Land Use Changes in the Alpine Area of Lombardy: A Challenge for AI

Luca Patelli[1,*,†], Paolo Maranzano[2,3,†] and Daniele Toninelli[1,†]

[1]Department of Economics, University of Bergamo, Via dei Caniana 2, Bergamo, 24127, Italy
[2]Department of Economics, Management and Statistics (DEMS), University of Milano-Bicocca, Piazza dell'Ateneo Nuovo n.1, Milano, 20126, Italy
[3]Fondazione Eni Enrico Mattei (FEEM), Corso Magenta n.63, Milano, 20124, Italy

Abstract

The sustainable development of remote and ecologically sensitive regions, such as mountain landscapes, depends also upon land use management. This study examines land use dynamics in the Alpine region of Lombardy, Italy, as delineated by the eco-regions defined by the Italian Statistical Institute (ISTAT). By analyzing land use data from Regional Agency for Services to Agriculture and Forestry (ERSAF), we assess the 2013 and the 2018 status for 19 relevant land use categories, along with the identification of net changes observed during this period. This exploratory work, based on the use of high volume data, lays the foundations for a wider project that aims at modeling spatial autocorrelation, outlining the challenges this estimation would lead to, when managing such a huge volume of data via AI algorithms.

Keywords

Land use, Status and net change, Lombardy (Italy), Eco-regions, Sustainable development, AI modeling.

1. Introduction

Sustainable land use has become a key issue in mitigating the effects and in adapting to the climate change. In particular, this issue is even more crucial in the case of remote and ecologically sensitive areas, such as mountainous regions. These regions, characterized by fragile ecosystems and complex socio-economic dynamics, are particularly vulnerable to the increasing pressure of urbanization and soil loss caused by hydro-geological instability. Consequently, the preservation of biodiversity, the mitigation of land degradation and the promotion of sustainable development require efficient planning for the conservation and management of human activity [1, 2].

Developing such plans and formulating public policies that concurrently support ecological integrity and socio-economic resilience present significant challenges. This underscores the need for targeted research into land use dynamics, a task that often exceeds the capacities of local administrations due to resource constraints. Effectively addressing these challenges requires adopting a spatially nuanced approach to land use management. For this purpose, we need a shift from administrative borders to the concept of eco-regions, also referred to as ecological regions. These are defined as relatively extensive units of land or water that contain a distinct composition of species and natural communities [3]. These kind of boundaries approximate the original extent of natural communities prior to substantial changes in land use. These regions offer a valuable framework for analyzing trends and the impacts of land use changes, identifying areas prioritized for conservation, and facilitating environmental risk assessment. Within the Italian context, the classification of territories is structured within a four-level eco-region system, comprising two divisions (first level), seven provinces (second level), eleven sections (third level), and thirty-three subsections (fourth level). The subdivision of the Italian territory, provided by the Italian Statistical Institute (ISTAT), has been undertaken in accordance to specific combinations of

2nd Workshop "New frontiers in Big Data and Artificial Intelligence" (BDAI 2025), May 29-30, 2025, Aosta, Italy
*Corresponding author.
†These authors contributed equally.
luca.patelli@unibg.it (L. Patelli); paolo.maranzano@unimib.it (P. Maranzano); daniele.toninelli@unibg.it (D. Toninelli)
0000-0001-5536-1047 (L. Patelli); 0000-0002-9228-2759 (P. Maranzano); 0000-0002-3158-1982 (D. Toninelli)

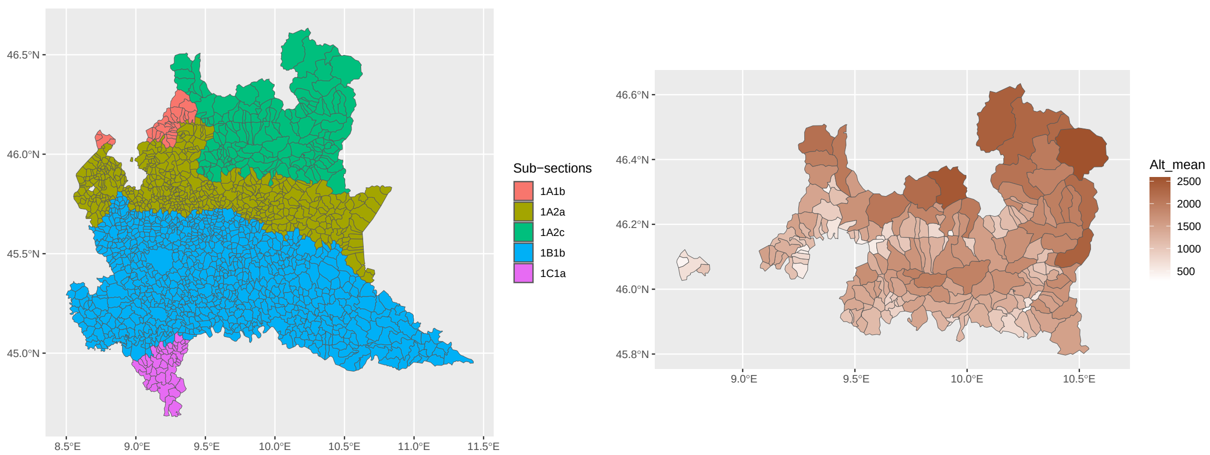
climatic, bio-geographic, physiographic, and hydrographic factors. This framework has been designed aiming to provide a framework to support the design and the implementation of environmental-related policies [4].

The objective of this work is to explore the possible integration of the Italian eco-regions system and the land use data available for Lombardy. We focus on the specific case of the Alpine province for 2013 and 2018. In addition, the challenges associated with the modeling of this type of data are discussed, with a particular focus on AI algorithms implementation.

Following a brief description of the study area, of the data, and of the adopted methodology, we discuss our early results. Finally, future research directions will be outlined.

2. Data and Methods

The study area we analyze encompasses 175 municipalities within the Italian Lombardy region, which have been classified by ISTAT into the Alpine Province. Specifically, at the fourth level of the hierarchy, 26 municipalities are classified as "Sottosezione Alpi Nord-Occidentali" (Northwestern Alps Subsection), while the remaining 149 municipalities belong to the "Sottosezione Alpi Nord-Orientali" (Northeastern Alps Subsection). To comprehend the magnitude of this area, it is necessary to consider that the Lombardy region encompasses an area of approximately 24,000 km^2, and the region covered by our study (the Alpine region) constitutes approximately one-fourth of that size. The geographical location of the study area is illustrated in panel (a) of Figure 1. This depicts the Lombardy region by municipal boundaries, with colors representing the respective subsections. The two Alpine subsections analyzed, the Northwestern (identified by the code 1A1b and color orange) and the Northeastern (identified by the code 1A2c and color green) correspond to the northernmost. Given the recognized role of altitude in delineating alpine provinces, the mean altitude for each municipality is reported in panel (b) of Figure 1. The darker shades of color in the figure correspond to higher mean altitudes. In particular, norther and northeastern municipalities, which represent the Italian border with Switzerland, appear to exhibit higher mean altitudes. We expect that this feature of the area will affect the spatial patterns of the phenomena we are going to study.



(a) Municipalities of Lombardy by ecological subsections.

(b) Mean altitude per municipality of the Alpine subsections.

Figure 1: Study area.

The land use data studied have been extracted from the Regional Agency for Services to Agriculture and Forestry (ERSAF) dataset. This dataset is obtained by combining land use and land cover data with information coming from the Lombardy Region's Agricultural Information System (SIARL) database and from the Intended Use of Agricultural and Forest Land (DUSAF) land cover dataset.

The SIARL database, which is maintained by the Lombardy Region, contains detailed agricultural data updated on an annual basis through agricultural funding procedures. It is based on precise geometries derived from the Italian cadaster, wherein land use is formally assigned by firm declarations for each cadastral parcel.

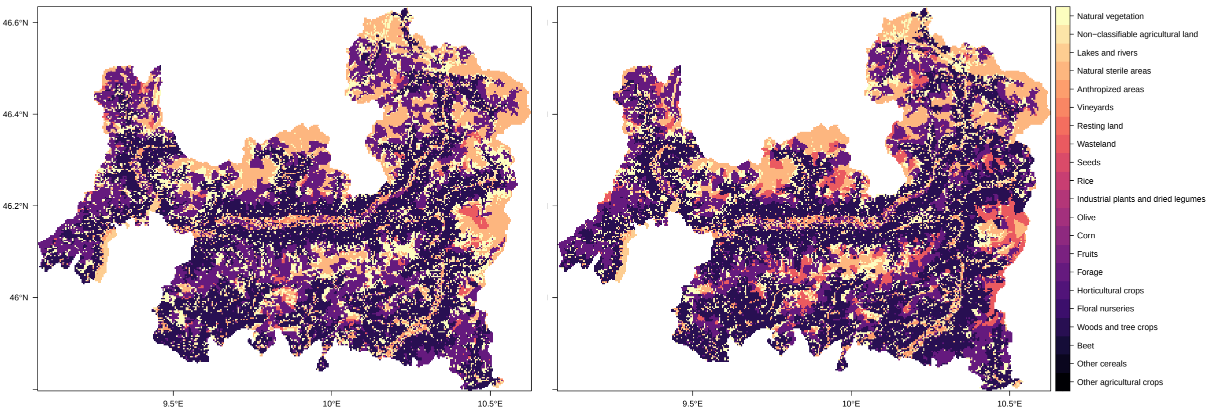
The DUSAF dataset is a land use and land cover classification product obtained from the labeling of aerial photogrammetry. This is available in shape files provided by the Lombardy region.

The objective of combining SIARL and DUSAF data is to map agro-forestry land use across the region. The integration of these data sources aims to provide a comprehensive cartographic representation of agro-forestry surface investments.

The final ERSAF dataset is characterized by the classification of land use into 21 distinct categories. Of these, 16 are derived from SIARL data, representing macro-agricultural land uses of SIARL activities aggregated from multiple sub-categories. The remaining 5 classes are sourced from the DUSAF and, with the exception of the “Non-classifiable agricultural land” label, account for non-agricultural land uses.

The resulting cartographic product represents the dominant land use for each parcel based on declared surface area. The dataset is available in GRID format, with cells measuring 20x20 meters (m). Further methodological details can be found at the link https://www.geoportale.regione.lombardia.it/.

For the purpose of this study, the ERSAF data from 2013 and 2018 were extracted and filtered to obtain only cells covering the Alpine Province of Lombardy. In particular, the number of cells for the two years are 14,441,192 and 14,421,052, respectively. It is noteworthy that the three municipalities not associated with the main Alpine Province were excluded from the analysis as lack of direct spatial adjacency. The resulting raster is shown in Figure 2: left panel (a) refers to 2013, and right panel (b) to 2018. Among the 21 defined land use categories, only 19 have been actually observed in these two years and their labels are reported on the right side of the figure. “Beet” and “Rice” land use categories were not observed in these two years (in the area under study).



(a) 2013.   (b) 2018.

Figure 2: Land use in the two considered years.

Our exploratory analysis involves the estimation of two types of statistics, one focused on the stock and the other on the flow: status and net change, respectively.

The *status* share is used to describe the proportion of total area that has been classified as a particular land use category. The status share s for a specific category k at a certain time point t can be defined as:

$$s_{k,t} = \frac{1}{n} \sum_{(y_{i,t}=k)\in D} y_{i,t}, \tag{1}$$

where $y_{i,t}$ represents the single i-th unit classified in the category k at time t, and n is the number of cells belonging to the total area D.

The *net change* is defined as the difference in the status share for a given land use category k between time points t and $t-c$. This can be computed as:

$$\Delta_{k,t,c} = s_{k,t} - s_{k,t-c}. \tag{2}$$

These two measures are commonly adopted to estimate land use and land cover [5].

3. Results and Discussion

The results pertaining to the status and the net change can be found in Table 1. Based on the status share (expressed as a percentage of the total cells) for 2013 and 2018, the four most prevalent categories, namely, *Woods and tree crops* (35.6 and 40.5 for the two years), *Forage* (27.0 and 23.2), *Natural sterile areas* (13.9 and 11.7) and *Natural vegetation* (12.8 and 9.8), collectively account for more than the 80% of the study area.

The dynamics between 2013 and 2018 can be observed by checking the "Change" section (last three columns of Table 1). In particular, it is reported in net terms, shown by Eq.(2), and, to provide a more comprehensive analysis, in year-to-year relative change (in percentage) and in squared kilometers (km^2) terms. By examining the net change values, it is possible to observe how *Woods and tree crops* and *Wasteland* are the land use that increased the most, in terms of status share (4.9 and 4.3, respectively), while *Forage*, *Natural vegetation* and *Natural sterile areas* are those that decreased the most (-3.8, -2.9 and -2.1, respectively). By considering the year-to-year percentage change, the land use categories subject to the highest positive changes are *Resting land* (+3425.0%), *Industrial plants and dried legumes* (+601.2%) and *Wasteland* (+207.4%). Negative signs are observed for *Seeds* (-67.3%), *Natural vegetation* (-23.2%), and *Natural sterile areas* (-15.6%). Similar conclusions, to those discussed when focusing on net changes, are obtained by also considering the dynamics in squared kilometers terms between 2018 and 2013, allowing for a more understandable unit of measurement. The calculation of this value can be achieved for each land use category by computing the differences between the two years number of cells (# cells) and by subsequently converting the value from squared meters to squared kilometers. Given that each individual cell is 20x20 m, the total number of cells is multiplied by 400 and divided by 10^6. By employing this method, it is possible to ascertain that moving from 2013 to 2018 the areas dedicated to *Forage*, *Natural vegetation* and *Natural sterile areas* showed a loss of 222.34 km^2, 171.99 km^2, and 125.59 km^2, respectively. Concurrently, the land dedicated to *Woods and tree crops* and *Wasteland* increased by 283.31 km^2 and 250.96 km^2, respectively.

In summary, an examination of the statistics presented in Table 1 reveals a decline in naturally vegetated areas (vegetation and sterile areas) and in the land for forage production. Concurrently, there has been an increase in areas allocated to wood and tree crop production, as well as in wasteland areas.

To obtain a general measure of variability in the land use able to quantify the total variation observed by moving from 2013 to 2018, we compute a weighted mean of the absolute relative changes using as weights the number of cells observed for each category in 2013. This statistic, equal to 0.1874, points out that in 5 years the 18.74% of the cells changed their land use (corresponding to an annual average percentage of 3.75%). This information is useful for future comparisons between other time points or windows in order to understand if the observed changes are more or less pronounced.

To further enhance the findings presented herein, it would be beneficial to consider the estimates of transitional changes in land use categories. This would permit the tracking of a cell's transition from one category to another. This approach would improve the comprehension and reconstruction of the dynamics that the territory have experienced between 2013 and 2018. However, the implementation of additional data harmonization techniques is necessary, given that a direct overlap of grids referred to the two years is not feasible. For instance, a possible strategy could be the establishment of a fixed common cell structure over the years. This can be achieved through the realignment of the cells with or without transitioning to a lower level of detail, i.e. aggregating 20x20 m cells to obtain a less refined grid. For these purposes, AI solutions could be a viable option. An example of developed AI strategy helpful in this framework can be found in the paper of [6]. In such a work, a two-stage convolutional

Table 1
Number of cells (# cells) and status share (*100) for 2013 and 2018; change expressed in Net (Eq.(2)), year-to-year relative change in percentage (%), and squared kilometers (km^2) for land use category.

	Land Use Category	2013		2018		Change		
		# cells	Status	# cells	Status	Net	(%)	km^2
1	Anthropized areas	335561	2.324	337103	2.338	0.014	0.5	0.62
2	Corn	19817	0.137	18298	0.127	-0.010	-7.7	-0.61
3	Floral nurseries	302	0.002	424	0.003	0.001	40.4	0.05
4	Forage	3903246	27.029	3347403	23.212	-3.817	-14.2	-222.34
5	Fruits	45428	0.315	48534	0.337	0.022	6.8	1.24
6	Horticultural crops	1559	0.011	2307	0.016	0.005	48.0	0.30
7	Industrial plants and dried legumes	82	0.001	575	0.004	0.003	601.2	0.20
8	Lakes and rivers	170128	1.178	158027	1.096	-0.082	-7.1	-4.84
9	Natural sterile areas	2011747	13.931	1697762	11.773	-2.158	-15.6	-125.60
10	Natural vegetation	1851372	12.820	1421410	9.856	-2.964	-23.2	-171.99
11	Non-classifiable agricultural land	607616	4.208	564154	3.912	-0.296	-7.2	-17.39
12	Olive	400	0.003	877	0.006	0.003	119.3	0.19
13	Other agricultural crops	1065	0.007	1413	0.010	0.002	32.7	0.14
14	Other cereals	304	0.002	617	0.004	0.002	103.0	0.13
15	Resting land	8	<0.001	282	0.002	0.002	3425.0	0.11
16	Seeds	141	0.001	46	< 0.001	-0.001	-67.4	-0.04
17	Vineyards	44523	0.308	38249	0.265	-0.043	-14.1	-2.51
18	Wasteland	302470	2.094	929873	6.448	4.354	207.4	250.96
19	Woods and tree crops	5145423	35.630	5853698	40.591	4.961	13.8	283.31

NN was adopted to align grids available for the same area at different time points. A potential drawback of this solution is linked to the very big size of the area that we need to re-align. Another solution that would improve the production of information useful for policy-making is the identification of possible neighborhood transition patterns (e.g. highlighting clusters of units).

A further research development would focus on distinguishing between structural land use changes and annual rotation of land use specific of agricultural practices. Finally, as already done in the literature, the observed categories could be collapsed into a reduced number of broader categories to better represent the dynamics of land use change [5].

Such a type of results would permit a general coordinated valorization of the whole area, such as interventions in favor of certain destination use, but also regarding the preservation of the land heterogeneity and of sustainable use of the territory. For these reasons it is necessary to acknowledge the importance of such type of data, that actually are available for a limited time window.

Finally, the next step of this research is the modeling of spatial autocorrelation, an aim that comes with several challenges. Firstly, it is important to note that classic AI algorithms have been developed on the basis of intrinsic i.i.d. assumptions. Consequently, these algorithms are not able to explicitly model dependence structures among the data. A variety of strategies for incorporating spatial information into AI models have been proposed. They range from more immediate approaches, such as the inclusion of coordinates as absolute spatial references, to more complex and challenging modeling techniques, involving the integration of the covariance matrix into AI algorithms [7]. The second challenge pertains to the volume of the adopted data: the estimation of the covariance matrix is computational expensive when dealing with such a large number of cells.

In conclusion, these efforts may also be oriented towards the development of an AI solution based on past cells occurred changes and spatial patterns. This method will allow to predict the cell land use across time points not only generally, but also referring to each specific cell (or neighborhood of cells). Such a solution would be useful, in a data-driven perspective, to guide policy-makers in setting regulations or in planning resources according to the expected patterns in land use.

Declaration on Generative AI

The authors have not employed any Generative AI tools in writing this paper or processing the data.

References

[1] E. Tasser, F. V. Ruffini, U. Tappeiner, An integrative approach for analysing landscape dynamics in diverse cultivated and natural mountain areas, Landscape ecology 24 (2009) 611–628. doi:10.1007/s10980-009-9337-9.

[2] L. Arcidiaco, M. Corongiu, Analysis of LULC Change Dynamics That Have Occurred in Tuscany (Italy) Since 2007, Land 14 (2025) 443. doi:10.3390/land14030443.

[3] C. Blasi, G. Capotorti, R. Copiz, D. Guida, B. Mollo, D. Smiraglia, L. Zavattero, Classification and mapping of the ecoregions of Italy, Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology 148 (2014) 1255–1345. doi:10.1080/11263504.2014.985756.

[4] ISTAT, Classification of Municipalities according to the Italian Ecoregions Methodological note, 2023. URL: https://www.istat.it/wp-content/uploads/2018/12/20232509-Methodological-Note.pdf.

[5] G. G. Moisen, K. S. McConville, T. A. Schroeder, S. P. Healey, M. V. Finco, T. S. Frescino, Estimating land use and land cover change in North Central Georgia: Can remote sensing observations augment traditional forest inventory data?, Forests 11 (2020) 856. doi:10.3390/f11080856.

[6] Y. Chen, J. Jiang, A two-stage deep learning registration method for remote sensing images based on sub-image matching, Remote Sensing 13 (2021) 3443.

[7] L. Patelli, M. Cameletti, N. Golini, R. Ignaccolo, A Path in Regression Random Forest Looking for Spatial Dependence: A Taxonomy and a Systematic Review, Springer Nature Switzerland, Cham, 2024, pp. 467–489. doi:10.1007/978-3-031-69111-9_23.