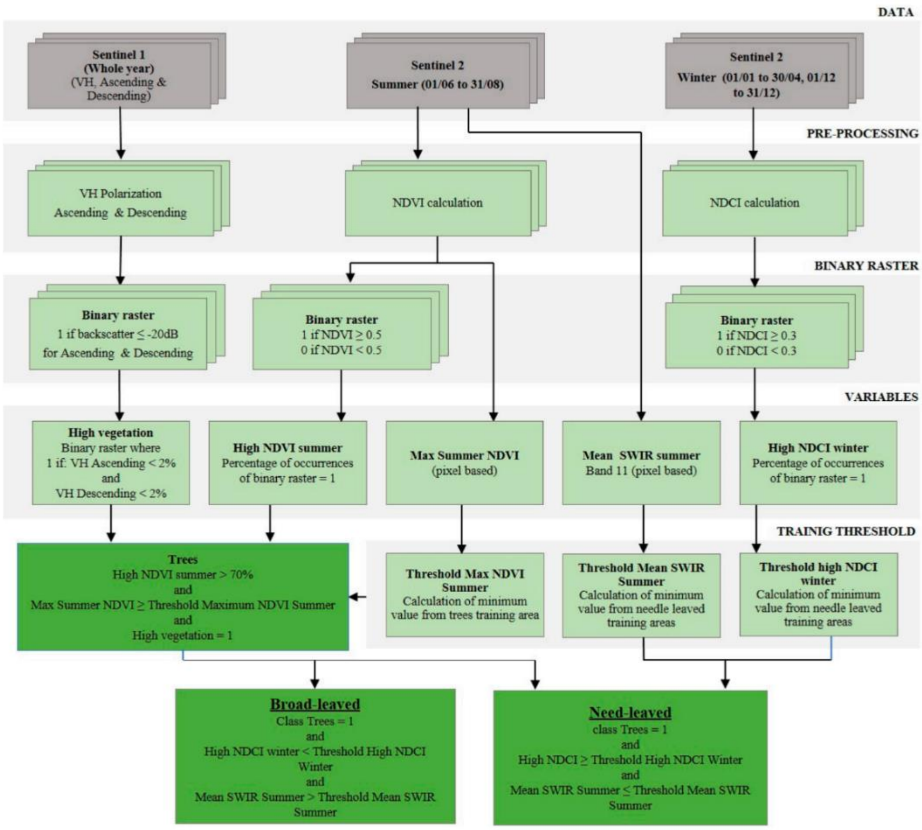

Figure 23 Workflow for woody vegetation classification.

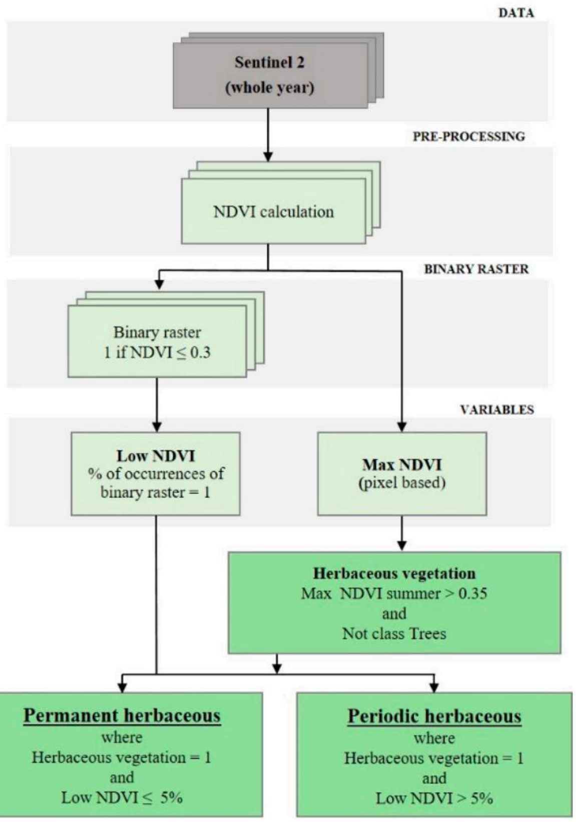

Figure 24 Workflow for herbaceous vegetation classification.

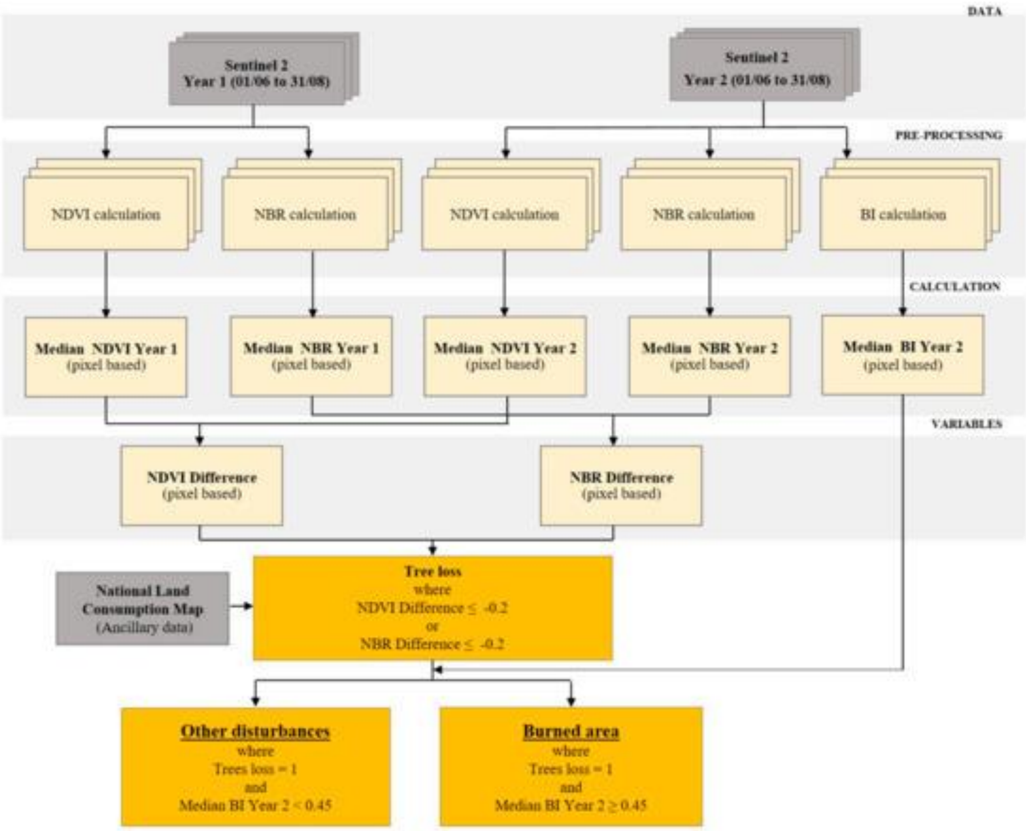

Figure 25 Workflow for forest disturbance detection.

4. Afforestation monitoring through automatic analysis of 36-years Landsat Best Available Composites

4.1. Introduction

According to Food and Agriculture Organization (FAO) definition, the forest is any "*territory with arboreal coverage greater than 10% compared to an extension greater than 0.5 ha, where the trees reach a minimum height of 5 m when ripe*", including also: "*[...] windbreak barriers and wooded strips of a width exceeding 20m, rubber tree plantations, and cork trees*" (FAO, 2001). Thirty-one percent of the World's total surface is occupied by forests (4.06 billion hectares) and Europe occupies a prominent position in this context since European forests account for 1,017,461 ha (FAO, 2020). Forests play a strategic role in terms of landscape, history, culture, and economy (RAF, 2018). They contribute significantly to human well-being, providing a multitude of ecosystem services among which wood production, carbon sink, food production, soil protection, climate control, biodiversity conservation, environmental decontamination, and hydrological cycle regulation (Spadoni *et al.*, 2020).

The Italian forest heritage is one of the most important in Europe - 11.8 million ha or 40% of the Italian land (Munafò et al. 2018) – but is exposed to many threats, in the majority of cases directly or indirectly related to anthropic activities. Indeed, great pressure is currently given by climate change, which undoubtedly led to an increase in the occurrence of extreme events such as storms, fires, and droughts (Forzieri *et al.*, 2021). Therefore, it results of primary importance to efficiently monitor forest ecosystems, to monitor their status, detect change processes and provide new tools for their conservation.

Forest change processes represent primary importance territorial transformations, and they have a central role in the definition of policies that aim to preserve forest resources and its ecosystem services provision. International organizations and other institutions at different levels have undertaken forest changes issues, giving definitions to frame this phenomenon.

United Nations Food and Agriculture Organization Forestry Department established a series of definitions, according to which, considering changes between forest and other land use classes, it is possible to distinguish, in addition to the loss of forest cover (deforestation), two different processes of tree cover growth: (i) *afforestation*, i.e. the conversion from other land uses into the forest, or the increase of the canopy cover to above 10%, and (ii) *reforestation*, i.e. the re-establishment of forest formations after less than 10 years with less than 10% canopy cover due to human-induced or natural perturbations. In this sense, reforestation is presented as a short-term process that, if longer than 10 years, falls within the definition of afforestation (FAO, 2001).

Forest changes occur heterogeneously around the globe, mainly because of different socio-economic backgrounds. While in Africa and South America and Eastern Europe deforestation trends up, in the European Mediterranean areas afforestation, generally resulting from the land-use change from agricultural to forest use, represents broadly the dominant process (Cervera et al., 2019; Gálos et al., 2013; FAO, 2001; Palmero-Iniesta et al., 2020). For what concerns the Italian territory, historical cartographic and inventory data confirms this assumption, showing that forest surface has been constantly increasing through the years, with an average annual growth rate of 0.3% (Angelucci 2011).

For this reason, afforestation represents a process of increasing importance, which certainly needs to be characterized to make the most of its potential and to correctly drive its development in line with forest sustainable management principles. Monitoring afforestation is therefore important for evaluating its positive effects. Afforestation processes increase the terrestrial carbon sink, create new ecological corridors, protect soil (Spadoni et al., 2020) and water, contribute to combating global warming, and increase biodiversity (Intergovernmental Panel on Climate Change, 2015; Shukla. et al., 2019). On the other hand, proper monitoring is also needed to assess the negative impacts that mismanaged afforestation processes may impose on a territory such as loss of grazing habitats, (Buscardo et al., 2008; IPCC 2015; Shukla et al., 2019; Veldman et al., 2015). Afforestation may also be a symptom of the effect of climate change, as in the case of migration to higher elevations or latitudes of the geographic range of tree species and of the tree line, the elevation or latitude above which trees give way to shrubs (Chersich et al., 2015). Finally, because of its important contribution to land use dynamics, afforestation processes have a central role in the implementation of policies that aim to preserve forest resources and provide ecosystem services (FAO, 2020).

The agreements stipulated by COP 26[10], the actions contained in the EU Forest Strategy (European Commission, 2021), the Italian national legislation on forestry and forestry supply chain (legislative decree 34/2018), and the Italian National Forest Strategy (MIPAAF, 2018) all raise the profile and priority of reforestation interventions and the construction of a series of cartographic products which collect information on forest heritage, including afforestation. Afforestation should be considered in reforestation programs, making the analysis of afforestation areas a central feature of forest and landscape planning and management. Because few existing datasets or products with appropriate spatial resolution can be used to monitor afforestation effectively, greater research attention should be focused on development of accurate and up-to-date tools to support afforestation-related decision-making processes.

[10] https://ukcop26.org/it/iniziale/

Since the afforestation process has a significant effect on the land cover dynamics, it is necessary to monitor its evolution, and for this reason, many different independent data in national and European environments are available, both inventory and cartographic. One of the most important is the Land Use Inventory of Italy (IUTI), an inventory data based on a sampling system and the classification in 6 macro-categories (Marchetti et al. 2012). Although IUTI is a valuable source of information, it is seldom updated (1990-2000-2008-2013-2016) and provides information on the predominant land cover in the area of analysis (in Italy 1,217,032 random points distributed in as many quadrants of 25 ha, with a minimum mapping unit of 5,000 m^2 and a minimum width of 20 m).

In this context, remote sensing (RS) certainly represents a high potential instrument (Filipponi *et al.*, 2018; Koch, 1999), since it allows to analyse and classify large areas enabling to obtain of frequently updatable results at different spatial scales (Agrillo *et al.*, 2021). Remote sensing data analysis can be improved thanks to cloud platforms as Google Earth Engine (GEE Gorelick et al., 2017), which allow managing large amounts of data and information even with small computational capabilities. Google Earth Engine provides numerous satellite missions as Landsat and the more recent European Sentinel Copernicus program which are by far the most used, being open-source data and covering many years of observations at different spatial resolution (Jönsson et al. 2018) .

Many information about land cover characteristics can be evaluated using the different bands of multispectral data, which can be summarised in single values, called Spectral Indices, or, when they are related to vegetated land cover, Spectral Vegetation Indices (SVI) (Oberti 2009, Spinsi et al. 2012). The most used SVIs in literature are the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI) and Normalized Burn Ratio (NBR). Besides the SVI, other indices, known as Tasseled Cap (TC), represent a valuable tool for this kind of analysis. These indices resume multispectral data into information about three dimensions of the land cover considered: Brightness (B), Wetness (W), Greenness (G) and Angle (A). SVI and TC acquire even more value when referred to different periods, allowing to build temporal trends which represent important additional information about the phenology of a given territory (Persson et al. 2018). Hence, to elaborate a land cover classification and to produce forest cartography, an efficient method consists in analysing multitemporal series of SVIs or TC, to dispose of very complete information synthetically. In doing so, it is possible to establish classification rules by imposing thresholds to discriminate different land cover classes and recognizing the ones of interest (De Fioravante et al., 2021; Luti et al., 2021; Smiraglia et al., 2020; Spadoni et al., 2020).

Landsat and Sentinel images find different applications considering their intrinsic characteristics. Landsat data are available since the 1970s, providing so very long time data series at medium spatial resolution (30 m), useful to analyse land cover trends and evolution. Sentinel instead, available since 2015, has, as its strong point, the high spatial resolution (10 meters) and short revisiting time (in Italy 2/3 days), which make it a tool of great utility for applications where high precision is needed. Sentinel-1 data and Sentinel-2 images, thanks to their spatial and temporal resolutions, have been recently used in many studies on the analysis of forest disturbances and land cover change detection. In (Laurin *et al.*, 2021), Sentinel-1 data were used to detect the damages caused by a wind storm in 2018 in Vaia (Italy) while Sentinel-2 data were used to map clearcuts in Tuscany (Francini *et al.*, 2021), and to estimate areas and relative confidence intervals of forest fires, wind damages and clearcuts in Italy for 2018 (Francini et al., 2022).

Other research activities involved topics in which the high temporal resolution of some satellite images allow to monitor land cover and land cover changes at the annual or sub-annual level. (De Fioravante et al., 2021) integrated the Sentinel-2 and Sentinel-1 data to assess land cover (2018) and land cover changes (2017-2018) in Italy, while (Francini *et al.*, 2020) used PlanetScope imagery to develop a methodology to detect near-real-time forest disturbances in Tuscany (Italy). While the land cover and its changes monitoring by remote sensing is a widely analysed topic, afforestation assessment remains a theme mainly unrepresented, being treated as a central subject only in a few studies (Raha, 2010; Shen *et al.*, 2019) in which relatively low accuracy was reached (less than 75%). For example, (Huang *et al.*, 2017) and (Qiu *et al.*, 2018) used MODIS images (500 m spatial resolution) to map forest growth areas in China. However, products with a spatial resolution of 500 meters are not useful for monitoring afforestation in fragmented and morphologically complex areas typical of the Italian territory, where afforestation processes are characterised by patches of limited size.

The present research aims to illustrate a new methodology that exploits the Landsat time series to demonstrate that RS can be valuable support not only for monitoring stable land cover classes (De Fioravante et al., 2021) and forest disturbances (Giannetti *et al.*, 2020), but also for predicting afforestation areas.

In the first section of this research, we describe the training data and Landsat images (1984-2020). Then we explain the use of the input data to assess the temporal trend difference of spectral response of the *classification periods (non-forest, afforestation* and *forest)*. Afterwards we focus on the use of the random forest (RF) algorithm to automatically distinguish the classes and to identify the combination of predictors which allows reaching the largest accuracy.

Finally, we discuss the results and how they support the use of remote sensing data in the distinction between *non-forest, afforestation* and *forest.*

4.2. Materials

4.2.1. Study area and reference data

The research interested sixty-one areas, where afforestation occurred between 1988-2020, distributed along the Italian peninsula, especially on the mountains range (Alps and Appennines). Forty-two points are located in areas with an elevation higher than 600 m above sea level, and twenty-six of them are located in areas above 1,000 m above sea level; the slope of almost 3/4 of the points is higher than 10% (Figure 26). These areas cover a surface of 269,197 ha, and each polygon is on average 4 ha large and they were defined using the IUTI dataset and other images, as explained in section 4.3 and 4.3.1.

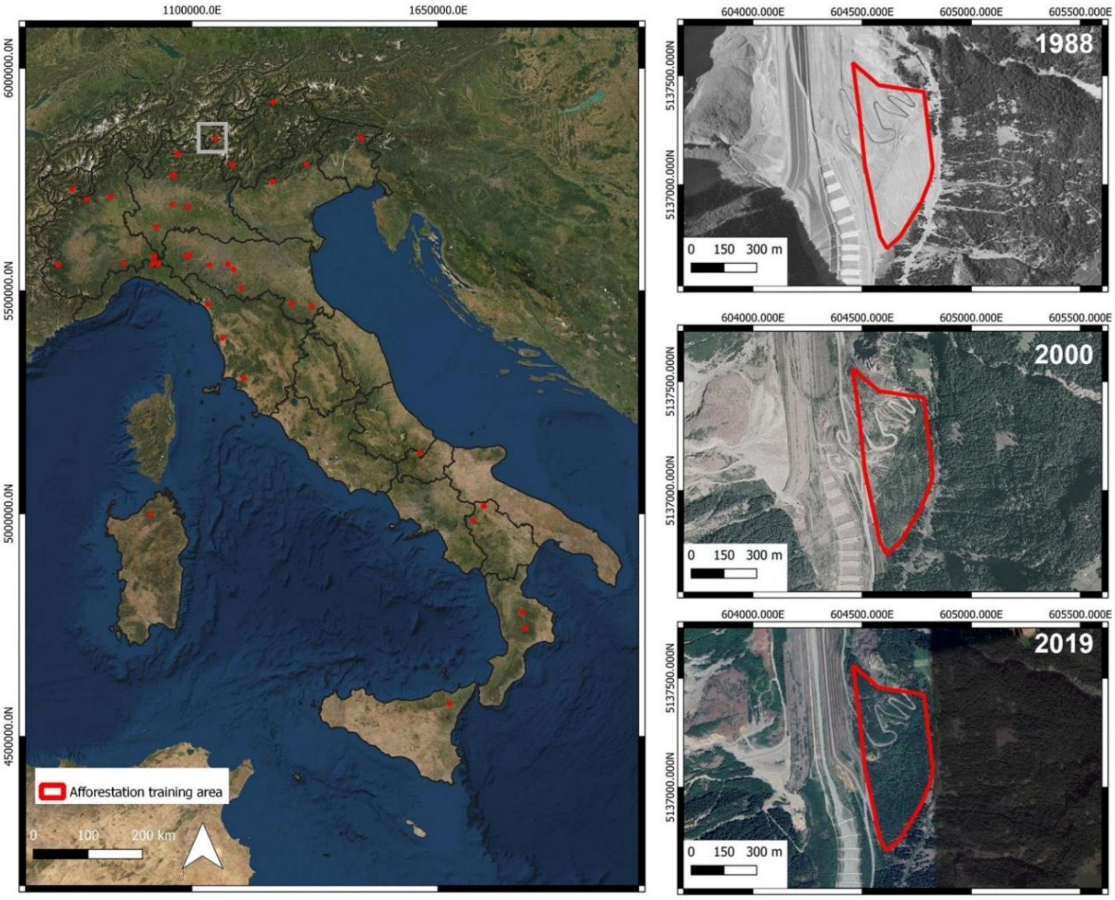

Figure 26 Study area and distribution of reference dataset. On the right there is an example of afforestation occurred between 1990-2000: in the upper image (1988) there is the non forest cover, in the centre (2000) the forest partially cover the area, in the last image the forest cover is complete (2019).

4.2.2. Land Use Inventory of Italy (IUTI)

Land Use Inventory of Italy IUTI was realized by the Italian Ministry of Environment and Protection of Land and Sea in 1990 (updated in 2000-2008-2013-2016) in the framework of the

Extraordinary Plan of Environmental Remote Sensing. It is composed of 1,217,032 randomly selected points, which have been classified through photointerpretation, considering a minimum mapping unit of 5,000 m^2 and a minimum width of 20 m. The classification system is based on the Intergovernal Panel on Climate Change guidelines, and it is composed of three hierarchical levels, which aim to identify the land cover categories which are important for the Kyoto Protocol and to integrate the National Inventory on Forest and Carbon Pools (INFC) results on woods and other woody areas categories based on the FAO definition (Sallustio *et al.*, 2016). Analyzing the evolution of the dataset it is possible to identify those areas where afforestation occurred.

4.2.3. Landsat data and Best Available Pixel Composite

In this study, the Landsat surface reflectance data available (for Landsat mission 5-7-8) in the Google Earth Engine (GEE) archive were used. We used Landsat missions 5, 7 and 8, because the product of the previous missions had a lower spatial and temporal resolution. Moreover, we analyzed data starting from 1985 because they are comparable with the training dataset. Landsat images are composed of six bands: three visible (blue, green, and red), one near-infrared (nir), and two short-wave infrared (swir1 and swir2). All of them are already processed to orthorectify surface and brightness (the thermal infrared) reflectance, atmospherically corrected, and cloud masked. Landsat images were used to create Italian annual composites from 1984 to 2020, from June 1st until August 31st, using the Best Available Pixel (BAP) procedure (White *et al.*, 2014).

The BAP aims to fill the final image mosaic with the composite best available pixel surface reflectance value. The selection of the best pixel of the images collection takes into account four different criteria: (i) sensor score, to penalize Landsat 7 images where the Scan Line Corrector malfunction (SLC-off) is present, (ii) target day score, to preferably select the images acquired close to a defined acquisition day (in this case 15th of August), (iii) distance to cloud/cloud shadow score, to decrease the scores of those pixels which are in the proximity the cloud cover, (iv) opacity score (calculated using opacity band produced by LEDAPS, Schmidt et al. 2013), to prefer the pixel with low opacity. Scores (i) and (ii) are applied to the whole image, while scores (iii) and (iv) are applied to each pixel. BAP composites calculation was performed using the GEE application developed by Francini (2021, http://code.earthengine.google.com/?accept_repo=users/sfrancini/bap) which allows the calculation and downloads of the BAP composites (Hermosilla *et al.*, 2015a, 2015b; White *et al.*, 2017). For more information on the criteria involved in the BAP pixels selection see the BAP-

GEE documentation[11]. For more info on the BAP algorithm see (Griffiths *et al.*, 2013; White *et al.*, 2014).

4.3. Methods

The methodology presented in this research employed Landsat images to define the feasibility of the assessment of afforestation areas in the last thirty-six years through remote sensing. The analysis of many years allows verifying the evolution of the *classification periods*, considering the afforestation, which can have a different duration depending on the species.

Using 61 training areas consisting of polygons defined using IUTI elements, orthophotos and very fine resolution images to identify the afforestation, a preliminary analysis of spectral response of *non-forest*, *afforestation* and *forest* areas was developed. Then RF algorithm, calibrated and validated, was applied to automatically distinguish the *non-forest*, *afforestation* and *forest* land cover (Figure 27). The processes were developed using Qgis, GEE and R, which are useful in processing spatial and statistical data.

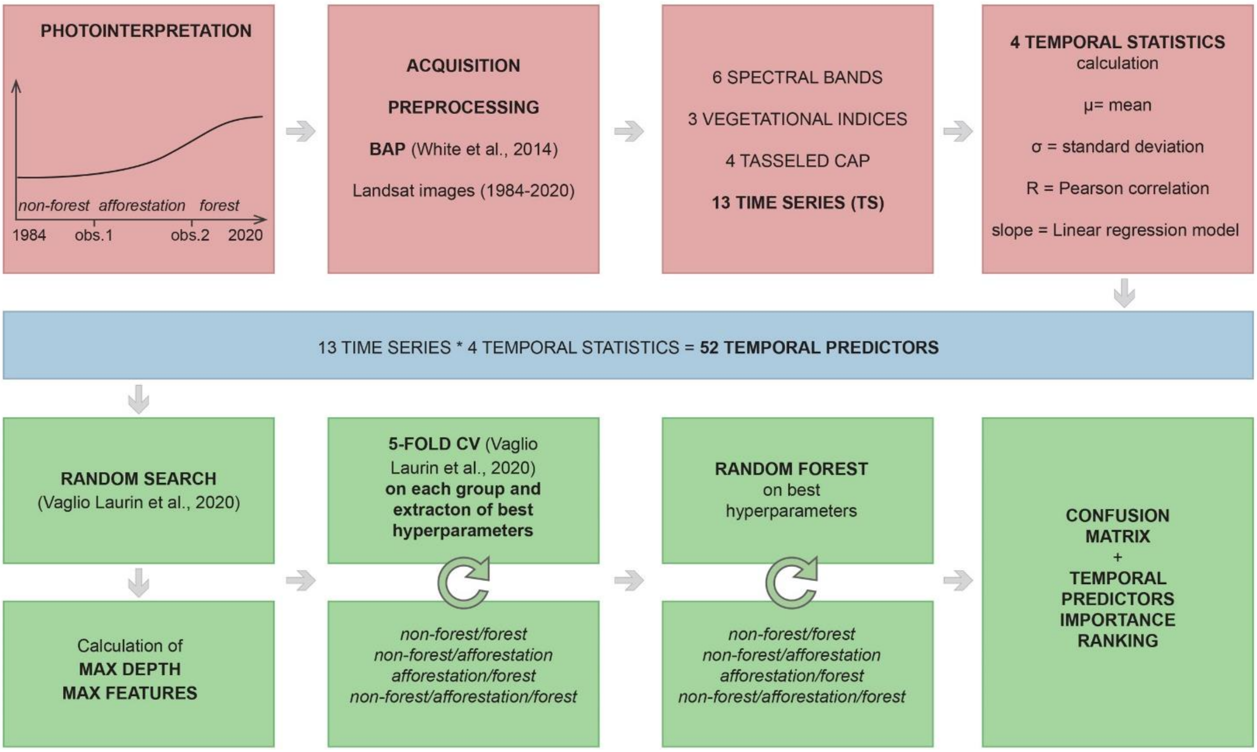

Figure 27 Workflow of the methodology. The steps to distinguish *non-forest, afforestation* and *forest* are briefly described.

[11] https://github.com/saveriofrancini/bap

4.3.1. Reference data acquisition

The orthophotos used for the photointerpretation process are available as Web Map Services (WMS) services of the Italian Ministry of Environment, Land and Sea, and they are referred to 1988-1989, 1994-1998 (black and white) and to 2000, 2006, 2012 (colour), with a spatial resolution of 50 cm. The more recent Very High Resolution images (30 cm spatial resolution, available since early 2000s) are freely available thanks to the web service of Qgis and Google Earth Pro.

The reference dataset (Figure 26) is produced through photointepretation of orthophotos and very fine resolution images, available from 1988 to 2020; in addition to known areas, to establish the afforestation areas, IUTI dataset has been used, considering the points which identify a forest change, from non forest to forest area. Sixty-one training areas have been collected and finally integrated with other information on observation date and polygon dimension.

Orthophotos and IUTI datasets were used to define the temporal range in which the afforestation process occurred. More specifically, two years were identified: (i) the year corresponding to the transition between *non-forest* and *afforestation* and (ii) the year corresponding to the transition between *afforestation* and *forest*.

4.3.2. ***Temporal predictors*** calculation and analysis

The indices calculation for each composite BAP and the analysis of the resulting temporal series were carried out to calculate the *temporal statistics.* More specifically, we augmented the six Landsat bands of our BAP composites by calculating seven additional indices: (i) Normalized Difference Vegetation Index NDVI, (ii) Normalized Burnt Ratio NBR, (iii) Enhanced Vegetation Index EVI, plus (iv) Tasseled Cap brightness B, (v) wetness W, (vi) greenness and (vii) angle A. As a result of this step, we obtained 13 temporal series of 36 years, hereinafter called Time Series (TS, Figure 28).

From the TS we calculated for each polygon, and for each of the *non-forest, afforestation, and forest* temporal ranges four *temporal statistics*: (i) the mean and (ii) the standard deviation of the TS, (iii) the slope, and (iv) the Pearson correlation coefficient (R) of the regression line obtained by linear interpolating the TS over time. As a result of this step, a set of $13\, TS * 4\, temporal\ statistics = 52\, temporal\ predictors$ were obtained to distinguish spectral response between *non-forest/afforestation/forest* areas and to investigate the spectral trend differences using density plots.

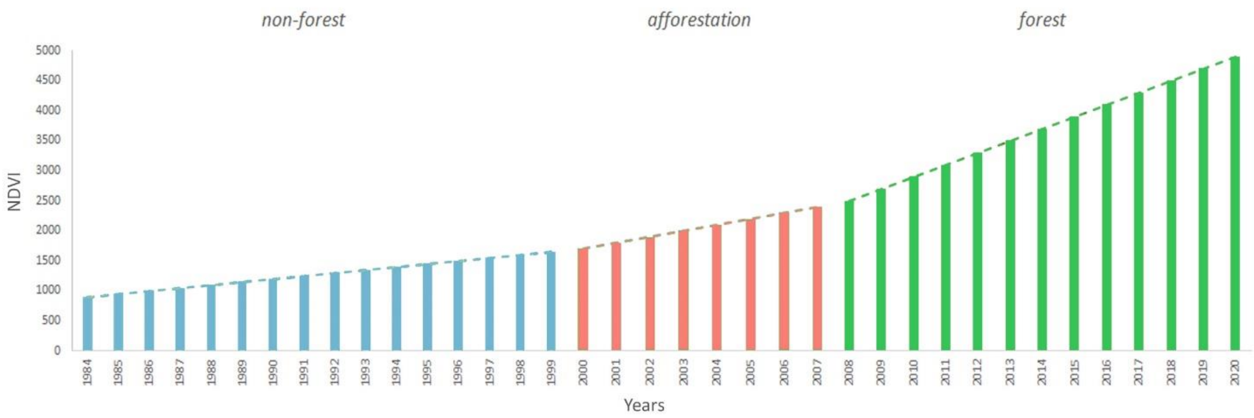

Figure 28 Example of NDVI series in an afforestation pixel. For each of the *classification period* was calculated the trend line to calculate the Pearson correlation coefficient.

4.3.3. Random Forests

4.3.3.1. Random Forests model and hyperparameters

To classify *forest, non-forest,* and *afforestation* areas we used a random forests model and the *temporal predictors* as variables. Random forests (RF) is a decision tree algorithm, often used for forest classification (Hawryło et al. 2020) and where a subset of predictors is used to build a set of regression trees, applying the Out-Of-Bag sample (OOB) procedure (Breiman, 2001). RF is also useful as it allows calculating the variable importance. RF algorithm was chosen because it achieves large accuracies when compared with other machine learning approaches (Nguyen *et al.*, 2020) and it is not influenced by overfitting (Belgiu and Drăgu, 2016). Random forests hyperparameters are model parameters that can be calibrated to obtain the minimum error rate (Laurin *et al.*, 2021). In this research, we tuned the following two: max features (the number of features considered) and max depth (the maximum depth of the tree, the longest path between the root node and the leaf node). The number of trees in the algorithm was set to 300 and the minimum number of samples for splitting a node was set to one.

4.3.3.2. Max depth and max features calibration

Max features and max depth hyperparameters were calibrated using the procedure named random search (Bergstra and Bengio, 2012; Laurin *et al.*, 2021). Specifically, we defined a grid of max features (1-52) and max depth (1-40) ranges, from which a subset of 100 combinations were randomly selected. To assess the performance of each tested combination we used k-fold Cross-Validation (CV) with k = 5.

Cross Validation serves to train and validate the model on different data and it is useful to reduce the overfitting (Whittaker *et al.*, 2010) and to test the robustness of the model (Laurin *et al.*, 2021). Using k-fold CV, the dataset is divided into k groups, which are composed of the same quantity of data: one of these folds is defined as the validation group, while the k-1 groups are considered as training data. The procedure is then repeated k times, each time using a different fold as a validation group. Finally, the validation data is merged and the k-fold CV Overall Accuracy (OA_{cv}) is calculated. As a result of this step, we obtained the max features- max depth combination that allowed to obtain larger OA_{cv}.

4.3.3.3. Performance assessment

Although the k-fold CV procedure allowed to rigorously estimate the accuracy (OA_{cv}), we further assessed the accuracy of the model exploiting the best max features-max depth combination using Out Of Bag (OOB) data, which allows getting an unbiased estimate of the error (Breiman, 2001). During random forests training, two independent sets are created: (i) the bootstrap sample, i.e. the data chosen to be "in-the-bag" by sampling with replacement, and (ii) the out-of-bag (OOB) sample, i.e. all data not chosen in the sampling process. During model training, this process is repeated and many OOB samples are created. Since they are never used to train the model, they are *never-seen-before* data and can thus be used to assess the performance of the model. Finally, all OOB samples are aggregated and OOB predictions are compared with reference data by constructing a confusion matrix from which we calculated the OOB. Overall Accuracy (OA_{oob}), omission errors ($omissions_{oob}$), and commission errors ($commissions_{oob}$). The procedure was repeated four times, to compare the *classification models* as follows: (i) *non-forest/afforestation,* (ii) *afforestation/forest,* (iii) *non-forest/forest,* and (iv) *non-forest/afforestation/forest.* For each of them, classification omission and commission error were calculated, to respectively identify the pixel non-classification and the misclassification.

4.3.3.4. Variables importance calculation

To identify *temporal predictors* that mostly contributed to increasing the accuracy of the model we analyzed the variable importance ranking outputted by random forests. The variable importance was expressed in terms of the Mean Decrease of the Gini index (MDG, Nicodemus, 2011). MDG is the measure of the variable contribution to the homogeneity of the random forests trees component. Large MDG values correspond to variables that relevantly contributed to increasing the performance of the model and thus the accuracy achieved. Conversely, small MDG values correspond to variables that did not contribute to increasing the accuracy of the model.

4.4. Results

4.4.1. Classification accuracies

Random forests validation using Out-Of-Bag data (OA_{oob}) was performed on all the *temporal predictors* of the best iteration of 5-fold CV (Table 12). The largest OA_{oob} was reported by the model classifying between *afforestation* and *forest* (80%), followed by the classification between *non-forest* and *forest* (77%). *Non-forest* and *afforestation* and *non-forest/afforestation/forest* classification model reached an overall accuracy smaller than 70%.

For each class of the four *classification models* omission ($omissions_{oob}$) and commission ($commissions_{oob}$) have been calculated (Table 13). Both omission and commission errors are larger than 10%, with the maximum value (40%) registered by non forest class in the classification between *non-forest*, *afforestation* and *forest*; the smallest error values are registered in the *afforestation* and *forest classification*, where omission error of the *forest* class is 11% and the *afforestation* one is 27%; as regard commission error the *afforestation* registered the smallest value (14%), while the *forest* one is 24%.

The large random forests OA_{oob} values were reached thanks to the use of the hyperparameters with the largest OA_{cv}, calculated using random search and 5-fold CV on the classification models, to identify the couple of hyperparameters which allow to reach the best results. Figure 29 underlines that the OA_{cv} is larger than 65% in every hyperparameters combination considered. Regarding the best iteration (value "max" in the figure) the smallest OA_{cv} is registered in the *non-forest, afforestation and forest* (72%) and *non-forest* and *afforestation* (75%) classification models. The model performance is larger in the classification of *non-forest*/*forest* and *afforestation*/*forest*, where the best iteration OA_{cv} reaches 83% in the *non-forest/forest* classification and 87% in *afforestation/forest* one.

Table 12 *Classification models* Overall Accuracy of random forests (OAoob), considering all the *temporal predictors*

classification models	***max features***	***max depth***	***OA_{oob}***
non-forest/afforestation	24	21	0.70
afforestation/forest	7	5	0.80
non-forest/forest	52	30	0.78
non-forest/afforestation/forest	24	28	0.67

Table 13 Omission and commission error for each classification model

	***omissions$_{oob}$* (%)**				***commissions$_{oob}$* (%)**			
	non-forest / afforestation	*afforestation / forest*	*non-forest / forest*	*non-forest / afforestation / forest*	*non-forest / afforestation*	*afforestation / forest*	*non-forest / forest*	*non-forest / afforestation / forest*
non-forest	26.2		23.0	36.1	31.8		21.7	40.0
afforestation	34.4	27.9		39.3	28.6	13.7		32.7
forest		11.5	21.3	23.0		23.9	22.6	25.4

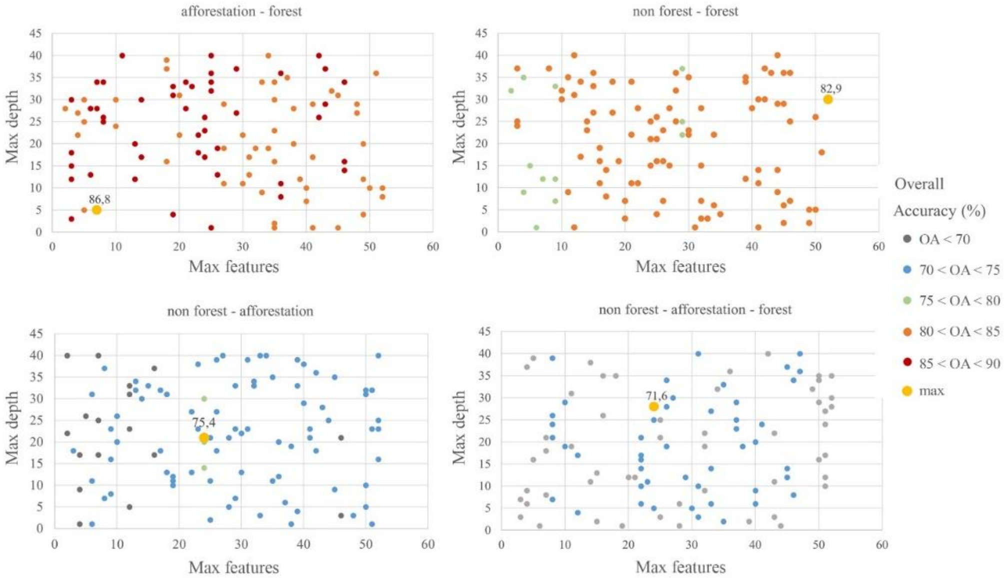

Figure 29 Five-fold cross validation overall accuracy (OA_{cv}), the value “max” is referred to the best iteration. Each OA_{cv} of every couple of hyperparameters of each *classification models* is illustrated, divided in accuracy range.

4.4.2. Random forests *temporal predictors* importance and *classification periods* differences analysis

The *temporal predictors* importance ranking is illustrated in Figure 30. The most important predictors for the model *afforestation/forest* were the median value of the green (MDG= 2.62) and the blue (MDG=2.62) bands, the Tasseled Cap Greenness R (MDG= 3.04), and the slope value of the Tasseled Cap Brightness (MDG=1.76) and the green band (MDG=1.72). The other predictors had smaller importance scores. In the *non-forest* and *forest* models, only two predictors were indicated as more important, that is the median value of the green (MDG= 10.35) and the blue

(MDG= 19.36) bands; the same ranking was registered in the *non-forest*, *afforestation*, and *forest* models, even if the other predictors had a larger score than the ones of *forest*, *non-forest* classification. Finally, the first predictor in the importance ranking of *non-forest* and *afforestation* models is the R value of the EVI (MDG= 9.90) and, to a lesser extent, the slope value of the blue band (MDG=2.69).

In Figure 31 are presented the *temporal predictors* of the *non-forest*, *afforestation* and *forest*, related to the variables which were considered as most important in the random forests. Tasseled Cap Greeness, R value (TCG, Figure 31d), registers the largest difference between data, in particular the *non-forest* reaches the largest values, as in the EVI (c). As regard the median values of blue (a) and green (b) bands and the slope values of green (e) band and TCB (f), the *forest* reaches the largest values, and the *afforestation* and the *non-forest* are quite similar.

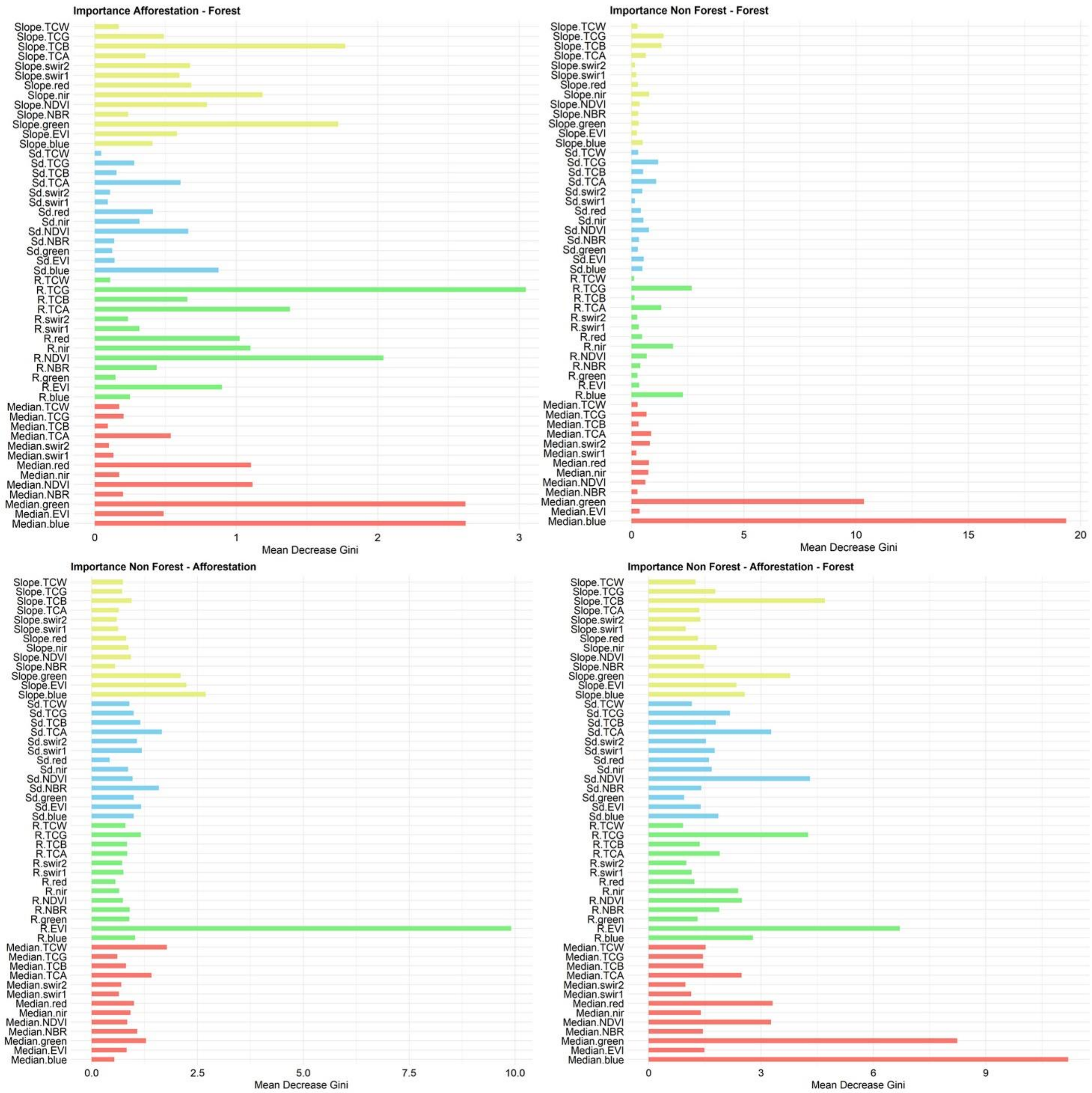

Figure 30 *Temporal predictors* importance ranking. The importance value (Mean Decrease Gini) of each *temporal predictors* in every *classification model* is illustrated.

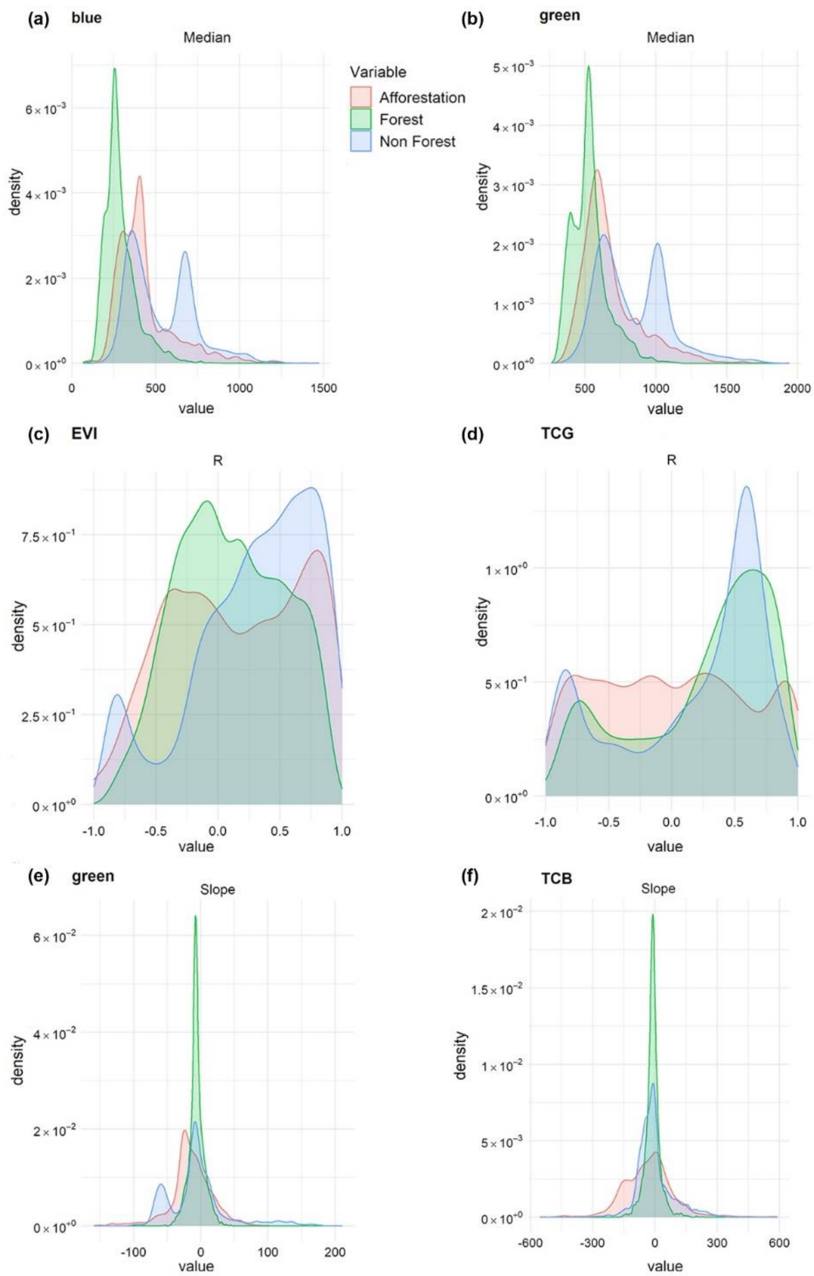

Figure 31 Density distribution plots. There are the most important *temporal predictors* of *non-forest*, *afforestation* and *forest*; plots allow to estimate the differences of the *classification periods*.

4.5. Discussion

Afforestation monitoring is important because it influences ecosystem services, but it can also cause some issues, as land fragmentation and habitat loss. In this research we aimed to demonstrate the efficacy of remote sensing data in the assessment of the difference in spectral response and in the automatic classification of *non-forest*, *afforestation* and *forest* areas. To classify them, we analyzed 36 years of Landsat images using sixty-one training areas and a random forests algorithm. The large overall accuracy reached with the classification algorithm supports the chosen analysis approach, confirming the efficacy of the remote sensing in the automatic classification between *non-forest*, *afforestation* and *forest*.

The methodology by using Landsat images allows analyzing long time series of images, which is essential to assess forest evolution, which generally is slow. Moreover, both the first analysis on *temporal predictors*, and the random forests algorithm are easily replicable, also using different input images. To limit the overfitting due to the use of neighbouring pixels as training and validation data the data were elaborated at polygon level, considering the training data boundaries. Furthermore, random search and k-fold cross validation have been useful to identify the combination of random forests hyperparameters with the highest accuracy, to ensure the maximum performance of the algorithm. Considering the results of the random forests validation, the difference in OA_{oob} values can be due to the similarity in spectral response in *non-forest* and *afforestation* land covers, which registered smaller accuracy values, while *forest* and *non-forest* or *forest* and *afforestation* classification models gives better results because there is a stronger difference in the spectral response of the two classes considered. It is notable that, even if the k-fold cross validation and the Out Of Bag accuracy results are not comparable, the OA_{cv} and OA_{oob} ranking is the same, with the largest accuracy value registered in the distinction of *afforestation* and *forest* and the smallest ones in the distinction of *non-forest*, afforestation and *forest*. This is important to support the results and the methodology, because the same result is obtained from two independent validation models.

The analysis of the *temporal predictors* importance allowed us to identify those variables which were more significant to separate the three *classification periods*. The median value of the green and the blue bands was part of the main predictors in almost every distinction considered; this probably concerned the fact that the blue band is associated with soil moisture, while the green band is associated with the reflectance value of the vegetation: the different water content and the different reflectance value are therefore two important factors to examine to classify *non-forest*, *afforestation* and *forest*. Considering the other predictors in the importance ranking, the Vegetation Spectral Indices and the Tasseled Cap in the first positions were those ones which are related to the blue band (as TCB, EVI), to the green band (TCG) and, to a lesser extent, those related to nir and red bands, as the Normalized Difference Vegetation Index.

These are important information to set the best combination of parameters to obtain the maximum accuracy in the distinction of *non-forest*, *afforestation* and *forest*, using a small number of input data. The results of the analysis of the *temporal predictors* confirm that the remote sensing data allow to assess the differences in the spectral response in relation to *non-forest*, *afforestation* and *forest*. The median value of blue and green bands, R value of EVI, TCG, TCB and green slope values, which were the most important variables in the random forests important ranking, showed the most significant dissimilarity in the spectral response. This can be due to several reasons, not only to the spectral response of the different vegetation type, but also to their phenology; in fact, a non-forest land cover, as pastures or other grasslands, could be more influenced by the microclimate or the soil moisture than the afforestation or forest land cover.

Similarly, shrublands could be more susceptible to local weather condition than the forest, which is the most stable land cover.

There are few studies on the use of remote sensing to assess afforestation areas, and most of them evaluate the forest gain by comparing different land cover maps. For example, Shen et al. (2019) estimated afforestation areas at a regional scale using annual forest mask and combining them to extract the afforestation areas. Brovelli et al., (2020); and Townshend et al., (2012) assessed forest cover change using machine learning and remote sensing data through the comparison of two forest cover maps. Other research examine the forest gain at global level using remote sensing, comparing forest and non-forest land cover, but they do not consider the afforestation land cover (Hansen *et al.*, 2013; Kim *et al.*, 2014; Shimada *et al.*, 2014).

One of the limits of the method proposed is the medium spatial resolution of images: in fact, Landsat images resolution is 30 meters, and this could cause the omission of the smaller elements. This could be overcome by integrating other data, like radar, lidar, or other multispectral images. Another issue could occur increasing the area of study outside the training areas, in fact it will be necessary to consider potential commission or omission errors which could be due for example to the orography, which could reduce the accuracy because of slope and shadows.

The results of the present research offer the possibility to deepen the existing analysis approaches, directly highlighting the *afforestation*, a land cover class which is not stable, but represent a transition from non-forest to forest cover. Furthermore, it will be possible to elaborate an afforestation areas map, at national or European level. The new cartographic data could satisfy the requirements proposed by European legislation for monitoring activities, being aligned to the EIONET Action Group on Land monitoring in Europe (EAGLE) Concept. EAGLE is based on a classification system which considers the distinction between land cover and land use and it is composed of three descriptors (land cover components, land use attributes and further characteristics), combined to define a classification system able to be integrated to existing classes, maintaining the three components independent (Arnold *et al.*, 2013). The classification structure of the EAGLE matrix specifically includes the afforestation land cover class. Considering this new classification method, the map will be aligned not only to the existing cartographic data, but also to the future steps of the land cover monitoring and classification planned by Europe and Italy, being easily updated in all the descriptors.

4.6. Conclusions

Non-forest, *afforestation* and *forest* can be automatically predicted using Landsat data. The classification models aimed to evaluate the different spectral response between *non-forest, afforestation* and *forest.* Since this approach is suitable for application at different scales of analysis and thanks to the results obtained in this study, the methodology will be applied to the entire Italian national area, to locate the potential afforestation areas outside the training areas. This step will allow to create a map of afforestation areas. This product follows the indications given in article 15 of the Italian national legislation on forestry and forestry supply chain (legislative decree 34/2018), which encourages the creation of a geo-referenced forest mapping and contributes advantageously to forest planning and management activities. In-fact the expansion of forest areas is often due to the agricultural and pasture surfaces loss (RAF, 2018; Munafò, 2018), and this could bring to an uncontrolled vegetation growth, which needs to be located and monitored, to avoid negative consequences, such as fires, that could be more destructive if the vegetation is not managed.

The afforestation areas assessment can be important to analyze the forest growth trend, allowing to make forecasts on the forest evolution, both in quantitative and qualitative terms. Moreover, the integration between the map and other data will allow to study the evolution of the forest areas to assess the forest ecosystem services.

5. Estimating afforestation area using Landsat time series and photointerpreted datasets

5.1. Introduction

Changes between forest and other land use classes include the gain and the loss of tree cover. Afforestation constitutes the gain of arboreal cover and is characterized by the conversion of other land uses to forest, while deforestation constitutes the loss of arboreal cover and is characterized as the conversion of forest to other land uses (FAO, 2020). According to the FAO definition, afforestation can be natural via spontaneous colonization of vegetation or anthropogenic through trees planting or direct sowing on land that was not previously forested. As on one hand, afforestation certainly involves all the benefits connected to forest ecosystem services, on the other hand, if not well managed, it can cause land fragmentation and other related issues (Czimczik et al., 2005, United Nations Framework Convention on Climate Change 2013). In other words, most of the issues and challenges associated with afforestation depend on its location (FAO, 2001).

Remotely sensed data are well-suited for land monitoring applications at both global and local scales, facilitating analysis of numerous elements at fine spatial resolution (Nabuurs *et al.*, 2022) preferably using open access data (Francini, McRoberts, *et al.*, 2022; Wulder and Coops, 2014), cloud computing platforms and Big Data processing (Gorelick *et al.*, 2017).

Afforestation can be analysed using remotely sensed data. An afforestation classification model can be constructed using observations of changes between two dates in both the land cover response variable and changes in remotely sensed predictor variables (McRoberts *et al.*, 2015). This method is based on the analysis of trends in variables or indices such as the Normalized Difference Vegetation Index (NDVI) linked to afforestation during the observation period; the variables or indices are then used to identify and classify the afforestation spectral response (Bollandsås *et al.*, 2013; Fuller *et al.*, 2003; Skowronski *et al.*, 2014). .

A currently popular algorithm for the automated classification of remotely sensed images is random forests (RF, Breiman, 2001), a classification algorithm that uses a series of decision trees based on a sample of data. RF can be used with large datasets because it is not severely adversely affected by overfitting (Belgiu and Drăgu, 2016).

Although forest disturbance monitoring through automated analysis of remotely sensed imagery is a common topic (Coops *et al.*, 2020; Hansen *et al.*, 2013; Kennedy *et al.*, 2010; Laurin *et al.*, 2021; Murillo-Sandoval *et al.*, 2020), afforestation assessment is rarely investigated, being treated as a central subject in only a few studies. Among the studies that consider afforestation prediction, Qiu et al., 2018 used MODIS images (500 m spatial resolution) to map afforestation areas between 2001 and 2016 in China. They proposed the Automatic Method for detecting Multiple vegetation Changes (AMCC) based on five temporal indices. Yin et al., (2018) used RF with MODIS data and a modified version of LandTrendr with Landsat images to monitor the effect of re-vegetation programs between 2000 and 2014 in China, while Ramírez-Cuesta et al., (2021) used MODIS-based NDVI seasonal variables to analyze the main land cover changes that occurred in Europe between 2000 and 2018, focusing on the maximum value of the NDVI that accurately predicts locations where main changes such as afforestation occurred.

Even if these studies have achieved large accuracies, products with a spatial resolution of 500 meters are not useful for monitoring afforestation in the fragmented and morphologically diverse areas typical of Italy and many other Mediterranean countries. Moreover, afforestation processes are characterized by patches of limited size whose detection requires finer resolution images.

Zhu and Woodcock, (2014) illustrated a method for land cover classification and continuous land cover change detection between 1986 and 2010 using Landsat images. Huang et al., (2017) used NDVI to analyze land cover dynamics in China, detecting vegetation gain and loss from 1985 to 2015 and then classifying the land cover in 2015. However, this approach is not useful for our purposes because continuous change detection is not effective for predicting afforestation which is characterized by longer development times which, therefore, requires longer observation periods than other land-cover change approaches. Further, the complex Italian morphology requires a larger number of vegetation indices for accurate classification, not a single index as illustrated in Huang approach. For example, a larger number of vegetation indices is needed to accommodate shaded areas created by mountains which can adversely affect the classification by limiting the vegetation spectral response.

Cartographic data derived from satellite images are therefore effective sources of information for predicting locations of phenomena related to land cover and its changes and for facilitating direct calculation of multiple useful indicators (Luti *et al.*, 2021). However, the variability of forest changes which differ in type, magnitude and definition, can contribute to inaccurate map products and make maps less suitable for direct estimation of the spatial extent of the phenomena under investigation (Chirici *et al.*, 2020). Integration of maps and sample reference data is a useful and potentially unbiased approach for increasing the precision of estimates of the phenomenon of interest. Moreover, using map classes as strata for stratified sampling and estimation can reduce the required sample size and, thus, accelerate the data acquisition and analysis procedures (Marcelli *et al.*, 2020; Olofsson *et al.*, 2013, 2014).

Multiple issues related to mapping and monitoring afforestation require additional attention. First, afforestation prediction using remotely sensed data is seldom investigated. Second, stratified sampling which is based on selection of a simple random sample from within each stratum (Wagner and Stehman, 2015) supports both accuracy assessment and rigorous and unbiased statistical estimation (Olofsson *et al.*, 2013) in support of remote sensing-assisted land cover classification. Third, training data sample sizes needed for calibration of prediction techniques, tend to be very large, but increasing the sample size increases acquisition costs and the time needed for analysis. Fourth stratified random sampling and estimation and can increase accuracies with only moderate sample sizes.

The objectives of this study were two-fold: (1) to develop and illustrate a method that exploits the 1985-2019 Landsat time series for predicting afforestation areas at 30 m resolution at a national scale, and (2) to statistically rigorously estimate afforestation areas within Italian administrative Regions and land elevation classes using a unique validation sample. We consider two forms of afforestation, both natural through spontaneous colonization of vegetation succession and anthropogenic through planting or direct sowing. Because afforestation contributes to increasing the carbon sink, we also estimate the potential carbon (t of C) that will be sequestered once the forest is mature in areas in which afforestation occurred between 1985 and 2019.

The research was constructed as follows: the study area was the whole area of Italy (section 5.2.1). To define forest boundaries, we used the forest mask; to define boundaries of Italian administrative Regions we used the corresponding shapefile and to define elevation class boundaries, we used a Digital Elevation Model (DEM) (section 5.2.2). Because vegetative recovery of forest disturbed areas has a spectral response which could be similar to afforestation and, thereby, decrease model prediction accuracy, and because forest disturbance does not involve any land use change, we considered disturbed forest areas as forest (section 5.2.5). Afforestation areas were mapped using RF in combination with Landsat Best Available Pixel (BAP) composites (section 5.2.3). We used photointerpetations of reference polygons (section 5.2.4) in combination with a corresponding subset of BAP composites as a training dataset to calibrate the RF classification algorithm (section 5.3.2.2) which was then used (section 5.3.2.1) to construct an afforestation map. We then selected a dataset of sample points using stratified random sampling with afforestation-based map classes as strata (section 5.3.3). This sample dataset was used with the map to estimate afforestation area at the national level and for both administrative Regions and elevation classes (section 5.3.4). Finally, we estimated the carbon that will potentially have been sequestered in the afforestation areas (section 5.3.5).

Our results included the afforestation map (section 5.4.1), the accuracy assessment (section 5.4.2), the afforestation estimates obtained from the sample data and the map at the national level and for administrative Regions and elevation classes (section 5.4.3 and appendix D), and the estimate of potential C-sequestration in the afforested areas (section 5.4.4). In the Discussion (section 5.5) and Conclusions (section 5.6) we present the primary results and suggest future research.

5.2. Data

5.2.1. Study area

The study area included the whole of Italy which covers 301,338 km^2 and is characterized by mountains (35%), hills (42%) and plains (23%). The maximum elevation of more than 4,000 m above sea level (asl) is observed in the Alps. The climate varies considerably along the Italian peninsula: alpine in the north, continental in internal areas and Mediterranean along the coasts. Forests occupy almost 35% of the area, about 11 million ha (RAF, 2018). In some administrative Regions, forests cover more than 50% of the territory as in Trentino-Alto Adige, Tuscany, Liguria, Umbria and Sardinia.

Italian forests are characterised by a rich diversity of plant species, in which deciduous broadleaf forests (*Fagus sylvatica, Quercus spp., Castanea sativa*, and others), spruce (*Picea abies*), larch (*Larix decidua*), black pines *(Pinus nigra*) and Mediterranean pines forests predominate. Shrublands cover about one million ha of which Mediterranean shrublands are the most widespread, followed by thermophilic and Alpine shrublands (RAF, 2018).

5.2.2. Forest mask, Italian administrative Regions and Digital Elevation Model

The forest mask of D'amico *et al.*, (2021) was used to distinguish forest from non-forest areas, so that the afforestation process could be analyzed separately inside and outside forest area defined by the forest mask. The analysis inside the forest mask served to identify areas that were not forest and are now forest; the analysis outside the mask, on the other hand, served to identify ongoing afforestation, which is not yet classified by the mask as forest. The mask is based on forest maps for 16 administrative Regions at scales between 1:5,000 and 1:25,000 and was supplemented with Corine Land Cover or Corine Biotopes data for administrative Regions where data were not available. The data were standardised in the reference system, rasterised and reclassified into "forest" and "non-forest" where FAO defines forest as having canopy cover greater than 10% and minimum area of 0.5 ha (FAO, 2001). These operations resulted in a 23x23 meter binary forest/non-forest map with accuracy greater than 85% (Vangi *et al.*, 2021). Moreover, to estimate afforestation areas within administrative Regions and elevation classes, the ISTAT administrative Regions boundary shapefile (ISTAT, 2019) and the TINITALY/01 DEM with a spatial resolution of 10 meters (S. Tarquini *et al.*, 2009; Tarquini *et al.*, 2012) were used, respectively.

5.2.3. Landsat Best Available Pixel (BAP) composite

We used Landsat Best Available Pixel (BAP) composite data as input data to calibrate the RF classification model and to classify afforestation between 1985 and 2019 analyzing the trend of 234 *temporal predictors* (as described in section 5.3.2). The BAP composite used as RF calibration data are composed of a subset of BAP data used to apply the classifier to the study area.

In particular the BAP data associated with the 1,578 training dataset polygons (section 5.2.4) and the *temporal predictors* calculated inside training dataset polygons are part of the RF calibration data, while the rest of the BAP composite and the *temporal predictors* are used as input to the RF classification model applied to the study area.

The BAP procedure was implemented in 2021 in Google Earth Engine (GEE), and the code is openly available[12]. GEE is a cloud platform for analyzing and processing geospatial data which, by exploiting the computational capacity of Google systems, facilitates the study and monitoring of important environmental phenomena such as changes in land cover including afforestation (Gorelick *et al.*, 2017). GEE makes available and enables the integration of numerous image collections and datasets ready to be processed (Francini, McRoberts, *et al.*, 2022). GEE permits multiple programming languages, such as Javascript and Python, so that data processing systems can be constructed that are easily replicable and can be adapted to different situations and needs (Gomes *et al.*, 2020).

GEE provides preprocessed images at different levels, as surface reflectance and brightness orthorectification (the thermal infrared), atmospheric correction, and clouds masking. We used images derived from Landsat 5, 7 and 8 and all 30 m spatial resolution bands, in the visible spectrum and in the infrared spectrum, to construct the Italian annual composites from 1985 to 2019, from 1st June to 31st August, using the BAP procedure (White *et al.*, 2014). The period between June and August is usually characterized by the absence of clouds which facilitates acquisition of many clear images. Moreover, the boreal summer is the season in which the photosynthetic activity is greater, allowing to obtain clearer spectral response. With BAP, it is possible to obtain a final image that is a composite of the best pixel reflectance values where the best pixel is selected using four criteria: (i) sensor score to penalize Landsat 7 images where the Scan Line Corrector malfunction (SLC-off) is present, (ii) target day score, to preferably select the images acquired close to a defined acquisition day (in this case 15th of August), (iii) distance to cloud/cloud shadow score, to decrease the scores of those pixels which are in the proximity the cloud cover, (iv) opacity score (calculated using opacity band produced by LEDAPS), to prefer the pixel with low opacity. Sensor scores and target day are applied to the whole image, while the other two scores are applied to each pixel.

[12] https://code.earthengine.google.com/?accept_repo=users/sfrancini/bap

The BAP application is described on GitHub[13]. For more details on the BAP algorithm see (Griffiths *et al.*, 2013; Hermosilla *et al.*, 2015a, 2015b; White *et al.*, 2014, 2017).

5.2.4. Training dataset

The training dataset (Figure 32) was used to calculate the *temporal predictors* together with the associated BAP composite to calibrate the random forests classification model, as described in section 5.3.2. The training dataset includes information for 1,578 selected training polygons with sizes ranging from 0.09 ha, the Landsat pixel resolution, to 60 ha, distributed as follows:

- 526 polygons (A) that experienced a change from *non-forest* to *forest* between 1985 and 2019,
- 526 polygons (B) in *non-forest* areas that did not change between 1985 and 2019
- 526 polygons (C) in *forest* areas that did not change between 1985 and 2019.

The training dataset was constructed using data from the Land Use Inventory of Italy (IUTI, Sallustio et al., 2016), which was conducted as part of the Extraordinary Plan of Environmental Remote Sensing by the Italian Ministry of Environment and Protection of Land and Sea in 1990 and was updated between 2000 and 2016. The inventory consists of 1,217,032 randomly selected points that were photointerpreted using a Minimum Mapping Unit (MMU) of 0.5 ha and a Minimum Mapping Width (MMW) of 20 m (RSE < 3%). The classification system considers different land use categories including forest based on FAO definitions (Marchetti *et al.*, 2012a). The 30-year historical dataset allows identification of areas where afforestation has occurred. We used IUTI data which identifies afforestation and then we added other training data through photointerpretation of composite of BAP (1985-1988), aerial imagery (1988-2012, spatial resolution 50x50 cm) and very fine-resolution images (2012-2019, spatial resolution 30x30 cm).

The training data were selected to represent the diversity of the national Italian forest area and include other cartographic data such as Corine Land Cover, both conifers and broadleaf forest types and multiple environmental conditions. Photointerpretation of aerial images and very fine resolution images for different years facilitated prediction of areas that have been subject to afforestation and areas that have not been affected by such changes but have maintained either stable forest or stable non-forest cover during the study period. The photointerpretation incorporated the FAO minimum canopy cover percentage of 10% when attributing the forest class (C) to a polygon. A square-meshed grid aligned with the Landsat pixels was constructed so that the photointerpreted polygons corresponded to group of Landsat pixels. Each polygon was drawn following the shape of the afforestation area and classified as either afforestation, forest or non-forest. The afforestation class

[13] https://github.com/saveriofrancini/bap

was assigned to polygons for which afforestation occurred during the study period and included all areas that satisfied the definition of forest including canopy cover >10% at the end of the afforestation process but excluded areas that did not satisfy the definition such as agricultural areas, mainly olive groves and orchards.

In the afforestation class (A), 44% of the polygons were located at altitudes below 1,000 meters asl, a percentage that increased to 55% in the forest class (C) and to 72% in the non-forest class (B).

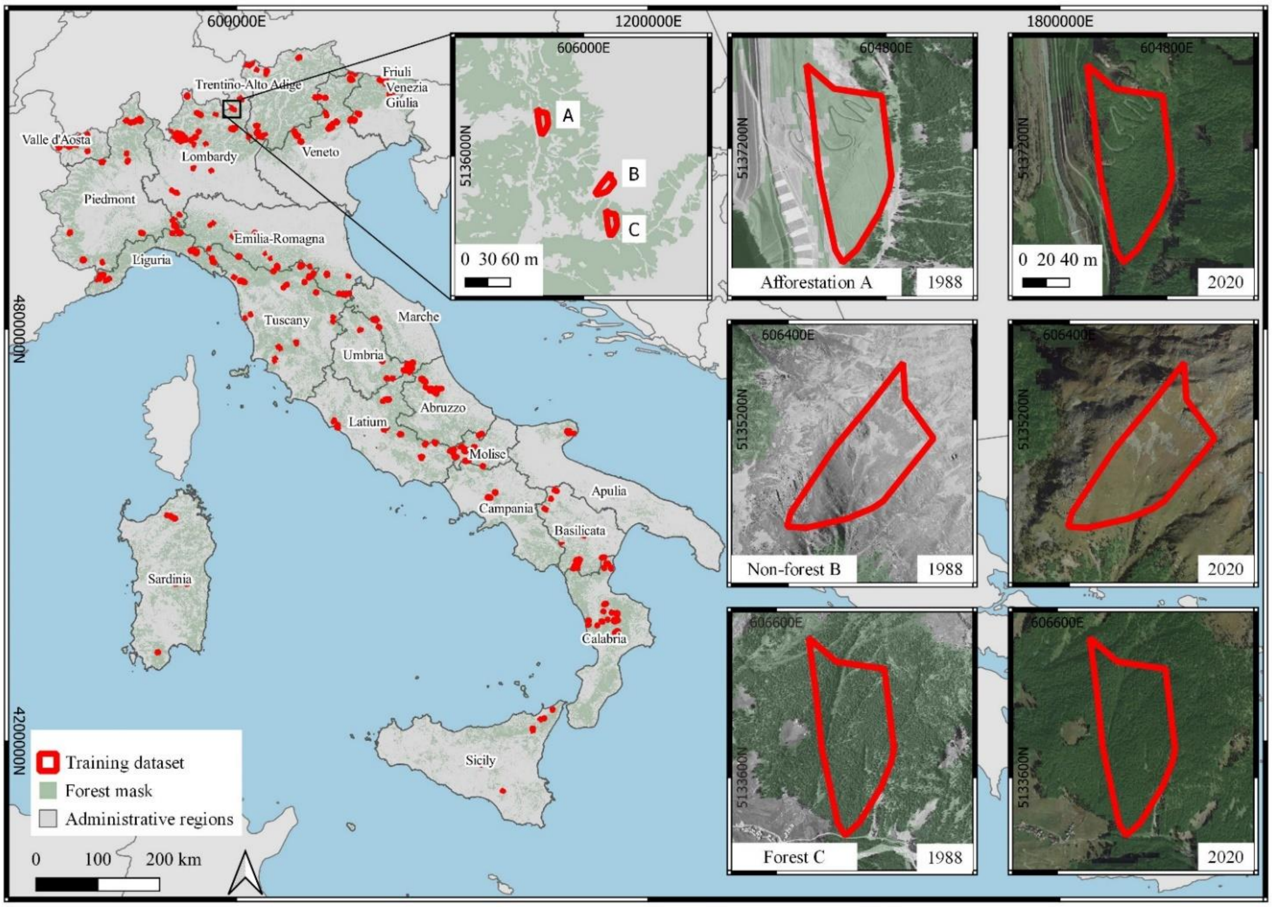

Figure 32 Training dataset. On the right are three examples of the classes considered: afforestation (A), non-forest (B) and forest (C). Polygons A, B and C have the same shape because for better calibrate RF it is important to use training areas with similar area. The figure also illustrates the Italian administrative Regions boundaries. CRS: WGS84/UTM Zone 32N.

5.2.5. Forest disturbance data

Because vegetative recovery of disturbed forest areas (both anthropogenic and natural) that does not involve a conversion of land use is not considered afforestation, we considered disturbed forest areas as forest. The identification of disturbed forest areas was necessary because spectral recovery for these forest disturbances may differ from spectral trends for afforestation so that failure to distinguish them for afforestation can adversely influence and decrease the accuracy of the afforestation classification model. To predict forest disturbances that occurred during the study period, we used the forest disturbance map constructed with the unsupervised algorithm 3I3D algorithm (Francini et al., 2022; Francini et al., 2021). 3I3D was recently implemented on GEE (Francini, McRoberts, *et al.*, 2022), where forest disturbances that occurred in Italy between 1985 and 2019 were mapped using Landsat data, with an OA=99.79% (Francini et al., 2022). Forest disturbances predicted by 3I3D include all changes characterized by relevant decreases in the photosynthetic activity including clearcuts, forest fires, wind damage, drought, frost, and pest disease (Francini, McRoberts, *et al.*, 2022).

5.3. Methods

5.3.1. Overview

The analyses focused on constructing a map of afforestation that occurred between 1985 and 2019 (section 5.2.1). The BAP composites, the training dataset (sections 5.2.3 and 5.2.4) and RF *temporal predictors* (section 5.3.2.1) were used to calibrate RF classification model and to construct an afforestation map (section 5.3.2). Accuracy was assessed and the map was used to support estimation of the area of afforestation, using an estimation dataset selected using stratified sampling technique (section 5.3.3 and 5.3.4). Finally, the potential C-sequestration that will occur in areas afforested between 1985 and 2019 when the forest in those areas mature was estimated (section 5.3.5).

5.3.2. Afforestation map construction

5.3.2.1. RF temporal predictors

The six bands of the Landsat BAP composites acquired between 1985 and 2019 were augmented with seven vegetation indices (VI): Normalized Difference Vegetation Index NDVI (Rouse, 1974), Normalized Burnt Ratio NBR (Key and Benson, 2006), Enhanced Vegetation Index EVI (Huete, 1999), and Tasseled Cap Brightness B, Wetness W, Greeness G and Angle A (Kauth, 1976). The six bands together with the seven VIs comprised a dataset of 13 Times Series (TS) of 35 years of bands and indices.

Then, 18 temporal statistics were calculated for each TS of each polygon of the training dataset: mean, standard deviation, Kendall correlation (Knight, 1966), 11 percentiles from 0 to 100, mean of the first five years, mean of the last five years year max and year min. The result was 13 TS * 18 temporal statistics = 234 *temporal predictors* which were used with RF, to calibrate the RF classification model, to predict afforestation areas in Italy and to construct the afforestation map.

5.3.2.2. Random forests

For classifying Italy into afforestation and non-afforestation areas, we used RF which was first calibrated using the training dataset and the associated BAP composite using the 234 *temporal predictors* and then applied to Italy using the *temporal predictors* calculated for the rest of the BAP composite. Even though RF algorithm accuracy is similar to accuracies for linear models, k-Nearest Neighbors (Chirici et al., 2016; Tomppo, 1991) and more complex models such as neural networks, it was chosen because it can achieve large accuracies as illustrated in (Nguyen *et al.*, 2020; Thanh Noi and Kappas, 2017) and because it is not excessively influenced by overfitting (Belgiu and Drăgu, 2016; Hawryło *et al.*, 2020).

The RF hyperparameters are parameters that can be calibrated to minimize error; two were calibrated for this study using a random search procedure (Bergstra and Bengio, 2012; Laurin *et al.*, 2021): max features, the maximum number of *temporal predictors* considered and maximum depth, the maximum depth of the tree, the longest path between the root node and the leaf node. In random search we defined a grid of max features (1-52) and max depth (1-40) ranges, from which a subset of 100 combinations were randomly selected. The performance of each combination was assessed using k-fold Cross Validation (CV) with k=5; finally, the k-fold CV Overall Accuracy were calculated, obtaining the max features-max depth combination that allowed to obtain larger accuracy. The number of trees in the algorithm was set to 500 and the minimum number of training sample units for splitting a node was set to one (Laurin *et al.*, 2021).

Finally, RF importance ranking was used to estimate the degree to which predictor variables contributed to increasing the accuracy of the map, to assess the most suitable variables for the classification model and to verify their usefulness. The importance of the variables was calculated using the Mean Decrease of the Gini Index (MDG), which estimates the contribution of the variable to increasing the classification model accuracy. MDG is directly proportional to the contribution of the variable within the algorithm (Nicodemus, 2011).

5.3.3. Selection of the estimation sample

To increase the precision of afforestation area estimates and to increase the objectivity of the statistical inference from a sample (Fattorini, 2014), sample units were concentrated in the areas near the forest/non-forest boundaries where the largest classification errors were expected. To this end, we used the forest/non-forest mask of Italy (D'amico et al., 2021, described in section 5.2.2). It was chosen because it has large accuracy (85%) and because when compared to other forest data (as Corine Land Cover and Corine Biotopes) it was the most accurate (Vangi *et al.*, 2021). We resampled the forest mask to 30 m to be consistent with the afforestation map and to define a buffer of 120 m (four Landsat pixels) on each side of the forest/non-forest boundary to facilitate increasing the sampling intensity in the areas where greater error was expected (Francini, McRoberts, *et al.*, 2022). The buffer distance is related to the afforestation patch dimension whose area mean value is 11,104 m^2 and median is 1,512 m^2. Intersecting the buffer and the two classes of the afforestation map we obtained a map, hereafter designated the afforestation-buffer (AB) map, with four map classes: (i) afforestation inside the forest buffer, (ii) non-afforestation inside the forest buffer, (iii) afforestation outside the forest buffer and (iv) non-afforestation outside the forest buffer.

Stratified random sampling with the four AB map classes as strata was then implemented in three phases.

In the first phase, an initial random sample of points was selected from the four AB map classes inside the forest buffer where greatest numbers of classification errors were expected. A total of 2,000 sample points was selected: 660 points were randomly selected in each of the afforestation and non-afforestation map classes inside the buffer, each representing 33% of the total sample, while 340 points were randomly selected in each of the afforestation and non-afforestation map classes outside the buffer, each representing 17% of the total sample.

These 2,000 points were photointerpreted following a procedure similar to the procedure described for the training dataset using fine resolution images (section 5.2.4) and were used to construct a confusion matrix for the afforestation/non-afforestation classification. Unlike the training dataset, for which the photointerpretation involved polygons, because the classification of BAP composite to construct the afforestation map is pixel based, the photointerpretation of the estimation sample dataset entailed classification of single Landsat pixels as afforestation or non afforestation. The matrix facilitates use of the stratified estimator described in (Francini, McRoberts, *et al.*, 2022) for estimating the afforestation area.

The second phase aimed at increasing the precision of the estimates. A second sample of 2,000 points was selected with new within-class sample sizes proportional to the first phase within-class variances: AB map classes with greater first-phase variances were sampled with greater intensities, i.e., more sample points. The second phase distribution was 194 points in afforestation inside the forest buffer class, 835 points in non-afforestation inside the forest buffer class, 75 points in afforestation outside the forest buffer class and 896 points in non-afforestation outside the forest buffer class. The second phase sample was photointerpreted and the results were merged with the results of the first phase sample photointerpretation, resulting in a total sample of 4,000 points (Table 16 appendix B). As for the first sample, a confusion matrix was constructed.

The third phase augmented the first and second phase samples to facilitate analyses of afforestation for (i) administrative Regions and (ii) elevation classes. To align the administrative Regions boundary shapefile and the DEM raster with the afforestation map, the boundary shapefile was rasterized at 30 meters, while the DEM was resampled at 30 meters and reclassified into 14 elevation classes (200 meter intervals). The 20 administrative Regions and the 14 elevation classes were intersected with the four AB map classes to produce a map with respectively 80 classes and 56 classes. Then the 4,000 points were assigned to their respective map classes obtained from the intersection of the four AB map classes and the administrative Regions and the intersection of the four AB map classes and the elevation classes. The within-class samples were augmented with additional randomly selected points where necessary to obtain at least 30 points per class as a means of assuring reliable estimates of within-class variances.

The original estimation sample dataset of 4,000 points was augmented as follows:

- **(i) administrative Regions sample**: we assigned the 4,000 points to the 80 map classes obtained by intersecting the four AB map classes and the 20 administrative Regions and augmented the samples for within-class estimation datasets that had fewer than 30 points. For this sample, we photointerpreted 254 additional points for a total of 4,254 points.
- **(ii) elevation classes sample**: we assigned the 4,000 points to the 56 map classes obtained by intersecting the four AB map classes and the 14 elevation classes from the DEM and augmented the samples for within-class estimation datasets that had fewer than 30 points. For this sample, we photointerpreted 401 additional points for a total of 4,401 points.

The two separate samples for administrative Regions and elevation classes were used to estimate afforestation areas inside and outside the forest buffer for each administrative Region and each elevation class.

5.3.4. Afforestation area estimation and accuracy assessment

Using the estimation sample of size 4,000, the stratified estimator was used to estimate the national afforestation area and to estimate overall accuracy (OA). Then, we used the stratified estimators to estimate afforestation areas for each Italian administrative Region using the estimation sample of size 2,254 and for each elevation class using the estimation sample of size 4,401. For individual administrative Regions and for the national estimate over administrative regions we used the stratified estimation with intersections of the 20 administrative regions and the four AB map classes as strata. In the same way for individual elevation classes and for the national estimate over elevation classes we used the stratified estimation with intersections of the 14 elevation classes and the four AB map classes as strata.

Following the photointerpretation, similar to that of the first phase (section 5.2.4), the confusion matrices for administrative Regions and elevation classes were constructed and used with the stratified estimator (Table 14). The area of afforestation that occurred between 1985 and 2019, was then estimated. The 95% confidence interval (CI) for the estimated area of each class was calculated following the estimator in (Francini, McRoberts, *et al.*, 2022) as illustrated in Table 14. Moreover, the stratified estimator is unbiased, meaning that the average of estimates over all possible samples of the same size obtained using the same sampling design equals the true value (Francini, McRoberts, *et al.*, 2022; Wagner and Stehman, 2015), although the estimate for any particular sample may deviate from the true value.

National estimates were calculated by applying the stratified estimator to the confusion matrix for the map. For administrative Regions, the intersection of the 20 administrative Regions raster and the four AB map classes served as strata for the stratified estimator, and for the elevation classes, the intersection of the 14 elevation classes from the DEM raster and the four AB map classes from the afforestation map served as strata for the stratified estimator.

Table 14 Stratified estimator, confusion matrix to area estimates and CI width.

Map class	Reference class		Sum	w_j [#1]	$\hat{p}_j$	$w_j * \hat{p}_j$	$\widehat{Var}(\hat{p}_j)$	$w_j^2 * \widehat{Var}(\hat{p}_j)$
	Afforest-ation	Non-afforest-ation						
Afforestation outside buffer	n_{11}	n_{12}	$n_1 = n_{11} + n_{12}$	w_1	$\hat{p}_1 = \frac{n_{11}}{n_1}$	$w_1 * \hat{p}_1$	$\widehat{Var}(\hat{p}_1) = \frac{\hat{p}_1 * (1 - \hat{p}_1)}{n_1}$	$w_1^2 * \widehat{Var}(\hat{p}_1)$
Afforestation inside buffer	n_{21}	n_{22}	$n_2 = n_{21} + n_{22}$	w_2	$\hat{p}_2 = \frac{n_{21}}{n_2}$	$w_2 * \hat{p}_2$	$\widehat{Var}(\hat{p}_2) = \frac{\hat{p}_2 * (1 - \hat{p}_2)}{n_2}$	$w_2^2 * \widehat{Var}(\hat{p}_2)$
Non-afforestation outside buffer	n_{31}	n_{32}	$n_3 = n_{31} + n_{32}$	w_3	$\hat{p}_3 = \frac{n_{31}}{n_3}$	$w_3 * \hat{p}_3$	$\widehat{Var}(\hat{p}_3) = \frac{\hat{p}_3 * (1 - \hat{p}_3)}{n_3}$	$w_3^2 * \widehat{Var}(\hat{p}_3)$
Non-afforestation inside buffer	n_{41}	n_{42}	$n_4 = n_{41} + n_{42}$	w_4	$\hat{p}_4 = \frac{n_{41}}{n_4}$	$w_4 * \hat{p}_4$	$\widehat{Var}(\hat{p}_4) = \frac{\hat{p}_4 * (1 - \hat{p}_4)}{n_4}$	$w_4^2 * \widehat{Var}(\hat{p}_4)$
	n_1	n_2	$n = \sum_j^n n_j$	$\sum_j^4 w_j = 1$		$\hat{p} = \sum_j^4 w_j * \hat{p}_j$ [#2]		$\widehat{Var}(\hat{p}) = \sum_j^4 w_j^2 * \widehat{Var}(\hat{p}_j)$ [#3]

[#1] w_j is the proportion of the map in each afforestation/non-afforestation inside/outside buffer class

[#2] $\hat{p} = \sum_j^4 w_j * \hat{p}_j$ is the estimate of the proportion of the total area in one of the afforestation/non-afforestation inside/outside buffer classes

[#3] $\widehat{Var}(\hat{p}) = \sum_j^4 w_j^2 * \widehat{Var}(\hat{p}_j)$ is the variance of the estimate of the total proportion of the area in one of the afforestation/non-afforestation inside/outside buffer classes (Francini et al., 2022a)

5.3.5. Potential carbon sequestration

To estimate the potential carbon that will be sequestered in areas afforested between 1985 and 2019 once the forest is mature, we used the estimates of total biomass per unit of forest area over all carbon pools (aboveground biomass, belowground biomass, litter, necromass and organic and mineral soil) published by the National Inventory of Forests and Carbon Sinks (INFC, 2005) for individual administrative Regions (RAF, 2018). For each Region, we multiplied the estimated area of afforestation by the INFC estimate of biomass per unit area to obtain an estimate of total biomass once the forest is mature.

To estimate the potential C-sequestration (t), we multiplied the estimate of total biomass by the default carbon fraction in biomass of 0.5 (IPCC, 1997; Vitullo et al., 2008). The 95% confidence interval estimates were also calculated for the carbon estimates at national and administrative Region levels; as we did for the potential C-sequestration estimate, we multiplied the CI of the afforestation estimates (ha) and the total carbon stock (t/ha) given by national forest inventory and then we multiplied the total biomass CI (t) and the default carbon fraction of 0.5. It is important to emphasize that this estimate is affected by increased uncertainty due to the error associated with the INFC per unit area biomass estimates and the assumption of a common biomass to carbon conversion factor.

5.4. Results

5.4.1. Afforestation map

The afforestation map (Figure 33) predicted two classes of land cover: afforestation between 1985 and 2019 and non-afforestation between 1985 and 2019. The afforestation map was intersected with the forest map buffer of 120 m (section 5.3.3) to obtain the four AB map classes used to assess map accuracy and to estimate afforestation areas.

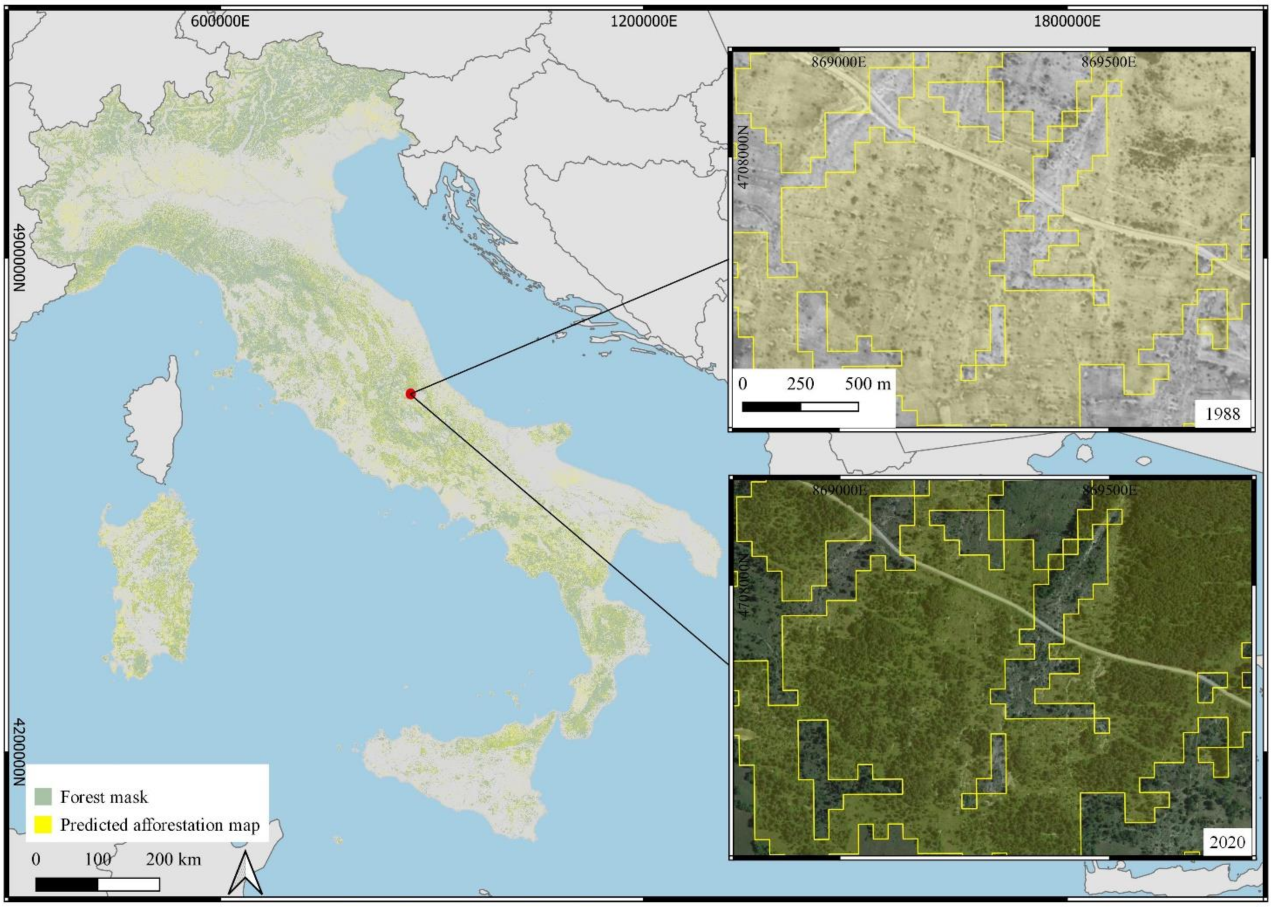

Figure 33. Afforestation map on the left and a highlighted portion and orthophoto on the rights. The figures on the right show the area in 1988 and 2020. The afforestation can be seen in the areas mapped in yellow, while the other areas do not change. CRS: WGS84/UTM Zone 32N.

5.4.2. Accuracy assessment

The estimated map OA was 87% with afforestation classification accuracy of 49% inside the forest buffer class 26% outside the buffer class. Non-afforestation classification accuracy estimates were larger both inside and outside the forest buffer, reaching 90% or greater (appendix C).

5.4.3. Area estimates

Afforestation did not occur uniformly at the national level; in fact, when considering administrative Region boundaries (Figure 34 A-A1), estimates of afforestation areas for four administrative Regions (Abruzzo, Calabria, Latium and Sardinia) were greater than 200,000 ha, with the largest estimate of 260,653 ± 58,522 ha (10.80% of the administrative Region area) for Sardinia. Estimates were less than 90,000 ha for Apulia, Friuli Venezia Giulia, Molise, Valle d'Aosta and Veneto, with the least estimate of 28,644 ± 12,114 ha (8.78% of the administrative Region area) for Valle d'Aosta. In the remaining administrative Regions, the estimated afforestation areas were between 100,000 and 200,000 ha (appendix D). Abruzzo and Liguria had the greatest percentage afforestation estimates relative to the administrative Region area.

The estimated afforestation area for the elevation classes and at national level are illustrated in Figure 34 (B-B1) and appendix D. The intermediate elevation classes are where the largest afforestation area between 1985 and 2019 was estimated. Below 1,200 m asl, estimates of afforested areas were greater than 200,000 ha with the greatest estimate 549,497 ± 84,979 ha (13.94%) for the 400-600 m asl class. For higher altitudes, estimates were less than 100,000 ha, a finding that aligns with the decrease in forest habitats. The greatest percentage afforestation area estimate was ~20% for the 600-1200 m asl class.

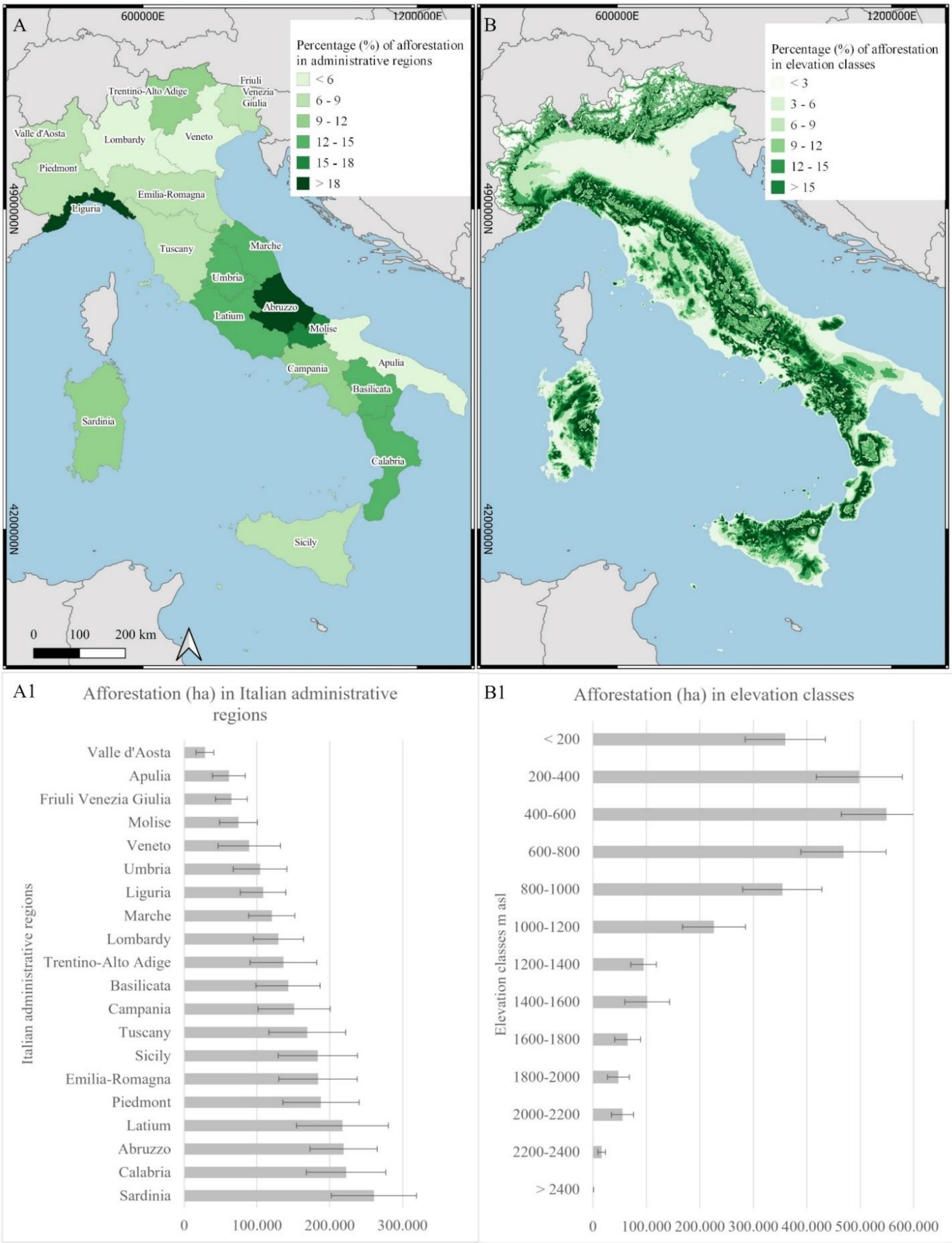

Figure 34 The upper maps above depict the percentage of afforestation (1985-2019) in each Italian administrative Region (A) and in each elevation class (B). The lower graphs report the afforestation area estimates and corresponding 95% confidence intervals (ha) (1985-2019) in Italian administrative Regions (A1) and in each elevation class (B1). CRS: WGS84/UTM Zone 32N.

With the final estimation sample dataset we estimated national afforestation area using the stratified estimator three ways, each with a different set of map classes serving as strata: the four original AB map classes, the 80 classes obtained by intersecting the four AB map classes and the 20 administrative Regions, and the 56 classes obtained by intersecting the four AB map classes and the 14 elevation classes. As expected, because the stratified estimator is unbiased and also because the estimates are based on the same afforestation map and on a common estimation sample, all three national estimates for the period 1985-2019 were very similar, approximately 2.8 ± 0.2 million ha as shown in Table 15.

Table 15 National afforestation area estimates

	Area (ha)	***SE (ha)***	***SE (%)***	***CI width (ha)***	***CI width (%)***
National map	2,833,365	101,125	3.57	202,250	7.14
Elevation classes	2,801,050	94,647	3.38	189,293	6.76
Administrative Regions	2,855,009	98,087	3.43	196,175	6.87

5.4.4. Potential C-sequestration assessment

Considering administrative Regions (Figure 35, Appendix F), the Region with the largest estimate of potential sequestered carbon was Calabria with 17,479,338 ± 4,289,144 t, while the Region with the least estimate of potential carbon sequestrated was Valle d'Aosta with 1,468,005 ± 620,842 t. In 11 of the 20 administrative Regions, estimates were greater than 10 million t of potential carbon sequestrated, and nationwide, 202,420,138 ± 13,908,807 t of potential carbon will be sequestrated when the trees in the afforestation areas mature.

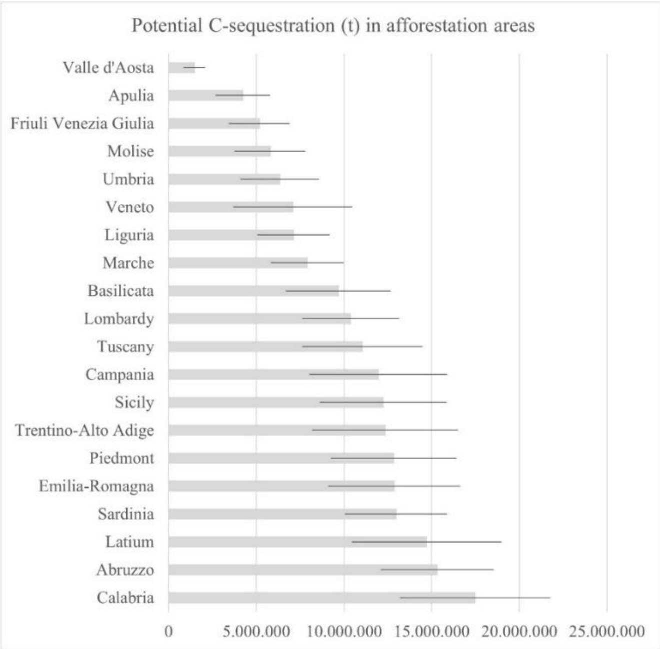

Figure 35 Potential C-sequestration (t) and 95% CI width (t) for administrative Regions in afforestation areas between 1985 and 2019. The value represents the potential carbon that would be sequestrated when the trees in the afforestation areas mature.

5.5. Discussion

Afforestation processes can be natural via spontaneous colonization of vegetation succession or anthropogenic through planting or direct sowing. Afforestation is a profound and potentially long-lasting land use transformation that has implications for ecosystems and their associated services. It is therefore extremely important to monitor afforestation and to locate its occurrence as a means of facilitating development of effective and timely management plans.

5.5.1. Random forests for afforestation map construction

Using the RF model with Landsat BAP composite (spatial resolution 30 meters) data, we constructed a map that predicted areas in Italy where afforestation occurred between 1985 and 2019. Implementation of the classification method on GEE produced the afforestation map for the entirety of Italy in about three hours. The methodology is based on the use of Landsat BAP composites because their long historical series facilitates prediction of slow phenomena such as afforestation. In addition, the use of multiple temporal predictors produced large accuracy estimates. Afforestation classification has focused on defining and estimating the area where afforestation has occurred in the past 35 years, because defining the exact year in which the forest was established involves considering many different variables as the forest type, elevation, latitude, aspect.

Finally, the analysis of the importance of the variables (illustrated in appendix E) used in the RF model allows us to identify the variables that contributed most to the classification, thus reducing the number of variables and speeding up classification operations. A variable with a greater weight in the classification means that it more accurately distinguishes among the afforestation/non-afforestation classes considered.

The most important predictor in the RF model was Kendall’s correlation (tau) of several bands and indices (mean value MDG > 8) and also the average of the values of several bands and indices (mean value MDG > 1). The predictor variables that had the greatest weight in the algorithm were Swir2 tau (MDG =15.22), Swir1 tau (MDG=14.79), TCW and green band tau (MDG =13.60).

5.5.2. Afforestation map validation and accuracy assessment

The validation of the map was carried out in two phases. A qualitative phase involved the visual analysis of the map and identified the major issues in the mapping process such as afforestation of agricultural areas including orchards, agroforestry and olive groves planted during the observation period, or even trees planted in urban areas. However, the visual analysis also made it possible to observe the accuracy with which the map predicted areas of afforestation due to agricultural abandonment or movement of the tree line to higher altitudes, as well as areas of afforestation, which involved many reforestations.

In the second stage of validation, we estimated the OA=87%. The estimates of accuracies for the individual classes were quite large in non-afforestation classes (>95%) and smaller for the afforestation class outside the forest buffer where accuracy was 26% and 50% for the afforestation class inside buffer; this difference is probably due to the fact that outside the forest buffer the classification method was more influenced by the spectral response of other land cover types such as agricultural areas which can have a similar spectral response as afforestation. These results confirm the utility of constructing the 120 m buffer to increase the precision of the estimates. As before, increasing the sampling intensity where smaller accuracy is expected (the non-afforestation classes) allow increasing the precision of the estimates. Accordingly, we concentrated 18% of points in the class with accuracy 26%, 21% in the class with accuracy 49%, and 15% in each of the two classes with accuracies >90% (the two non-afforestation classes).

The map has some systematic commission errors, including agricultural areas, some urban areas and linear elements such as roads. In general, agricultural areas are most prone to these errors. We have observed this feature in other maps and attribute it to crop variability which is subject to periodic cycles of growth. This error is also present in the afforestation inside the forest buffer class which could explain the smaller accuracies for classes inside the buffer compared to those outside the buffer.

Map errors can also be traced back to the input data. Although the Landsat images have the advantage of a long time series, they can be affected by the diversity of the sensors mounted on different satellites which, in turn, can influence the output of the classification. The analysis of accuracy is important for focusing future research on the classes that have smaller accuracies, looking for an effective method to increase it.

5.5.3. Afforestation area estimates

The use of the statistically rigorous, unbiased stratified estimator contributed to more accurate and precise estimates of afforestation areas. The intermediate result based on the first phase sample of 2,000 random points showed that predicted afforestation between 1985 and 2019 covered 3,273,496 ± 401,557 ha. Following the second phase sample which produced a total of 4,000 random points, the estimate of the area subject to afforestation was 2,833,365 ± 202,250 ha. With the final sample, the area of afforestation between 1985 and 2019 was estimated to be 2,801,050 ± 189,293 ha when using the intersection of the 14 elevation classes and four AB map classes as strata, and 2,855,009 ± 196,175 ha when using the intersection of the 20 administrative Regions and the four AB map classes as strata, or an average of ±82,000 ha per year (0.3% of the national area). It is important to note that the increase in the sample size reduced the estimates of the variances and SEs. Overall, all the estimates are reliable, and the SEs are similar and small (from 3.30 to 3.50%), moreover the estimate for any estimate is included in the CIs for the other estimates, e.g., the CI of the national afforestation estimate assessed using elevation classes (189,239 ha) and administrative Regions (196,175 ha) are included in the CI of the national afforestation (202,250 ha). Confidence interval widths as percentage of the estimates were between 6.76% for the elevation classes, 6.87% for the administrative Regions, and 7.14% for the national map.

Afforestation estimates at the national level corresponded well with results from the National Inventory of Forests and Carbon Sinks (INFC) which estimated an annual percentage increase of 0.2-0.3% of the national area, equal to approximately 77,960 ha between 1985 and 2005 and 52,856 ha between 2005-2015 (RAF, 2018). The estimates for this study and the INFC estimates are not completely comparable because the INFC estimates afforestation for forest and other wooded land following the FAO definition, whereas this study focused on the definition of afforestation proposed by (FAO, 2020). Nevertheless, the general comparability of the estimates lends credibility to them.

Studying afforestation within individual administrative Regions and elevation classes was important because it facilitated assessment of afforestation in Italy in which it generally differs, not only in terms of the physical characteristics of the land, but also in terms of forest management systems. The DEM facilitates analyses of afforestation in the different elevation classes, so that it is then possible to reflect on the forests and species most subject to afforestation.

The afforestation area estimates for the elevation classes show that the greatest increase in area has been in deciduous or evergreen broadleaf forests at lower altitudes, while conifers areas have expanded at higher altitudes. Above 1,200 meters, but below the tree line, afforestation has been less, so presumably, the expansion of conifers forests such as *Larix decidua* has been less than in forests at lower altitudes.

The distribution of afforestation in the Italian administrative Regions showed greater estimates in the southern Regions, where an increase of 1,315,280 ha was estimated, while in the north and centre the estimates were 928 ha and 611 ha respectively.

These results can be attributed to agricultural abandonment which occurs at lower altitudes and in southern Italy which generally have fewer intensive forms of management than in the center-north where farms are more specialized.

In addition to INFC afforestation estimates, another result which is consistent with our afforestation results was reported by Pompei, (2005) for Abruzzo, specifically that between 1954 and 2002 estimated afforestation areas covered almost 18% of the administrative Region area with an annual percentage increase of 0.4%. These estimates correspond to our estimates in which the estimate of afforestation area between 1985 and 2019 in Abruzzo covered 20% of the administrative Region area with an annual percentage increase of 0.6%.

5.5.4. Potential C-sequestration

Afforestation estimates were also used to estimate the potential C-sequestration in areas in which afforestation occurred between 1985 and 2019. Potential C-sequestration refers to the amount of carbon (t) which will be absorbed by trees in the afforestation areas when they are mature. The estimates of the potential carbon sequestered in afforestation areas show a slight decoupling between administrative Region afforestation areas and the amount of carbon sequestered. In fact, while Sardinia had the greatest afforestation area, it was not the administrative Region with the greatest amount of potential carbon sequestered. This is because carbon content in forests depends on different distributions of carbon in different forest pools, a result influenced by both climatic and geographic factors.

Considering the average 50-year age of Italian forests, it is important to note that because afforestation areas are not yet mature forest, our results illustrate the potential C-sequestration, therefore the amount of carbon which a mature forest will sequester. To precisely assess the amount of C-sequestration, it will be important to know the forest age and other elements such as the geographical location, environmental and soil characteristics.

The large uncertainty in the potential C-sequestration estimates is due to the association of several errors arising from the uncertainty in the conversion of biomass to carbon which is used together with afforestation areas to approximately estimate the carbon from biomass values.

5.5.5. Future developments

Afforestation, in addition to being an extremely significant environmental phenomenon, is also a topical issue and it is included in many international and national programs and legislations.

The analysis of afforestation areas can be strengthened by improving the classification to decrease omission and commission errors using additional auxiliary data related to the afforestation process such as aspect, slope, pedology, geographical area. The analysis of afforestation could be also useful to accurately predict the vegetation associations and, therefore, the species that have been the protagonists of afforestation over the last 35 years and establish the year when the afforestation process ended. This could be useful also to define a minimum time range in which afforestation can be estimated with consideration of the different growth rates for different species.

The method of afforestation classification can be the basis for studying afforestation dynamics. Future research can involve the classification method to assess many ecosystem services such as C-sequestration which can be estimated using both the classification method and chronosequences (Alberti *et al.*, 2008). The classification method can be useful to predict where afforestation is occurring in support of *ad hoc* forest management strategies. Since the beginning of the 20th Century until the 1970s, afforestation in Italy has occurred mainly by tree planting. Currently, afforestation in Italy is mostly via spontaneous vegetation succession (Agnoletti *et al.*, 2022), therefore, using the classification method and the analysis of the afforestation spectral response, anthropogenic and natural afforestation can be distinguished. Finally, the classification method facilitates assessment of the rate of afforestation within elevation classes or ecoregions, and to characterize afforestation trends

5.6. Conclusions

In this study, we constructed a 30-meter map of afforestation areas that occurred between 1985 and 2019 using random forests and Landsat composite BAP data and we estimated afforestation area using a statistically rigorous stratified estimator. Three main conclusions were drawn from the study. Firstly, thanks to remotely sensed data with fine spatial resolution and available for long time series, it is possible to predict locations of afforestation areas by analyzing their evolution over time. Secondly, predicted afforestation maps can be integrated with a photointerpretation phase and auxiliary information such as administrative Region boundaries and the DEM to facilitate analyses of territorial differences of the afforestation phenomenon, offering estimates that can be important from a management point of view. Finally, integration of the map with sample data in combination with the statistically rigorous stratified estimators has shown that afforestation in Italy over the last 35 years has been about 2 million ha, associated with 202 million t of potential carbon sequestrated. These results contribute to greater understanding of the evolution of tree cover which is fundamental for monitoring climate change. For these reasons it will be important to strengthen the analysis and to improve the quality of the information that can be extracted from cartographic data in support of decision-making.

Appendix

B. Intermediate results of the sample selection

Table 16 Selection of first and second estimation sample dataset. The first estimation sample dataset composed of 2,000 sample points. Per-class variance and resulting second sample selection. Through this procedure the 2,000 random points were selected in direct proportion to the variance of the first sample of points. At the end, the estimation sample dataset of 4,000 points used for the next steps was extracted.

AB Map class	***Reference class***					
	Afforestation	Non-afforestation	First sample total	% class variances	Second sample total	Final sample
afforestation inside buffer (i)	164	496	660	9.68	194	854
non-afforestation inside buffer (ii)	329	331	660	41.76	835	1495
afforestation outside buffer (iii)	9	331	340	3.75	75	415
non-afforestation outside buffer (iv)	26	314	340	44.80	896	1236
Total	528	1,472	2,000	100	2,000	4,000

C. Accuracy assessment

Table 17 Overall accuracy assessment for map. The table shows the overall and the class accuracies for the map, estimated from the confusion matrix and the sample of 4,000 points.

AB map Classes	*Reference*		*Total*	*Accuracy*	*Weight (Wt)*	*Wt*Acc*
	Afforestation	Non-afforestation				
afforestation inside buffer (i)	418	436	854	0.49	0.11	0.05
non-afforestation inside buffer (ii)	60	1115	1175	0.95	0.30	0.29
afforestation outside buffer (iii)	191	544	735	0.26	0.08	0.02
non-afforestation outside buffer (iv)	16	1220	1236	0.99	0.52	0.51
Overall Accuracy						**0.87**

D. Afforestation estimates

Afforestation estimates in administrative Regions

Table 18 Stratified area estimates for Italian administrative Regions. This table shows the afforestation estimates for each Italian administrative Region in ha and percentage, with the associated standard error and 95% confidence interval width.

Administrative Region	***Afforestation (ha)***	***Afforestation (%)***	***SE (ha)***	***SE (%)***	***CI width (ha)***	***CI width (%)***
Abruzzo	218,839	20.27	23,035	10.53	46,071	21.05
Apulia	61,313	3.17	11,254	18.36	22,509	36.71
Basilicata	142,766	14.29	22,028	15.43	44,057	30.86
Calabria	222,525	14.75	27,302	12.27	54,604	24.54
Campania	151,133	11.11	24,805	16.41	49,610	32.83
Emilia-Romagna	183,743	8.19	26,858	14.62	53,716	29.23
Friuli Venezia Giulia	64,950	8.20	10,864	16.73	21,727	33.45
Latium	217,336	12.63	31,482	14.49	62,964	28.97
Liguria	108,395	20.00	15,639	14.43	31,277	28.85
Lombardy	129,382	5.42	17,173	13.27	34,346	26.55
Marche	120,337	12.82	15,784	13.12	31,569	26.23
Molise	74,489	16.78	12,987	17.44	25,975	34.87
Piedmont	187,841	7.40	26,166	13.93	52,333	27.86
Sardinia	260,653	10.81	29,261	11.23	58,522	22.45
Sicily	183,561	7.14	27,181	14.81	54,361	29.61
Trentino-Alto Adige	136,171	10.01	22,885	16.81	45,771	33.61
Tuscany	169,211	7.36	26,327	15.56	52,653	31.12
Umbria	104,304	12.34	18,450	17.69	36,900	35.38
Valle d'Aosta	28,644	8.78	6,057	21.15	12,114	42.29
Veneto	89,398	4.88	21,365	23.90	42,731	47.80
National	**2,855,009**	**9.53**	**98,087**	**3.44**	**196,175**	**6.87**

Afforestation estimates in elevation classes

Table 19 Stratified area estimates for elevation classes. This table shows the afforestation estimates for each elevation class in ha and percentage, with the associated standard error and 95% confidence interval width

Elevation classes	***Afforestation (ha)***	***Afforestation (%)***	***SE (ha)***	***SE (%)***	***CI width (ha)***	***CI width (%)***
< 200	359,769	3.37	37,499	10.42	74,998	20.85
200-400	498,396	8.20	40,165	8.06	80,330	16.12
400-600	549,497	13.94	42,489	7.73	84,979	15.46
600-800	468,891	17.49	39,854	8.50	79,708	17.00
800-1000	354,437	19.55	37,065	10.46	74,131	20.92
1000-1200	226,546	18.53	29,391	12.97	58,782	25.95
1200-1400	94,722	10.56	11,994	12.66	23,988	25.33
1400-1600	101,565	15.18	20,946	20.62	41,891	41.25
1600-1800	64,805	12.55	12,066	18.62	24,133	37.24
1800-2000	47,693	11.26	10,289	21.57	20,578	43.15
2000-2200	55,295	16.01	10,420	18.84	20,841	37.69
2200-2400	16,226	5.92	3,620	22.31	7,240	44.62
> 2400	474	0.11	441	93.14	883	186.27
National	**2,801,050**	**9.35**	**94,647**	**3.38**	**189,293**	**6.76**

E. Random forests importance ranking

The *temporal predictors* importance ranking for the classification approach are illustrated in Figure 36. Because the predictors were numerous, the mean values of MDG for each band/index and for each predictor are shown in the figure, while the importance rankings for all the temporal predictors are illustrated in Figure 37.

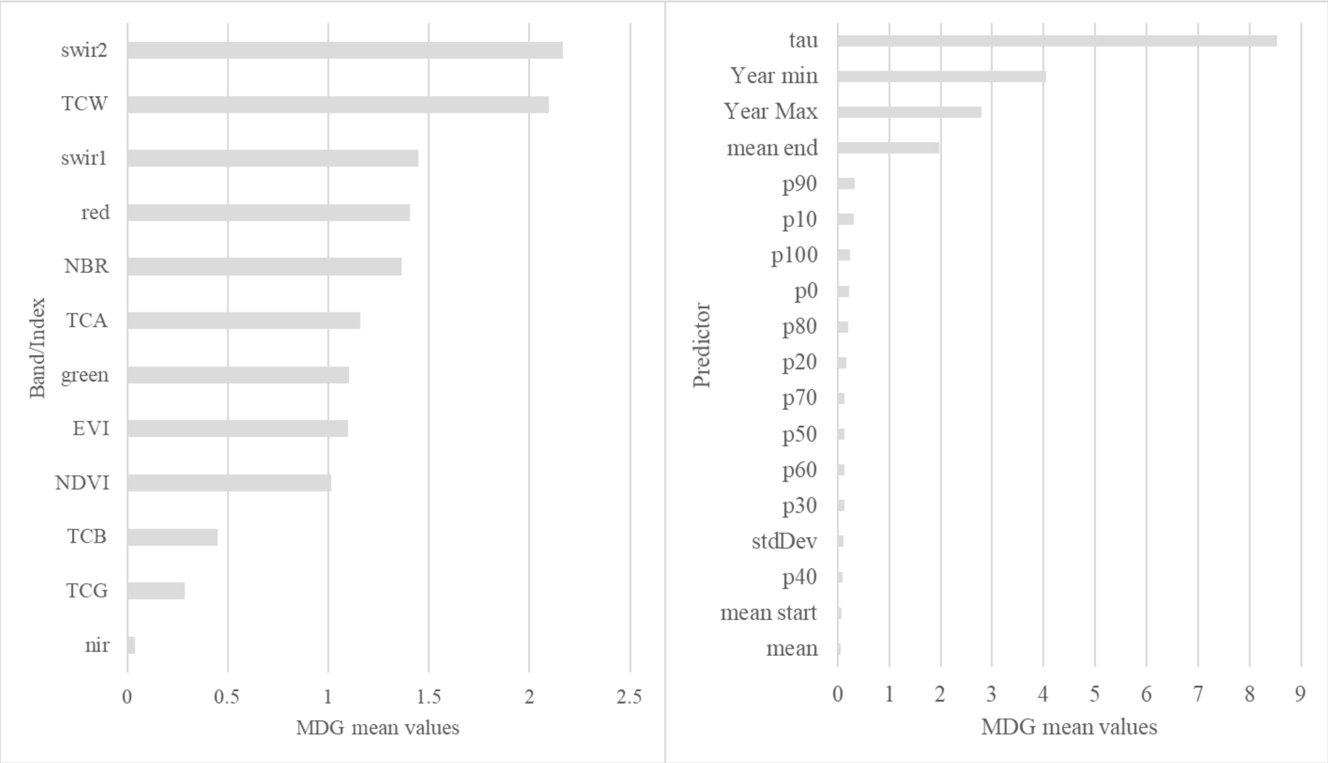

Figure 36 Temporal predictors importance ranking for the classification approach. In the figure are illustrated mean values of the Mean Decrease Gini Index for each band, index and predictor.

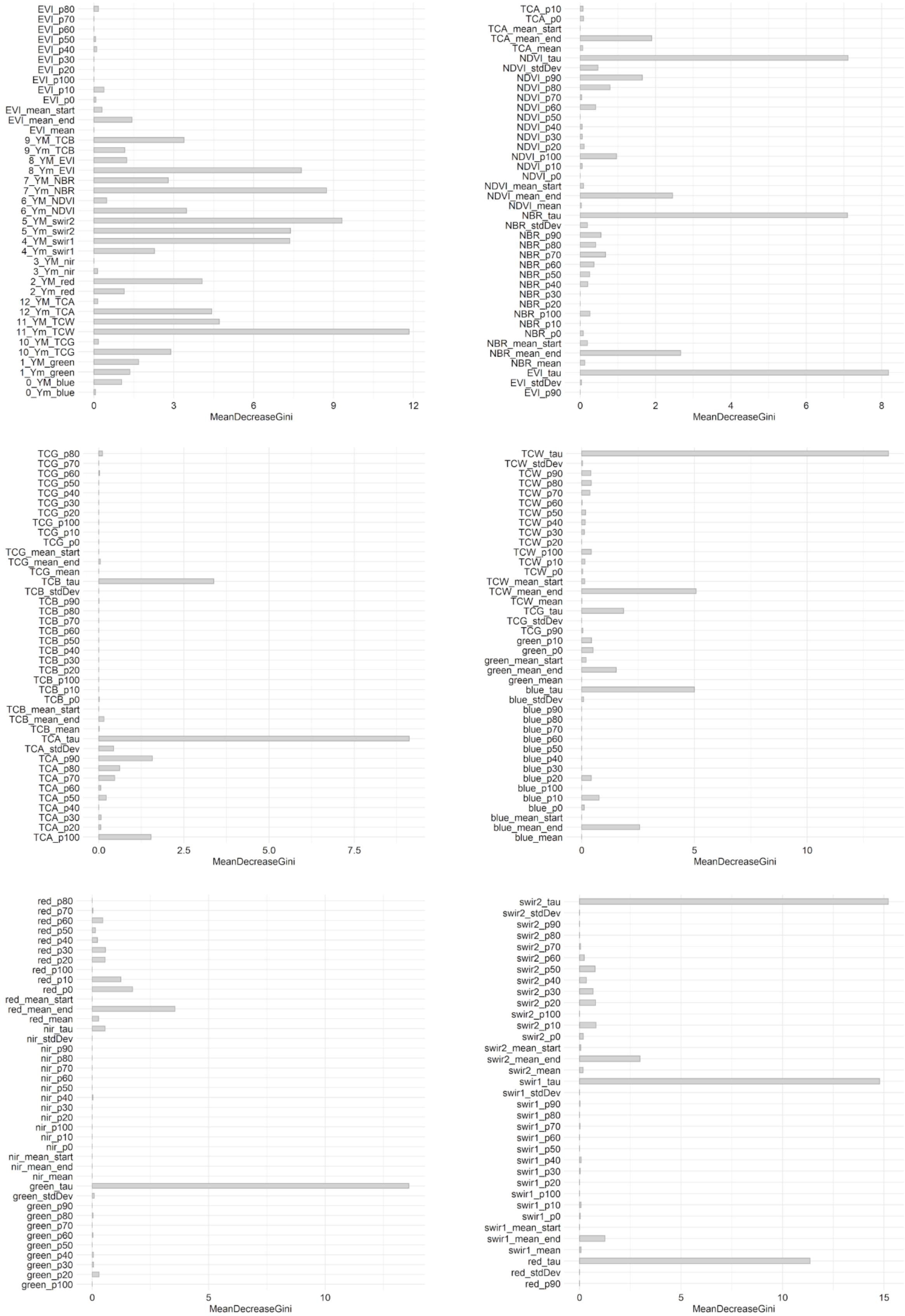

Figure 37 Importance ranking of MDG values of the 234 temporal predictors.

F. Potential C-sequestration estimation

Table 20 Potential C-sequestration (t) and CI width (t) estimated in administrative Regions in afforestation areas between 1985-2019.

Administrative Region	*Afforestation (ha)*	*Afforestation CI width (ha)*	*total biomass stock (t/ha)*	*Afforestation potential biomass (t)*	*Potential C (t)*	*Potential biomass CI width (t)*	*Potential C CI width (t)*
Calabria	222,525	54,604	157.10	34,958,678	17,479,339	8,578,288	4,289,144
Abruzzo	218,839	46,071	140.00	30,637,460	15,318,730	6,449,940	3,224,970
Latium	217,336	62,964	135.40	29,427,294	14,713,647	8,525,326	4,262,663
Sardinia	260,653	58,522	99.50	25,934,974	12,967,487	5,822,939	2,911,470
Emilia-Romagna	183,743	53,716	140.00	25,724,020	12,862,010	7,520,240	3,760,120
Piedmont	187,841	52,333	136.70	25,677,865	12,838,932	7,153,921	3,576,961
Trentino-Alto Adige	136,171	45,771	181.30	24,687,802	12,343,901	8,298,282	4,149,141
Sicily	183,561	54,361	133.30	24,468,681	12,234,341	7,246,321	3,623,161
Campania	151,133	49,610	158.30	23,924,354	11,962,177	7,853,263	3,926,632
Tuscany	169,211	52,653	130.60	22,098,957	11,049,478	6,876,482	3,438,241
Lombardy	129,382	34,346	160.50	20,765,811	10,382,906	5,512,533	2,756,267
Basilicata	142,766	44,057	135.60	19,359,070	9,679,535	5,974,129	2,987,065
Marche	120,337	31,569	131.20	15,788,214	7,894,107	4,141,853	2,070,926
Liguria	108,395	31,277	131.40	14,243,103	7,121,552	4,109,798	2,054,899
Veneto	89,398	42,731	158.70	14,187,463	7,093,731	6,781,410	3,390,705
Umbria	104,304	36,900	121.50	12,672,936	6,336,468	4,483,350	2,241,675
Molise	74,489	25,975	155.60	11,590,488	5,795,244	4,041,710	2,020,855
Friuli Venezia Giulia	64,950	21,727	159.30	10,346,535	5,173,268	3,461,111	1,730,556
Apulia	61,313	22,509	138.20	8,473,457	4,236,728	3,110,744	1,555,372
Valle d'Aosta	28,644	12,114	102.50	2,936,010	1,468,005	1,241,685	620,843
Italy	**2,855,009**	**196,175**	**141.80**	**404,840,276**	**202,420,138**	**27,817,615**	**13,908,808**

6. Intersection of land cover map and afforestation data

To improve the analysis of afforestation one of the future developments illustrated in section 5 was tested. Land cover and afforestation data obtained in the previous phases of the research (section 3 and 5) were intersected to analyze the type of tree cover subject to afforestation between 1985 and 2019. The methods used and the results obtained, are described below.

6.1. Materials and methods

The national land cover map (section 3) has been reclassified considering the two tree cover classes, conifers and broadleaves, obtaining a tree cover map. To be consistent with the afforestation analysis we resampled the tree cover map at 30x30 meters, and we constructed a 120 m buffers (4 pixels) inside and outside the tree cover borders. Then we extracted the tree cover map values with the estimation dataset of 4,000 samples, obtaining an estimation dataset with the information on the vegetation type (conifers and broadleaves).

We then weighed the sample of estimation dataset classified as afforestation/non-afforestation against the total estimation dataset sample considering type of tree cover extracted from the cover map (broadleaves/conifers/unclassified). Finally, using the afforestation/non-afforestation values we calculated with the stratified estimator (section 5), we multiplied them by the weight of the sample points and extracted the afforestation/non-afforestation area divided into conifers and broadleaf, as illustrated in Table 21.

Table 21 Estimation sample dataset broadleaves and conifers weight calculation.

Estimation sample dataset										***Estimation sample dataset weight***					
	broadleaves		conifers		nc					broadleaves		conifers		nc	
AB map classes	non aff	aff	non aff	aff	non aff	aff	total aff	total non aff	total	non aff	aff	non aff	aff	non aff	aff
(i)*afforestation outside buffer*	480	166	20	22	44	3	191	544	735	0.88	0.87	0.04	0.12	0.08	0.02
(ii)*afforestation inside buffer*	400	374	35	44	1	0	418	436	854	0.92	0.89	0.08	0.11	0.00	0.00
(iii)*non afforestation outside buffer*	910	15	60	1	250	0	16	1220	1236	0.75	0.94	0.05	0.06	0.20	0.00
(iv)*non afforestation*	1003	57	87	3	25	0	60	1115	1175	0.90	0.95	0.08	0.05	0.02	0.00

inside buffer													
Total	**2793**	**612**	**202**	**70**	**320**	**3**	**685**	**3315**	**4000**	**0.84**	**0.89**	**0.06**	**0.10**	**0.10**	**0.00**

Afforestation = 2,833,365±202,250 ha; non afforestation = 27,128,594±202,250 ha

6.2. Results and discussion

As illustrated in Table 22 broadleaves were the vegetation type most affected by afforestation between 1985 and 2019, with 2,531,415±180,696 ha (89% of the total afforestation), 289,541±20,668 ha of afforestation can be associated to conifers expansion (10% of the total amount of afforestation). 12,409±886 ha were not classified in conifers and broadleaves, and they are associated both to commission/omission errors of both the input data, which misclassified afforestation or tree cover. These results are consistent with the analysis of afforestation and elevation classes, where most of the afforestation between 1985 and 2019 occurred at low-medium elevation, where broadleaves are more present. This result can be also linked to the agricultural abandonment, which mostly occurred at low altitudes. The conifers expansion could be more common at higher altitudes, also for pasture abandonment.

Table 22 Afforestation estimates and CI (ha) broadleaves and conifers

	broadleaves		*conifers*		*nc*	
AB map classes	non aff	aff	non aff	aff	non aff	aff
(i)*afforestation outside buffer*	23,936,995 ±178,456	2,462,506 ±175,778	997,375 ±7,436	326,356 ±23,296	2,194,225 ±16,358	44,503 ±3,177
(ii)*afforestation inside buffer*	24,888,618 ±185,550	2,535,116 ±180,961	2,177,754 ±16,236	298,249 ±21,289	62,222 ±464	0
(iii)*non afforestation outside buffer*	20,235,263 ± 150,859	2,656,280 ±189,609	1,334,193 ±9,947	177,085 ±12,641	5,559,138 ±41,445	0
(iv)*non afforestation inside buffer*	24,403,569 ±181,934	2,691,697 ±192,138	2,116,760 ±15,781	141,668 ±10,113	608,264 ±4,535	0
Total	**22,856,761 ±170,403**	**2,531,415 ±180,696**	**1,653,085 ±12,324**	**289,541 ±20,668**	**2,618,748 ±19,523**	**12,409 ±886**

7. Conclusion and Future Perspectives

The research has been focused on the analysis on the land cover and its changes, in particular afforestation. It has been part of a research project which involved not only University of Tuscia but also ISPRA, University of Florence, University of Molise and University of Minnesota. The main objectives were: (i) to develop a land cover and land cover change map of Italy using remotely sensed data, (ii) to develop and verify a new methodology to map afforestation occurred in the last 35 years using remotely sensed data, (iii) to construct a map of afforestation occurred between 1985 and 2019 in Italy, (iv) to assess the afforestation areas using statistically rigorous procedures, (v) to estimate the potential carbon sequestration in afforestation areas, (vi) to define the type of tree vegetation most affected by afforestation between 1985 and 2019. Here the conclusion and the future perspectives of the research are resumed.

In the first part of the research (section 3) the national land cover map produced using the supervised classification methods based on decision rules achieved a fine spatial resolution (10 m) and a large accuracy (83%, and 87% for broadleaves and 90% for conifers). The nationwide map can be used for operational purposes and it can be frequently updated, responding to the needs of Mirror Copernicus initiative. The decision rule approach allows to control the choice of the variables used for the classification, ad it can be modified by changing variable thresholds in relation to the characteristic of the area (i.e. elevation or latitude) or it can be integrated with other rules to identify other land cover classes. As regard the tree cover classes (both conifers and broadleaves) a preliminary research was conducted to study the NDVI trend to distinguish tree cover from other land cover; the results are illustrated in (Spadoni *et al.*, 2020): in this research we defined three training areas (three Sentinel-2 tiles) and we analysed 18 months of images. Then we assessed some thresholds values of NDVI, considering the Dinamic Thresholds Method which allowed us to define phenological parameters as Start of Season (SOS) and End of Season (EOS), and considering precise threshold related to the maximum NDVI value. This information together with multitemporal S2 time series were useful to classify broadleaves and conifers distinguishing them from other land cover classes, and thanks to this research we could define more thresholds, considering different parameters to construct the land cover map. The easy implementation of the decision rules approach (both in terms of method and computational cost) is also due to the classification system based on EAGLE, which has a hierarchical structure and in can be integrated or modified maintaining its structure.

In the second part of the research (section 4) we demonstrated that non-forest, afforestation and forest can be automatically predicted using Landsat data. It was a preliminary study which allowed

to apply the methodology to the whole Italy to construct the national afforestation map, to estimate afforestation area between 1985 and 2019 and to assess the potential carbon sequestration in afforestation areas.
The study described in section 5 followed the preliminary research illustrated in section 4. The 30 m afforestation map was constructed, considering the afforestation occurred between 1985 and 2019 and using Landsat composite BAP data. The main conclusions of the research are the following: (i) remotely sensed data with fine spatial resolution and available for long time series allow to predict afforestation areas; (ii) ancillary data as administrative Region boundaries and DEM can be used together with the afforestation map to improve the analysis of territorial differences of afforestation; (iii) the combination of remotely sensed map, sample data and the statistically rigorous stratified estimator allowed to estimate about 2,8 million ha of afforestation in the last 35 years, associated with 202 million t of potential carbon sequestrated (the carbon that will be sequestered in areas afforested between 1985 and 2019 once the forest is mature). These results contribute to explain the afforestation process which is important for monitoring the climate change. Therefore, it will be important to improve the quality of the information useful to decision making process.

Finally, we analyzed the tree cover types which were more interested by afforestation (section 6). Results are consistent with the afforestation analysis in elevation classes, in fact broadleaves are the tree cover which have been more interested by afforestation between 1985 and 2019 (2.5±0.2 million hectares, the 89% of the total afforestation) and they are more common at low and medium altitudes, where the most off afforestation was observed. This operation allowed us not only to successfully integrate two datasets constructed during the PhD studies and to improve the quality of afforestation information, but also to begin to analyze one of the future developments of the afforestation analysis, which aims to know the forest types in expansion.

Future developments

The national land cover map is unique in the panorama of European and national land cover products, being based on an easy replicable method, which allow to construct products maintaining a fine spatial resolution and to frequently update the map. It can be implemented by characterizing specific land cover classes, as the distinction of forest types, agricultural areas and to study other land cover dynamics.
Afforestation is a significant environmental phenomenon and it is included in many national and international programs and legislations. The classification of afforestation can be improved using auxiliary data (i.e. aspect, slope, pedology) which influence the distribution of afforestation, to decrease omission and commission errors. It could also be improved studying the vegetation

associations and the species which are involved in the afforestation process, as we started to study integrating the land cover map and the afforestation data. Moreover, it will be important to define the year when afforestation ended; this could be important to define a minimum time range of afforestation estimation, also considering the species growth rate.

The method presented in this research can be the basis for study afforestation dynamics. It could be used to assess other ecosystem services, and C-sequestration can be estimated combining the classification method and ancillary data as carbon chronosequences (Alberti *et al.*, 2008). Moreover, the classification method could be used to support forest management strategies, suitable for different areas. It could be improved by distinguishing the natural and anthropogenic forest expansion, which is important because in Italy between early 20th Century and 1970s afforestation occurred mainly by tree planting, while currently the spontaneous vegetation succession is more common.

In conclusion, given the importance of land monitoring, which finds application in many International and National programs (SDG, DEF, PNRR, COP 26, EU and national forest strategy, TUFF), this PhD thesis is part of a research which construct many tools useful to the land monitoring, management and planning scopes. Finally, since all the classification methods proposed are suitable to be replicated and integrated, they can be applied in other situations according to different needs.

8. References

Agnoletti, M., Piras, F., Venturi, M. and Santoro, A. (2022), "Cultural values and forest dynamics: The Italian forests in the last 150 years", *Forest Ecology and Management*, Elsevier B.V., Vol. 503, p. 119655.

Agrillo, E., Filipponi, F., Pezzarossa, A., Casella, L., Smiraglia, D., Orasi, A., Attorre, F., *et al.* (2021), "Earth observation and biodiversity big data for forest habitat types classification and mapping", *Remote Sensing*, Vol. 13 No. 7, available at: https://doi.org/10.3390/rs13071231.

Alberti, G., Peressotti, A., Piussi, P. and Zerbi, G. (2008), "Forest ecosystem carbon accumulation during a secondary succession in the Eastern Prealps of Italy", *Forestry*, Vol. 81 No. 1, pp. 1–11.

Arnold, S., Kosztra, B., Banko, G., Smith, G., Hazeu, G. and Bock, M. (2013), "The EAGLE concept – A vision of a future European Land Monitoring Framework", *EARSeL Symposium Proceedings 2013, Towards Horizon 2020*, pp. 551–568.

Bajocco, S., Ferrara, C., Alivernini, A., Bascietto, M. and Ricotta, C. (2019), "Remotely-sensed phenology of Italian forests: Going beyond the species", *International Journal of Applied Earth Observation and Geoinformation*, Elsevier, Vol. 74 No. September 2018, pp. 314–321.

Banskota, A., Kayastha, N., Falkowski, M.J., Wulder, M.A., Froese, R.E. and White, J.C. (2014), "Forest Monitoring Using Landsat Time Series Data: A Review", *Canadian Journal of Remote Sensing*, Vol. 40 No. 5, pp. 362–384.

Belgiu, M. and Drăgu, L. (2016), "Random forest in remote sensing: A review of applications and future directions", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 114, pp. 24–31.

Berger, M. and Aschbacher, J. (2012), "Preface: The Sentinel missions-new opportunities for science", *Remote Sensing of Environment*, Elsevier Inc., Vol. 120 No. 2012, pp. 1–2.

Bergstra, J. and Bengio, Y. (2012), "Random search for hyper-parameter optimization", *Journal of Machine Learning Research*, Vol. 13, pp. 281–305.

Berhane, T.M., Lane, C.R., Wu, Q., Autrey, B.C., Anenkhonov, O.A., Chepinoga, V. v. and Liu, H. (2018), "Decision-tree, rule-based, and random forest classification of high-resolution multispectral imagery for wetland mapping and inventory", *Remote Sensing*, Vol. 10 No. 4, available at: https://doi.org/10.3390/rs10040580.

Bhattarai, R., Rahimzadeh-Bajgiran, P., Weiskittel, A., Meneghini, A. and MacLean, D.A. (2021), "Spruce budworm tree host species distribution and abundance mapping using multi-temporal Sentinel-1 and

Sentinel-2 satellite imagery", *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier B.V., Vol. 172 No. November 2020, pp. 28–40.

Bollandsås, O.M., Gregoire, T.G., Næsset, E. and Øyen, B.H. (2013), "Detection of biomass change in a Norwegian mountain forest area using small footprint airborne laser scanner data", *Statistical Methods and Applications*, Vol. 22 No. 1, pp. 113–129.

Breiman, L. (2001), Random Forests. Machine Learning 45,5-32.

Brovelli, M.A., Sun, Y. and Yordanov, V. (2020), "Monitoring forest change in the amazon using multi-temporal remote sensing data and machine learning classification on Google Earth Engine", *ISPRS International Journal of Geo-Information*, Vol. 9 No. 10, pp. 1–21.

Buscardo, E., Smith, G.F., Kelly, D.L., Freitas, H., Iremonger, S., Mitchell, F.J.G., O'Donoghue, S., *et al.* (2008), "The early effects of afforestation on biodiversity of grasslands in Ireland", *Biodiversity and Conservation*, Vol. 17 No. 5, pp. 1057–1072.

Cervera, T., Pino, J., Marull, J., Padró, R. and Tello, E. (2019), "Understanding the long-term dynamics of forest transition: From deforestation to afforestation in a Mediterranean landscape (Catalonia, 1868–2005)", *Land Use Policy*, Elsevier Ltd, Vol. 80, pp. 318–331.

Chatziantoniou, A., Petropoulos, G.P. and Psomiadis, E. (2017), "Co-Orbital Sentinel 1 and 2 for LULC mapping with emphasis on wetlands in a mediterranean setting based on machine learning", *Remote Sensing*, Vol. 9 No. 12, available at: https://doi.org/10.3390/rs9121259.

Chen, Y. and Gong, P. (2013), "Clustering based on eigenspace transformation - CBEST for efficient classification", *ISPRS Journal of Photogrammetry and Remote Sensing*, International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS), Vol. 83, pp. 64–80.

Chersich, S., Rejšek, K., Vranová, V., Bordoni, M. and Meisina, C. (2015), "Climate change impacts on the Alpine ecosystem: An overview with focus on the soil - A review", *Journal of Forest Science*, Vol. 61 No. 11, pp. 496–514.

Chirici, G., Giannetti, F., Mazza, E., Francini, S., Travaglini, D., Pegna, R. and White, J.C. (2020), "Monitoring clearcutting and subsequent rapid recovery in Mediterranean coppice forests with Landsat time series", *Annals of Forest Science*, Annals of Forest Science, Vol. 77 No. 2, available at: https://doi.org/10.1007/s13595-020-00936-2.

Chirici, G., Mura, M., McInerney, D., Py, N., Tomppo, E.O., Waser, L.T., Travaglini, D., *et al.* (2016), "A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data", *Remote Sensing of Environment*, Elsevier Inc., Vol. 176, pp. 282–294.

Clerici, N., Valbuena Calderón, C.A. and Posada, J.M. (2017), "Fusion of sentinel-1a and sentinel-2A data for land cover mapping: A case study in the lower Magdalena region, Colombia", *Journal of Maps*, Taylor & Francis, Vol. 13 No. 2, pp. 718–726.

Cochran, W.G. (1977), *Sampling Techniques, 3rd Edition*.

Colditz, R.R. (2015), "An evaluation of different training sample allocation schemes for discrete and continuous land cover classification using decision tree-based algorithms", *Remote Sensing*, Vol. 7 No. 8, pp. 9655–9681.

Colson, D., Petropoulos, G.P. and Ferentinos, K.P. (2018), "Exploring the Potential of Sentinels-1 & 2 of the Copernicus Mission in Support of Rapid and Cost-effective Wildfire Assessment", *International Journal of Applied Earth Observation and Geoinformation*, Vol. 73 No. July, pp. 262–276.

Coluzzi, R., Imbrenda, V., Lanfredi, M. and Simoniello, T. (2018), "A first assessment of the Sentinel-2 Level 1-C cloud mask product to support informed surface analyses", *Remote Sensing of Environment*, Elsevier, Vol. 217 No. June, pp. 426–443.

Coops, N.C., Shang, C., Wulder, M.A., White, J.C. and Hermosilla, T. (2020), "Change in forest condition: Characterizing non-stand replacing disturbances using time series satellite imagery", *Forest Ecology and Management*, Elsevier, Vol. 474 No. June, p. 118370.

Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., *et al.* (1998), "The value of the world's ecosystem services and natural capital", *Ecological Economics*, Vol. 25 No. 1, pp. 3–15.

Czimczik, C.I., Mund, M., Schulze, E.D. and Wirth, C. (2005), "Effects of reforestation, deforestation, and afforestation on carbon storage in soils.", *SEB Experimental Biology Series*, No. June, pp. 319–330.

D'amico, G., Vangi, E., Francini, S., Giannetti, F., Nicolaci, A., Travaglini, D., Massai, L., *et al.* (2021), "Are we ready for a national forest information system? State of the art of forest maps and airborne laser scanning data availability in Italy", *IForest*, Vol. 14 No. 2, pp. 144–154.

Dozier, J. (1989), "Spectral signature of alpine snow cover from the landsat thematic mapper", *Remote Sensing of Environment*, Vol. 28 No. C, pp. 9–22.

Drusch, M., del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., *et al.* (2012), "Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services", *Remote Sensing of Environment*, Vol. 120, pp. 25–36.

Duong, P.C., Trung, T.H., Nasahara, K.N. and Tadono, T. (2018), "JAXA high-resolution land use/land cover map for Central Vietnam in 2007 and 2017", *Remote Sensing*, Vol. 10 No. 9, pp. 1–23.

Dusseux, P., Corpetti, T., Hubert-Moy, L. and Corgne, S. (2014), "Combined use of multi-temporal optical and Radar satellite images for grassland monitoring", *Remote Sensing*, Vol. 6 No. 7, pp. 6163–6182.

Erinjery, J.J., Singh, M. and Kent, R. (2018), "Mapping and assessment of vegetation types in the tropical rainforests of the Western Ghats using multispectral Sentinel-2 and SAR Sentinel-1 satellite imagery", *Remote Sensing of Environment*, Vol. 216 No. July, pp. 345–354.

European Environment Agency. (2020), "The European Environment - State and Outlook 2020", p. 496.

European Parliament, Council of the E.U. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Off. J. Eur. Union 2007, 1–14.

FAO. (2001), "Global Forest Resources Assessment 2000 Main Report. Land use policy", 20, 195, doi:10.1016/s0264-8377(03)00003-6.

FAO. (2016), "Map Accuracy Assessment and Area Estimation : A Practical Guide. Food and Agriculture Organization of the United Nations, Rome", No. 46, p. 69.

FAO. (2020), "Global Forest Resources Assessment 2020 - Guidelines and specifications", *Forest Resources Assessment*, p. 42.

Fassnacht, F.E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L.T., Straub, C., *et al.* (2016), "Review of studies on tree species classification from remotely sensed data", *Remote Sensing of Environment*, Vol. 186, pp. 64–87.

Fattorini, L. (2014), "Design-based methodological advances to support national forest inventories: A review of recent proposals", *IForest*, Vol. 8, pp. 6–11.

Ferretti, F., Alberti, G., Badalamenti, E., Campagnaro, T., Corona, P., Garbarino, M., Mantia, T. la, *et al.* (2019), "Boschi di neoformazione in Italia : approfondimenti conoscitivi e orientamenti gestionali Boschi di neoformazione in Italia : approfondimenti conoscitivi e orientamenti gestionali", pp. 1–33.

Filipponi, F., Valentini, E., Xuan, A.N., Guerra, C.A., Wolf, F., Andrzejak, M. and Taramelli, A. (2018), "Global MODIS fraction of green vegetation cover for monitoring abrupt and gradual vegetation changes", *Remote Sensing*, Vol. 10 No. 4, available at: https://doi.org/10.3390/rs10040653.

De Fioravante, P., Luti, T., Cavalli, A., Giuliani, C., Dichicco, P., Marchetti, M., Chirici, G., *et al.* (2021), "Multispectral sentinel-2 and sar sentinel-1 integration for automatic land cover classification", *Land*, Vol. 10 No. 6, pp. 1–35.

Feranec J et al., (2016). Overview of Changes in Land Use and Land Cover in Eastern Europe. In: Gutman G., Radeloff V. (eds) Land-Cover and Land-Use Changes in Eastern Europe after the Collapse of the Soviet Union in 1991. *Springer*. Pag 13-33.

Flores, A., Herndon, K., Thapa, R.B. and Cherrington, E.A. (2019), "The SAR Handbook : Comprehensive Methodologies for Forest Monitoring and Biomass Estimation The SAR Handbook : Comprehensive Methodologies for Forest Monitoring and Biomass Estimation Take a look and enjoy these free resources developed by the SERVIR Glob", No. September, available at:https://doi.org/10.25966/nr2c-s697.

Foody, G.M. and Mathur, A. (2006), "The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM", *Remote Sensing of Environment*, Vol. 103 No. 2, pp. 179–189.

Formaggio, A.R., Epiphanio, J.C.N. and dos Santos Simões, M. (2001), "Radarsat backscattering from an agricultural scene", *Pesquisa Agropecuaria Brasileira*, Vol. 36 No. 5, pp. 823–830.

Forzieri, G., Girardello, M., Ceccherini, G., Spinoni, J., Feyen, L., Hartmann, H., Beck, P.S.A., *et al.* (2021), "Emergent vulnerability to climate-driven disturbances in European forests", *Nature Communications*, Springer US, Vol. 12 No. 1, pp. 1–12.

Francini, S., Amico, G.D., Vangi, E. and Borghi, C. (2022), "Integrating GEDI and Landsat : Spaceborne Lidar and Four Decades of Optical Imagery for the Analysis of Forest Disturbances and Biomass Changes in Italy".

Francini, S., McRoberts, R.E., D'Amico, G., Coops, N.C., Hermosilla, T., White, J.C., Wulder, M.A., *et al.* (2022), "An open science and open data approach for the statistically robust estimation of forest disturbance areas", *International Journal of Applied Earth Observation and Geoinformation*, Elsevier B.V., Vol. 106 No. December 2021, p. 102663.

Francini, S., McRoberts, R.E., Giannetti, F., Marchetti, M., Scarascia Mugnozza, G. and Chirici, G. (2021), "The Three Indices Three Dimensions (3I3D) algorithm: a new method for forest disturbance mapping and area estimation based on optical remotely sensed imagery", *International Journal of Remote Sensing*, Taylor & Francis, Vol. 42 No. 12, pp. 4697–4715.

Francini, S., McRoberts, R.E., Giannetti, F., Mencucci, M., Marchetti, M., Scarascia Mugnozza, G. and Chirici, G. (2020), "Near-real time forest change detection using PlanetScope imagery", *European Journal of Remote Sensing*, Taylor & Francis, Vol. 53 No. 1, pp. 233–244.

Fuller, R.M., Smith, G.M. and Devereux, B.J. (2003), "The characterisation and measurement of land cover change through remote sensing: Problems in operational applications?", *International Journal of Applied Earth Observation and Geoinformation*, Vol. 4 No. 3, pp. 243–253.

Gálos, B., Hagemann, S., Hänsler, A., Kindermann, G., Rechid, D., Sieck, K., Teichmann, C., *et al.* (2013), "Case study for the assessment of the biogeophysical effects of a potential afforestation in Europe", *Carbon Balance and Management*, Vol. 8 No. 1, pp. 1–12.

Gamfeldt, L., Snäll, T., Bagchi, R., Jonsson, M., Gustafsson, L., Kjellander, P., Ruiz-Jaen, M.C., *et al.* (2013), "Higher levels of multiple ecosystem services are found in forests with more tree species", *Nature Communications*, Vol. 4, available at: https://doi.org/10.1038/ncomms2328.

Gerola F., M., (2006), "Biologia vegetale – sistematica filogenetica". Utet Terza Edizione 712 p.

Giannetti, F., Pegna, R., Francini, S., McRoberts, R.E., Travaglini, D., Marchetti, M., Mugnozza, G.S., *et al.* (2020), "A new method for automated clearcut disturbance detection in mediterranean coppice forests using landsat time series", *Remote Sensing*, Vol. 12 No. 22, pp. 1–23.

Gomes, V.C.F., Queiroz, G.R. and Ferreira, K.R. (2020), "An overview of platforms for big earth observation data management and analysis", *Remote Sensing*, Vol. 12 No. 8, pp. 1–25.

Gómez, C., White, J.C. and Wulder, M.A. (2011), "Characterizing the state and processes of change in a dynamic forest environment using hierarchical spatio-temporal segmentation", *Remote Sensing of Environment*, Elsevier B.V., Vol. 115 No. 7, pp. 1665–1679.

Gómez, C., White, J.C. and Wulder, M.A. (2016), "Optical remotely sensed time series data for land cover classification: A review", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 116, pp. 55–72.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R. (2017), "Google Earth Engine: Planetary-scale geospatial analysis for everyone", *Remote Sensing of Environment*, The Authors, Vol. 202 No. 2016, pp. 18–27.

Griffiths, P., van der Linden, S., Kuemmerle, T. and Hostert, P. (2013), "Erratum: A pixel-based landsat compositing algorithm for large area land cover mapping (IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing)", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 6 No. 5, pp. 2088–2101.

Gulácsi, A. and Kovács, F. (2020), "Sentinel-1-imagery-based high-resolutionwater cover detection on wetlands, aided by google earth engine", *Remote Sensing*, Vol. 12 No. 10, pp. 1–20.

Haines-Young, R. and Potschin, M. (2018), "CICES V5. 1. Guidance on the Application of the Revised Structure", *Fabis Consulting*, No. January, p. 53.

Hansen, M.C., Potapov, P. v., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., *et al.* (2013), “High-resolution global maps of 21st-century forest cover change”, *Science*, Vol. 342 No. 6160, pp. 850–853.

Hawryło, P., Francini, S., Chirici, G., Giannetti, F., Parkitna, K., Krok, G., Mitelsztedt, K., *et al.* (2020), “The use of remotely sensed data and polish NFI plots for prediction of growing stock volume using different predictive methods”, *Remote Sensing*, Vol. 12 No. 20, pp. 1–20.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C. and Hobart, G.W. (2015a), “An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites”, *Remote Sensing of Environment*, Elsevier B.V., Vol. 158, pp. 220–234.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C. and Hobart, G.W. (2015b), “Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics”, *Remote Sensing of Environment*, Elsevier B.V., Vol. 170, pp. 121–132.

Holloway, J., Helmstedt, K.J., Mengersen, K. and Schmidt, M. (2019), “A decision tree approach for spatially interpolating missing land cover data and classifying satellite images”, *Remote Sensing*, Vol. 11 No. 15, available at: https://doi.org/10.3390/rs11151796.

Holtgrave, A.K., Röder, N., Ackermann, A., Erasmi, S. and Kleinschmit, B. (2020), “Comparing Sentinel-1 and -2 data and indices for agricultural land use monitoring”, *Remote Sensing*, Vol. 12 No. 18, available at: https://doi.org/10.3390/RS12182919.

Houhoulis, P.F. and Michener, W.K. (2000), “Detecting wetland change: A rule-based approach using NWI and SPOT-XS data”, *Photogrammetric Engineering and Remote Sensing*, Vol. 66 No. 2, pp. 205–211.

Huang, H., Chen, Y., Clinton, N., Wang, J., Wang, X., Liu, C., Gong, P., *et al.* (2017), “Mapping major land cover dynamics in Beijing using all Landsat images in Google Earth Engine”, *Remote Sensing of Environment*, Elsevier Inc., Vol. 202, pp. 166–176.

Huang, X., Liu, J., Zhu, W., Atzberger, C. and Liu, Q. (2019), “The optimal threshold and vegetation index time series for retrieving crop phenology based on a modified dynamic threshold method”, *Remote Sensing*, Vol. 11 No. 23, available at: https://doi.org/10.3390/rs11232725.

Huete, A.R. (1999), “MODIS VEGETATION INDEX ALGORITHM THEORETICAL BASIS v3”, *Environmental Sciences*, No. Mod 13.

Iannelli, G.C. and Gamba, P. (2018), “Jointly exploiting sentinel-1 and sentinel-2 for urban mapping”, *International Geoscience and Remote Sensing Symposium (IGARSS)*, Vol. 2018-July No. 1, pp. 8209–8212.

Ienco, D., Interdonato, R., Gaetano, R. and Ho Tong Minh, D. (2019), "Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series for land cover mapping via a multi-source deep learning architecture", *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier, Vol. 158 No. February, pp. 11–22.

Inventario Nazionale Delle Foreste e Dei Serbatoi Forestali Di Carbonio. (2005), INFC-2005.

Intergovernal Panel on Climate Change. (1997), Revised 1996 IPCC Guidelines for National Greenhouse Gas Emission Inventories. Three volumes: Reference Manual, Reporting Guidelines and Workbook.

Intergovernmental Panel on Climate Change. (2015), "Agriculture, Forestry and Other Land Use (AFOLU)", *Climate Change 2014 Mitigation of Climate Change*, pp. 811–922.

Intergovernal Panel on Climate Change (2019), "Foreword Technical and Preface", *Climate Change and Land: An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems*, pp. 35–74.

ISTAT. (2019), "Descrizione dei dati geografici dei confini delle unità amministrative a fini statistici", available at: https://www.istat.it/it/files//2013/11/2019.28.06-Descrizione-dei-dati.pdf.

Jönsson, P., Cai, Z., Melaas, E., Friedl, M.A. and Eklundh, L. (2018), "A method for robust estimation of vegetation seasonality from Landsat and Sentinel-2 time series data", *Remote Sensing*, Vol. 10 No. 4, available at: https://doi.org/10.3390/rs10040635.

Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M.R., *et al.* (2016), "A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring", *Remote Sensing*, Vol. 8 No. 1, pp. 1–23.

Karkauskaite, P., Tagesson, T. and Fensholt, R. (2017), "Evaluation of the plant phenology index (PPI), NDVI and EVI for start-of-season trend analysis of the Northern Hemisphere boreal zone", *Remote Sensing*, Vol. 9 No. 5, available at: https://doi.org/10.3390/rs9050485.

Kauth, R.J. (1976), "Tasselled Cap - a Graphic Description of the Spectral-Temporal Development of Agricultural Crops As Seen By Landsat.", pp. 41–51.

Kennedy, R.E., Yang, Z. and Cohen, W.B. (2010), "Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms", *Remote Sensing of Environment*, Elsevier Inc., Vol. 114 No. 12, pp. 2897–2910.

Key, C.H.; Benson, N.C. (1999), The Normalized Burn Ratio (NBR): A Landsat TM Radiometric Measure of Burn Severity; *USA Geological Survey Northern Rocky Mountain Science Center*: Bozeman, MT, USA.

Key, C.H. and Benson, N.C. (2006), “Landscape Assessment (LA) sampling and analysis methods”, *USDA Forest Service - General Technical Report RMRS-GTR*, No. 164 RMRS-GTR.

Khanolkar, R.A., Clark, S.T., Wang, P.W., Hwang, D.M., Yau, Y.C.W., Waters, V.J. and Guttman, D.S. (2020), “Ecological Succession of Polymicrobial Communities in the Cystic Fibrosis Airways”, *MSystems*, Vol. 5 No. 6, pp. 1–16.

Kim, D.H., Sexton, J.O., Noojipady, P., Huang, C., Anand, A., Channan, S., Feng, M., *et al.* (2014), “Global, Landsat-based forest-cover change from 1990 to 2000”, *Remote Sensing of Environment*, Elsevier Inc., Vol. 155, pp. 178–193.

Kleeschulte, S., Banko, G., Smith, G., Arnold, S., Scholz, J., Kosztra, B. and Maucha, G. (2017), “Technical specifications for implementation of a new land-monitoring concept based on EAGLE. D3: Draft design concept and CLC-Backbone, CLC-Core technical specifications, including requirements review. Version 3.0”, pp. 0–78.

Knight, W.R. (1966), A Computer Method for Calculating Kendall’s Tau with Ungrouped Data. J Am Stat Assoc, 61, 436–439, doi:10.1080/01621459.1966.10480879.

Kobayashi, N., Tani, H., Wang, X. and Sonobe, R. (2020), “Crop classification using spectral indices derived from sentinel-2a imagery”, *Journal of Information and Telecommunication*, Taylor & Francis, Vol. 4 No. 1, pp. 67–90.

Koch, B. (1999), “The Contribution of Remote Sensing for Afforestation and Forest Biodiversity Assessment”, No. March.

Kuenzer, C., Ottinger, M., Wegmann, M., Guo, H., Wang, C., Zhang, J., Dech, S., *et al.* (2014), “Earth observation satellite sensors for biodiversity monitoring: potentials and bottlenecks”, *International Journal of Remote Sensing*, Vol. 35 No. 18, pp. 6599–6647.

Kulkarni, A.D. and Lowe, B. (2016), “Random Forest Algorithm for Land Cover Classification”, *International Journal on Recent and Innovation Trends in Computing and Communication*, Vol. 4 No. 3, pp. 58–63.

Lary, D.J., Alavi, A.H., Gandomi, A.H. and Walker, A.L. (2016), “Machine learning in geosciences and remote sensing”, *Geoscience Frontiers*, Elsevier Ltd, Vol. 7 No. 1, pp. 3–10.

Laurin, G.V., Francini, S., Luti, T., Chirici, G., Pirotti, F. and Papale, D. (2021), “Satellite open data to monitor forest damage caused by extreme climate-induced events: A case study of the Vaia storm in Northern Italy”, *Forestry*, Vol. 94 No. 3, pp. 407–416.

Lefebvre, A., Sannier, C. and Corpetti, T. (2016), “Monitoring urban areas with Sentinel-2A data: Application to the update of the Copernicus High Resolution Layer Imperviousness Degree”, *Remote Sensing*, Vol. 8 No. 7, pp. 1–21.

Liu, K., Shi, W. and Zhang, H. (2011), “A fuzzy topology-based maximum likelihood classification”, *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier B.V., Vol. 66 No. 1, pp. 103–114.

Liu, Y., Gong, W., Hu, X. and Gong, J. (2018), “Forest type identification with random forest using Sentinel-1A, Sentinel-2A, multi-temporal Landsat-8 and DEM data”, *Remote Sensing*, Vol. 10 No. 6, pp. 1–25.

Luti, T., de Fioravante, P., Marinosci, I., Strollo, A., Riitano, N., Falanga, V., Mariani, L., *et al.* (2021), “Land consumption monitoring with sar data and multispectral indices”, *Remote Sensing*, Vol. 13 No. 8, pp. 1–26.

Luti, T., Segoni, S., Catani, F., Munafò, M. and Casagli, N. (2020), “Integration of remotely sensed soil sealing data in landslide susceptibility mapping”, *Remote Sensing*, Vol. 12 No. 9, available at: https://doi.org/10.3390/RS12091486.

Malenovský, Z., Rott, H., Cihlar, J., Schaepman, M.E., García-Santos, G., Fernandes, R. and Berger, M. (2012), “Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land”, *Remote Sensing of Environment*, Vol. 120, pp. 91–101.

Mansaray, L.R., Huang, W., Zhang, D., Huang, J. and Li, J. (2017), “Mapping rice fields in urban Shanghai, southeast China, using Sentinel-1A and Landsat 8 datasets”, *Remote Sensing*, Vol. 9 No. 3, available at: https://doi.org/10.3390/rs9030257.

Marcelli, A., Mattioli, W., Puletti, N., Chianucci, F., Gianelle, D., Grotti, M., Chirici, G., *et al.* (2020), “Large-scale two-phase estimation of wood production by poplar plantations exploiting sentinel-2 data as auxiliary information”, *Silva Fennica*, Vol. 54 No. 2, pp. 1–15.

Marchetti, M., Bertani, R., Corona, P. and Valentini, R. (2012a), “Changes of forest coverage and land uses as assessed by the inventory of land uses in Italy”, *Forest@ - Rivista Di Selvicoltura Ed Ecologia Forestale*, Vol. 9 No. 4, pp. 170–184.

Marchetti, M., Bertani, R., Corona, P. and Valentini, R. (2012b), “Changes of forest coverage and land uses as assessed by the inventory of land uses in Italy”, *Forest@ - Rivista Di Selvicoltura Ed Ecologia Forestale*, Vol. 9 No. 4, pp. 170–184.

Maxwell, A.E., Warner, T.A., Vanderbilt, B.C. and Ramezan, C.A. (2017), “Land cover classification and feature extraction from National Agriculture Imagery Program (NAIP) Orthoimagery: A review”, *Photogrammetric Engineering and Remote Sensing*, Vol. 83 No. 11, pp. 737–747.

McFeeters, S.K. (1996), “NDWI BY McFEETERS”, *Remote Sensing of Environment*, Vol. 25 No. 3, pp. 687–711.

McFeeters, S.K. (2013), “Using the normalized difference water index (ndwi) within a geographic information system to detect swimming pools for mosquito abatement: A practical approach”, *Remote Sensing*, Vol. 5 No. 7, pp. 3544–3561.

McRoberts, R.E., Næsset, E., Gobakken, T. and Bollandsås, O.M. (2015), “Indirect and direct estimation of forest biomass change using forest inventory and airborne laser scanning data”, *Remote Sensing of Environment*, Elsevier B.V., Vol. 164, pp. 36–42.

Millennium Ecosystem Assessment, 2005. Ecosystems and Human Well-being: Biodiversity Synthesis. World Resources Institute, Washington, DC.

Millard, K. and Richardson, M. (2015), “On the importance of training data sample selection in Random Forest image classification: A case study in peatland ecosystem mapping”, *Remote Sensing*, Vol. 7 No. 7, pp. 8489–8515.

MIPAAF. (2018), *Strategia Forestale Nazionale*.

Miura, S., Amacher, M., Hofer, T., San-Miguel-Ayanz, J., Ernawati and Thackway, R. (2015), “Protective functions and ecosystem services of global forests in the past quarter-century”, *Forest Ecology and Management*, Elsevier B.V., Vol. 352, pp. 35–46.

Modzelewska, A., Fassnacht, F.E. and Stereńczak, K. (2020), “Tree species identification within an extensive forest area with diverse management regimes using airborne hyperspectral data”, *International Journal of Applied Earth Observation and Geoinformation*, Vol. 84 No. April 2019, available at: https://doi.org/10.1016/j.jag.2019.101960.

Munafò, M. (a cura di). (2018) *Territorio, processi e trasformazioni in Italia. Edizione 2022, Report SNPA, Vol 296/2018*

Munafò, M. (a cura di). (2021), *Consumo Di Suolo, Dinamiche Territoriali e Servizi Ecosistemici. Edizione 2022, Report SNPA*, Vol. 22/21,

Munafò, M. (a cura di). (2022), *Consumo Di Suolo, Dinamiche Territoriali e Servizi Ecosistemici. Edizione 2022, Report SNPA*, Vol. 32/22,

Murillo-Sandoval, P.J., van Dexter, K., van den Hoek, J., Wrathall, D. and Kennedy, R. (2020), “The end of gunpoint conservation: forest disturbance after the Colombian peace agreement”, *Environmental Research Letters*, IOP Publishing, Vol. 15 No. 3, p. 34033.

Muro, J., Varea, A., Strauch, A., Guelmami, A., Fitoka, E., Thonfeld, F., Diekkrüger, B., *et al.* (2020), "Multitemporal optical and radar metrics for wetland mapping at national level in Albania", *Heliyon*, Vol. 6 No. 8, available at: https://doi.org/10.1016/j.heliyon.2020.e04496.

Muys, B. (2021), "Forest Ecosystem Services", No. c, pp. 386–395.

Nabuurs, G.-J., Harris, N., Sheil, D., Palahi, M., Chirici, G., Boissière, M., Fay, C., *et al.* (2022), "Glasgow forest declaration needs new modes of data ownership", *Nature Climate Change*, available at: https://doi.org/10.1038/s41558-022-01343-3.

Ngadze, F., Mpakairi, K.S., Kavhu, B., Ndaimani, H. and Maremba, M.S. (2020), "Exploring the utility of Sentinel-2 MSI and Landsat 8 OLI in burned area mapping for a heterogenous savannah landscape", *PLoS ONE*, Vol. 15 No. 5, pp. 1–13.

Nguyen, H.T.T., Doan, T.M., Tomppo, E. and McRoberts, R.E. (2020), "Land Use / Land Cover Mapping Using Multitemporal Sentinel-2 Imagery and Four Classification", *Remote Sens*, Vol. 12, pp. 1–27.

Nicodemus, K.K. (2011), "Letter to the editor: on the stability and ranking of predictors from random forest variable importance measures.", *Briefings in Bioinformatics*, Vol. 12 No. 4, pp. 369–373.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S. v., Woodcock, C.E. and Wulder, M.A. (2014), "Good practices for estimating area and assessing accuracy of land change", *Remote Sensing of Environment*, Elsevier Inc., Vol. 148, pp. 42–57.

Olofsson, P., Foody, G.M., Stehman, S. v. and Woodcock, C.E. (2013), "Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation", *Remote Sensing of Environment*, Elsevier Inc., Vol. 129, pp. 122–131.

Otukei, J.R. and Blaschke, T. (2010), "Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms", *International Journal of Applied Earth Observation and Geoinformation*, Elsevier B.V., Vol. 12 No. SUPPL. 1, pp. S27–S31.

Palmero-Iniesta, M., Espelta, J.M., Gordillo, J. and Pino, J. (2020), "Changes in forest landscape patterns resulting from recent afforestation in Europe (1990–2012): defragmentation of pre-existing forest versus new patch proliferation", *Annals of Forest Science*, Annals of Forest Science, Vol. 77 No. 2, available at: https://doi.org/10.1007/s13595-020-00946-0.

Pepe, M., Pompilio, L., Gioli, B., Busetto, L. and Boschetti, M. (2020), "Detection and classification of non-photosynthetic vegetation from prisma hyperspectral data in croplands", *Remote Sensing*, Vol. 12 No. 23, pp. 1–12.

Persson, M., Lindberg, E. and Reese, H. (2018), “Tree species classification with multi-temporal Sentinel-2 data”, *Remote Sensing*, Vol. 10 No. 11, pp. 1–17.

Pesaresi, S., Mancini, A., Quattrini, G. and Casavecchia, S. (2020), “Mapping mediterranean forest plant associations and habitats with functional principal component analysis using Landsat 8 NDVI time series”, *Remote Sensing*, Vol. 12 No. 7, available at: https://doi.org/10.3390/rs12071132.

Phiri, D., Simwanda, M., Salekin, S., Ryirenda, V.R., Murayama, Y., Ranagalage, M., Oktaviani, N., *et al.* (2019), “remote sensing Sentinel-2 Data for Land Cover / Use Mapping : A Review”, *Remote Sensing*, Vol. 42 No. 3, p. 14.

Pompei, E. (2005), “Espansione delle foreste italiane negi ultimi 50 anni: il caso della Regione Abruzzo”, p. 162.

Prach, K. and Walker, L.R. (2011), “Four opportunities for studies of ecological succession”, *Trends in Ecology and Evolution*, Elsevier Ltd, Vol. 26 No. 3, pp. 119–123.

Puletti, N. and Bascietto, M. (2019), “Towards a tool for early detection and estimation of forest cuttings by remotely sensed data”, *Land*, Vol. 8 No. 4, available at: https://doi.org/10.3390/land8040058.

Qiu, B., Zou, F., Chen, C., Tang, Z., Zhong, J. and Yan, X. (2018), “Automatic mapping afforestation, cropland reclamation and variations in cropping intensity in central east China during 2001–2016”, *Ecological Indicators*, Elsevier, Vol. 91 No. April, pp. 490–502.

Raha, A.K. (2010), “The Journal of Ecology Ph ton Time Series Analysis of Forest and Tree Cover of West Bengal from 1988 to Ph ton”.

RAF. RaFITALIA 2017-2018. Rapporto sullo stato delle foreste e del settore forestale in Italia. 2019 ISBN 9788898850341.

Ramírez-Cuesta, J.M., Minacapilli, M., Motisi, A., Consoli, S., Intrigliolo, D.S. and Vanella, D. (2021), “Characterization of the main land processes occurring in Europe (2000-2018) through a MODIS NDVI seasonal parameter-based procedure”, *Science of the Total Environment*, Vol. 799, available at: https://doi.org/10.1016/j.scitotenv.2021.149346.

Rast, M., Nieke, J., Adams, J., Isola, C. and Gascon, F. (2021), “Copernicus Hyperspectral Imaging Mission for the Environment (Chime)”, *International Geoscience and Remote Sensing Symposium (IGARSS)*, Vol. 2021-July, pp. 108–111.

Raven, P., H., Evert, R., F., Eichhorn, S. (2002) “Biologia delle piante”. Zanichelli Sesta Edizione, 1024 p.

Reba, M. and Seto, K.C. (2020), “A systematic review and assessment of algorithms to detect, characterize, and monitor urban land change”, *Remote Sensing of Environment*, Elsevier, Vol. 242 No. May 2019, p. 111739.

del Río-Mena, T., Willemen, L., Vrieling, A. and Nelson, A. (2020), “Understanding intra-annual dynamics of ecosystem services using satellite image time series”, *Remote Sensing*, Vol. 12 No. 4, available at: https://doi.org/10.3390/rs12040710.

Rodríguez, E., Arqués, J.L., Rodríguez, R., Nuñez, M., Medina, M., Talarico, T.L., Casas, I.A., *et al.* (1989), “We are IntechOpen , the world ’ s leading publisher of Open Access books Built by scientists , for scientists TOP 1 %”, *Intech*, Vol. 32 No. tourism, pp. 137–144.

Rodriguez-Galiano, V.F., Chica-Olmo, M., Abarca-Hernandez, F., Atkinson, P.M. and Jeganathan, C. (2012), “Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture”, *Remote Sensing of Environment*, Elsevier Inc., Vol. 121, pp. 93–107.

Rouse, J.W., Haas, R.H., Schell, J.A. and Deering, D.W. (1973), “Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation”, *Progress Report RSC 1978-1*.

Rouse, J.W. and Space, G. (2020), “S .876”.

Roy, D.P., Boschetti, L. and Trigg, S.N. (2006), “Remote sensing of fire severity: Assessing the performance of the normalized burn ratio”, *IEEE Geoscience and Remote Sensing Letters*, Vol. 3 No. 1, pp. 112–116.

Rüetschi, M., Small, D. and Waser, L.T. (2019), “Rapid detection of windthrows using Sentinel-1 C-band SAR data”, *Remote Sensing*, Vol. 11 No. 2, pp. 1–23.

S. Tarquini, I. Isola, M. Favalli, F. Mazzarini, M. Bisson, M. T. Pareschi and E. Boschi. (2009), “TINITALY/01: a new Triangular Irregular Network of Italy”, *Annals of Geophysics*, Vol. 50 No. 3, available at: https://doi.org/10.4401/ag-4424.

Sallustio, L., Munafò, M., Riitano, N., Lasserre, B., Fattorini, L. and Marchetti, M. (2016), “Integration of land use and land cover inventories for landscape management and planning in Italy”, *Environmental Monitoring and Assessment*, Vol. 188 No. 1, pp. 1–20.

Shen, W., Li, M., Huang, C., Tao, X., Li, S. and Wei, A. (2019), “Mapping annual forest change due to afforestation in Guangdong Province of China using active and passive remote sensing data”, *Remote Sensing*, Vol. 11 No. 5, pp. 1–21.

Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R. and Lucas, R. (2014), "New global forest/non-forest maps from ALOS PALSAR data (2007-2010)", *Remote Sensing of Environment*, Elsevier Inc., Vol. 155, pp. 13–31.

Sinha, P. and Kumar, L. (2013), "Independent two-step thresholding of binary images in inter-annual land cover change/no-change identification", *ISPRS Journal of Photogrammetry and Remote Sensing*, International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS), Vol. 81, pp. 31–43.

Siyag, P. (n.d.). "Afforestation and Reforestation Projects under the Clean Development Mechanism A Reference Manual".

Skowronski, N.S., Clark, K.L., Gallagher, M., Birdsey, R.A. and Hom, J.L. (2014), "Airborne laser scanner-assisted estimation of aboveground biomass change in a temperate oak-pine forest", *Remote Sensing of Environment*, Elsevier B.V., Vol. 151, pp. 166–174.

Smiraglia, D., Filipponi, F., Mandrone, S., Tornato, A. and Taramelli, A. (2020), "Agreement index for burned area mapping: Integration of multiple spectral indices using Sentinel-2 satellite images", *Remote Sensing*, Vol. 12 No. 11, available at: https://doi.org/10.3390/rs12111862.

Smith, T., M., Smith R., L., (2009), "Elementi di Ecologia", Pearson Sesta edizione, 726 p.

Spadoni, G.L., Cavalli, A., Congedo, L. and Munafò, M. (2020), "Analysis of Normalized Difference Vegetation Index (NDVI) multi-temporal series for the production of forest cartography", *Remote Sensing Applications: Society and Environment*, Vol. 20 No. August, available at: https://doi.org/10.1016/j.rsase.2020.100419.

Spisni, A., Marletto, V. and Botarelli, L. (2012), "Indici vegetazionali da satellite per il monitoraggio in continuo del territorio", *Italian Journal of Agrometeorology*, Vol. 3, pp. 49–55.

Starkey, H.C., Mountjoy, W. and Gardner, J.M. (1977), "REMOVAL OF FLUORINE AND LITHIUM FROM HECTORITE BY SOLUTIONS SPANNING A WIDE RANGE OF pH.", *J Res US Geol Surv*, Vol. 5 No. 2, pp. 235–242.

Stehman, S. v and Czaplewski, R.L. (1998), "Design and Analysis for Thematic Map Accuracy Assessment - an application of satellite imagery", *Remote Sensing of Environment*, Vol. 64 No. January, pp. 331–344.

Stehman, S. v. and Foody, G.M. (2019), "Key issues in rigorous accuracy assessment of land cover products", *Remote Sensing of Environment*, Elsevier, Vol. 231 No. December 2018, p. 111199.

Steinhausen, M.J., Wagner, P.D., Narasimhan, B. and Waske, B. (2018), “Combining Sentinel-1 and Sentinel-2 data for improved land use and land cover mapping of monsoon regions”, *International Journal of Applied Earth Observation and Geoinformation*, Elsevier, Vol. 73 No. April, pp. 595–604.

Strollo, A., Smiraglia, D., Bruno, R., Assennato, F., Congedo, L., de Fioravante, P., Giuliani, C., *et al.* (2020), “Land consumption in Italy”, *Journal of Maps*, Taylor & Francis, Vol. 16 No. 1, pp. 113–123.

Sun, Z., Xu, R., Du, W., Wang, L. and Lu, D. (2019), “High-resolution urban land mapping in China from sentinel 1A/2 imagery based on Google Earth Engine”, *Remote Sensing*, Vol. 11 No. 7, pp. 1–22.

Tarquini, S., Vinci, S., Favalli, M., Doumaz, F., Fornaciai, A. and Nannipieri, L. (2012), “Release of a 10-m-resolution DEM for the Italian territory: Comparison with global-coverage DEMs and anaglyph-mode exploration via the web”, *Computers and Geosciences*, Elsevier, Vol. 38 No. 1, pp. 168–170.

TEEB (2010), The Economics of Ecosystems and Biodiversity Ecological and Economic Foundations. Edited by Pushpam Kumar. Earthscan: London and Washington.

Teodoro, A. and Amaral, A. (2019), “A statistical and spatial analysis of portuguese forest fires in summer 2016 considering landsat 8 and sentinel 2A data”, *Environments - MDPI*, Vol. 6 No. 3, available at: https://doi.org/10.3390/environments6030036.

Thanh Noi, P. and Kappas, M. (2017), “Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery”, *Sensors (Basel, Switzerland)*, Vol. 18 No. 1, available at: https://doi.org/10.3390/s18010018.

Tomppo, E. (1991), Satellite image based national forest inventory of Finland. *International Archives of Photogrammetry and Remote Sensing*. 28(7-1), 419– 424.

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., *et al.* (2012), “GMES Sentinel-1 mission”, *Remote Sensing of Environment*, Elsevier Inc., Vol. 120, pp. 9–24.

Townshend, J.R., Masek, J.G., Huang, C., Vermote, E.F., Gao, F., Channan, S., Sexton, J.O., *et al.* (2012), “Global characterization and monitoring of forest cover using Landsat data: Opportunities and challenges”, *International Journal of Digital Earth*, Vol. 5 No. 5, pp. 373–397.

UNCCD. (2017), *United Nations Convention to Combat Desertification. 2017. The Global Land Outlook, First Edition.*

Vangi, E., D’amico, G., Francini, S., Giannetti, F., Lasserre, B., Marchetti, M. and Chirici, G. (2021), “The new hyperspectral satellite prisma: Imagery for forest types discrimination”, *Sensors (Switzerland)*, Vol. 21 No. 4, pp. 1–19.

Vangi, E., D’amico, G., Francini, S., Giannetti, F., Lasserre, B., Marchetti, M., McRoberts, R.E., *et al.* (2021), “The effect of forest mask quality in the wall-to-wall estimation of growing stock volume”, *Remote Sensing*, MDPI AG, Vol. 13 No. 5, available at: https://doi.org/10.3390/rs13051038.

Veldman, J.W., Overbeck, G.E., Negreiros, D., Mahy, G., le Stradic, S., Fernandes, G.W., Durigan, G., *et al.* (2015), “Where Tree Planting and Forest Expansion are Bad for Biodiversity and Ecosystem Services”, *BioScience*, Vol. 65 No. 10, pp. 1011–1018.

Vitullo, M., de Lauretis, R. and Federici, S. (2008), “La contabilità del carbonio contenuto nelle foreste italiane”, *Silvae*, Vol. 9 No. 3, pp. 91–104.

Wagner, J.E. and Stehman, S. v. (2015), “Optimizing sample size allocation to strata for estimating area and map accuracy”, *Remote Sensing of Environment*, Elsevier Inc., Vol. 168, pp. 126–133.

White, J.C., Wulder, M.A., Hermosilla, T., Coops, N.C. and Hobart, G.W. (2017), “A nationwide annual characterization of 25 years of forest disturbance and recovery for Canada using Landsat time series”, *Remote Sensing of Environment*, Elsevier Inc., Vol. 194, pp. 303–321.

White, J.C., Wulder, M.A., Hobart, G.W., Luther, J.E., Hermosilla, T., Griffiths, P., Coops, N.C., *et al.* (2014), “Pixel-based image compositing for large-area dense time series applications and science”, *Canadian Journal of Remote Sensing*, Vol. 40 No. 3, pp. 192–212.

Whittaker, G., Confesor, R., di Luzio, M. and Arnold, J.G. (2010), “Detection of overparameterization and overfitting in an automatic calibration of SWAT”, *Transactions of the ASABE*, Vol. 53 No. 5, pp. 1487–1499.

Wulder, M.A. and Coops, N.C. (2014), “Satellites: Make Earth observations open access.”, *Nature*, Vol. 513 No. 7516, pp. 30–31.

Wunder, S., Kaphengst, T. and Frelih-Larsen, A. (2018), “Implementing Land Degradation Neutrality (SDG 15.3) at National Level: General Approach, Indicator Selection and Experiences from Germany”, pp. 191–219.

Xin, Q., Li, J., Li, Z., Li, Y. and Zhou, X. (2020), “Evaluations and comparisons of rule-based and machine-learning-based methods to retrieve satellite-based vegetation phenology using MODIS and USA National Phenology Network data”, *International Journal of Applied Earth Observation and Geoinformation*, The Authors, Vol. 93 No. January, p. 102189.

Yin, H., Pflugmacher, D., Li, A., Li, Z. and Hostert, P. (2018), “Land use and land cover change in Inner Mongolia - understanding the effects of China’s re-vegetation programs”, *Remote Sensing of Environment*, Vol. 204 No. December 2016, pp. 918–930.

Zakeri, F. and Mariethoz, G. (2021), “A review of geostatistical simulation models applied to satellite remote sensing: Methods and applications”, *Remote Sensing of Environment*, Elsevier Inc., Vol. 259 No. February, p. 112381.

Zhang, F. and Yang, X. (2020), “Improving land cover classification in an urbanized coastal area by random forests: The role of variable selection”, *Remote Sensing of Environment*, Elsevier, Vol. 251 No. August, p. 112105.

Zhang, X., Schaaf, C.B., Friedl, M.A., Strahler, A.H., Gao, F. and Hodges, J.C.F. (2002), “MODIS tasseled cap transformation and its utility”, *International Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, Vol. 2 No. C, pp. 1063–1065.

Zhou, T., Pan, J., Zhang, P., Wei, S. and Han, T. (2017), “Mapping winter wheat with multi-temporal SAR and optical images in an urban agricultural region”, *Sensors (Switzerland)*, Vol. 17 No. 6, pp. 1–16.

Zhu, Z., Gallant, A.L., Woodcock, C.E., Pengra, B., Olofsson, P., Loveland, T.R., Jin, S., *et al.* (2016), “Optimizing selection of training and auxiliary data for operational land cover classification for the LCMAP initiative”, *ISPRS Journal of Photogrammetry and Remote Sensing*, International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS), Vol. 122, pp. 206–221.

Zhu, Z. and Woodcock, C.E. (2014), “Continuous change detection and classification of land cover using all available Landsat data”, *Remote Sensing of Environment*, Elsevier Inc., Vol. 144, pp. 152–171.

Zimmermann, N.E., Edwards, T.C., Moisen, G.G., Frescino, T.S. and Blackard, J.A. (2007), “Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah”, *Journal of Applied Ecology*, Vol. 44 No. 5, pp. 1057–1067.

9. List of Papers

- Spadoni, G.L., Cavalli, A., Congedo, L. and Munafò, M. (2020), "Analysis of Normalized Difference Vegetation Index (NDVI) multi-temporal series for the production of forest cartography", *Remote Sensing Applications: Society and Environment*, Vol. 20 No. August
- Riitano N., Dichicco P., De Fioravante P., Cavalli A., Falanga V., Giuliani C., Mariani L., Strollo A., Munafò M., (2020) Land consumption in italian coastal areas. *Environmental Engineering and Management Journal.* Vol 19, No 10, pp. 1857-1868.
- ISPRA - SNPA (2019) *Geosfera.* In: Stato dell'Ambiente 89/2020 (ed.) *Annuario dei dati ambientali*, 1-126. (Italian).
- ISPRA - SNPA (2020) *Consumo di suolo, dinamiche territoriali e servizi ecosistemici.* Rapporti 15/2020. (Italian).
- De Fioravante P., Luti T., Cavalli A., Giuliani C., Dichicco P., Marchetti M., Chirici G., Congedo L., Munafò M., (2021) Multispectral Sentinel-2 and SAR Sentinel-1 Integration for Automatic Land Cover Classification. *Land*. Vol. 10 No. 6, pp. 1–35
- C. Borghi, S. Francini, M. Pollastrini, F. Bussotti, D. Travaglini, M. Marchetti, M. Munafò, G. Scarascia-Mugnozza, D. Tonti, M. Ottaviano, C. Giuliani, A. Cavalli, E. Vangi, G. D'Amico, F. Giannetti, G. Chirici, (2021) Monitoring Thirty-Five Years Of Italian Forest Disturbance Using Landsat Time Series, AIT Conference – svolta 13-14 settembre 2021.
- ISPRA - SNPA (2021) *Consumo di suolo, dinamiche territoriali e servizi ecosistemici.* Rapporti 22/2021. (Italian).
- UO Statistica - Open Data Roma Capitale (2021) *Il consumo di suolo di Roma Capitale - Analisi della copertura del suolo nel territorio di Roma*. (Italian)
- Cavalli A., Francini S., Cecili G., Cocozza C., Congedo L., Falanga V., Spadoni G.L., Maesano M., Munafò M., Chirici G., Scarascia Mugnozza G. (2022) A*fforestation monitoring through automatic analysis of 36-years Landsat Best Available Composites*, iForest, vol 15, pp. 220-228
- Cavalli A., Francini S., McRoberts R.E., Falanga V., Congedo L., De Fioravante P., Maesano M., Munafò M., Chirici G., Scarascia Mugnozza G. (2023) *Estimating afforestation area using Landsat time series and photointerpreted datasets*, Remote Sensing, 15,923

- Assennato F., Smiraglia D., Cavalli A., Congedo L., Giuliani C., Riitano N., Strollo A., Munafò M. (2022) *The impact of Urbanization on Land: A Biophysical-Based Assessment of Ecosystem Services Loss Supported by Remote Sensed Indicators"* Land, 11, 236
- ISPRA - SNPA (2022) *Consumo di suolo, dinamiche territoriali e servizi ecosistemici.* Rapporti 32/2022. (Italian).
- UO Statistica - Open Data Roma Capitale (2021) *L'uso e il consumo di suolo di Roma Capitale - Analisi della copertura del suolo nel territorio di Roma*. (Italian)
- ISPRA-SNPA (2022) *Agricoltura e selvicoltura, in Stato dell'Ambiente.* Rapporti 97/2022. (Italian)
- ISPRA-SNPA (2022) *Biosfera, in Stato dell'Ambiente.* Rapporti 97/2022. (Italian)
- ISPRA-SNPA (2022) *Geosfera, in Stato dell'Ambiente.* Rapporti 97/2022. (Italian)
- De Fioravante P., Strollo A., Cavalli A., Cimini A., Smiraglia D., Assennato F., Munafò M., (2023) *Ecosystem mapping and accounting in Italy based on Copernicus and national data through integration of EAGLE and SEEA-EA frameworks.* Land, 12,286.