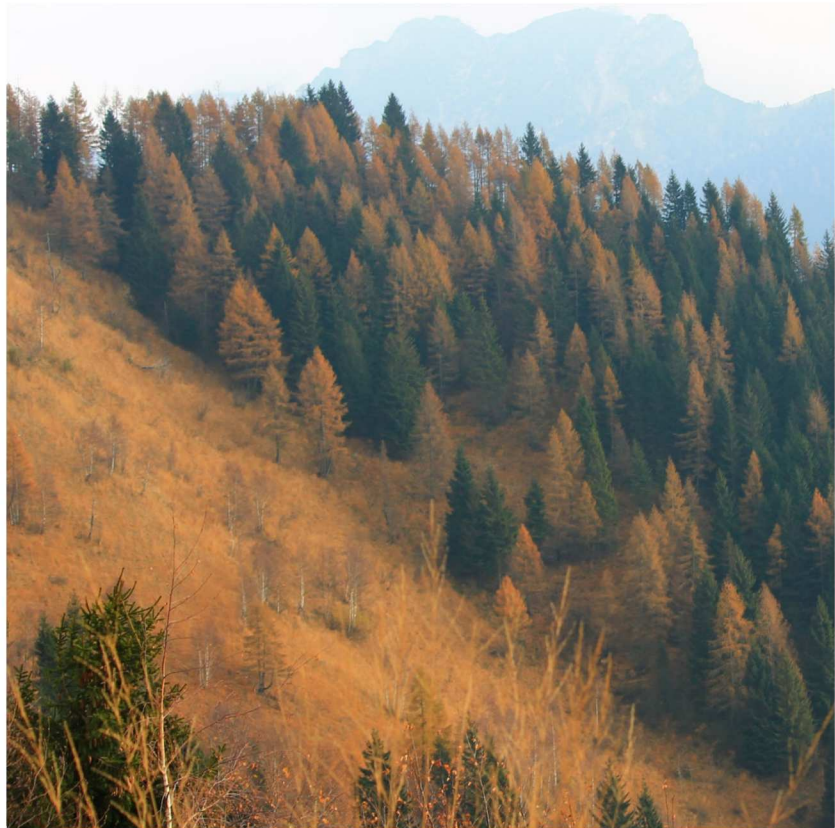




Stefano Gobbi

# **Fine spatial scale modelling of Trentino past forest landscape and future change scenarios to study ecosystem services through the years**



Credits of the cover image Stefano Gobbi



Contents on this book are licensed under a Creative Common Attribution  
Non Commercial - No Derivatives  
4.0 International License, except for the parts already published by other publishers.

University of Trento  
Doctoral School in Civil, Environmental and Mechanical Engineering  
<http://web.unitn.it/en/dricam>  
Via Mesiano 77, I-38123 Trento  
Tel. +39 0461 282670 / 2611 - [dicamphd@unitn.it](mailto:dicamphd@unitn.it)

FINE SPATIAL SCALE MODELLING OF TRENTINO PAST  
FOREST LANDSCAPE AND FUTURE CHANGES SCENARIOS TO  
STUDY ECOSYSTEM SERVICES THROUGH THE YEARS

STEFANO GOBBI

Supervisors:  
Marco Ciolli, Nicola La Porta, Maria Giulia Cantiani

## ABSTRACT

---

Landscape in Europe has dramatically changed in the last decades. This has been especially true for Alpine regions, where the progressive urbanization of the valleys has been accompanied by the abandonment of smaller villages and areas at higher elevation. This trend has been clearly observable in the Provincia Autonoma di Trento (PAT) region in the Italian Alps. The impact has been substantial for many rural areas, with the progressive shrinking of meadows and pastures due to the forest natural recolonization. These modifications of the landscape affect biodiversity, social and cultural dynamics, including landscape perception and some ecosystem services. Literature review showed that this topic has been addressed by several authors across the Alps, but their researches are limited in space coverage, spatial resolution and time span. This thesis aims to create a comprehensive dataset of historical maps and multitemporal orthophotos in the area of PAT to perform data analysis to identify the changes in forest and open areas, being an evaluation of how these changes affected landscape structure and ecosystems, create a future change scenario for a test area and highlight some major changes in ecosystem services through time.

In this study a high resolution dataset of maps covering the whole PAT area for over a century was developed. The earlier representation of the PAT territory which contained reliable data about forest coverage was considered is the Historic Cadastral maps of the 1859. These maps in fact systematically and accurately represented the land use of each parcel in the Habsburg Empire, included the PAT. Then, the Italian Kingdom Forest Maps, was the next important source of information about the forest coverage after World War I, before coming to the most recent datasets of the greyscale images of 1954, 1994 and the multiband images of 2006 and 2015.

The purpose of the dataset development is twofold: to create a series of maps describing the forest and open areas coverage in the last 160 years for the whole PAT on one hand and to setup and test procedures to extract the relevant information from imagery and historical maps on the other. The datasets were archived, processed and analysed using the Free and Open Source Software (FOSS) GIS GRASS, QGIS and R.

The goal set by this work was achieved by a remote sensed analysis of said maps and aerial imagery. A series of procedures were applied to extract a land use map, with the forest categories reaching a level of detail rarely achieved for a study area of such an extension (6200 km<sup>2</sup>). The resolution of the original maps is in fact at a meter level, whereas the coarser resampling adopted is 10m x 10m pixels.

The great variety and size of the input data required the development, along the main part of the research, of a series of new tools for automatizing the analysis of the aerial imagery, to reduce the user intervention. New tools for historic map classification were as well developed, for eliminating from the resulting maps of land use from symbols (e.g.: signs), thus enhancing the results.

Once the multitemporal forest maps were obtained, the second phase of the current work was a qualitative and quantitative assessment of the forest coverage and how it changed. This was performed by the evaluation of a number of landscape metrics, indexes used to quantify the compaction or the rarefaction of the forest areas.

A recurring issue in the current Literature on the topic of landscape metrics was identified along their analysis in the current work, that was extensively studied. This highlighted the importance of specifying some parameters in the most used landscape fragmentation analysis software to make the results of different studies properly comparable.

Within this analysis a set of data coming from other maps were used to characterize the pro-

cess of afforestation in PAT, such as the potential forest maps, which were used to quantify the area of potential forest which were actually afforested through the years, the Digital Elevation Model, which was used to quantify the changes in forest area at a different ranges of altitude, and finally the forest class map, which was used to estimate how afforestation has affected each single forest type.

The output forest maps were used to analyse and estimate some ecosystem services, in particular the protection from soil erosion, the changes in biodiversity and the landscape of the forests.

Finally, a procedure for the analysis of future changes scenarios was set up to study how afforestation will proceed in absence of external factors in a protected area of PAT. The procedure was developed using Agent Based Models, which considers trees as thinking agents, able to choose where to expand the forest area.

The first part of the results achieved consists in a temporal series of maps representing the situation of the forest in each year of the considered dataset. The analysis of these maps suggests a trend of afforestation across the PAT territory. The forest maps were then reclassified by altitude ranges and forest types to show how the afforestation proceeded at different altitudes and forest types. The results showed that forest expansion acted homogeneously through different altitude and forest types. The analysis of a selected set of landscape metrics showed a progressive compaction of the forests at the expenses of the open areas, in each altitude range and for each forest type. This generated on one hand a benefit for all those ecosystem services linked to a high forest cover, while reduced ecotonal habitats and affected biodiversity distribution and quality. Finally the ABM procedure resulted in a set of maps representing a possible evolution of the forest in an area of PAT, which represented a similar situation respect to other simulations developed using different models in the same area. A second part of the result achieved in the current work consisted in new open source tools for image analysis developed for achieving the results showed, but with a potentially wider field of application, along with new procedure for the evaluation of the image classification. The current work fulfilled its aims, while providing in the meantime new tools and enhancement of existing tools for remote sensing and leaving as heritage a large dataset that will be used to deepen the knowledge of the territory of PAT, and, more widely to study emerging pattern in afforestation in an alpine environment.

## CONTENTS

---

1	INTRODUCTION	1
2	MATERIALS AND METHODS	4
2.1	A brief description of the study area	4
2.1.1	Forest in PAT	4
2.1.2	Human population in PAT	8
2.2	Data sources	9
2.2.1	The 1859 dataset	9
2.2.2	The 1936 dataset	10
2.2.3	Aerial imagery of Provincia di Trento	11
2.2.4	The potential treeline maps	12
2.2.5	Other maps for the analysis	13
2.3	The Free And Open Source Software used for the analysis	17
2.4	Preparation of the images: orthorectification of 1954 dataset	26
2.5	A procedure for images classification	34
2.5.1	Image segmentation	35
2.5.2	Classification of objects	36
2.5.3	Distribution of radiometric and geometric parameters for the training areas	40
2.5.4	Classification of the imageries	49
2.5.5	The automatization of the procedure	51
2.5.6	The assessment of the results	52
2.6	Possible future changes scenarios: a single Agent Based Model Approach	54
2.6.1	Agent Based Models	54
2.6.2	Implementation of ABM in the current Case Study	54
2.6.3	Software used: GAMA	57
2.6.4	The calibration of the model	58
2.7	The <i>landscapemetrics</i> package in the R environment	62
2.7.1	The Landscapemetrics package in R	64
3	RESULTS AND DISCUSSION	65
3.1	The 1954 orthorectification of the dataset	65
3.1.1	Qualitative assessment of the precision	67
3.1.2	Statistical assessment of the precision	71
3.2	The digitization of 1859 datasets	81
3.2.1	<i>New Tools for the Classification and Filtering of Historical Maps</i>	85
3.2.2	Assessment of the results of the classification	118
3.3	The results of the classification	122
3.4	Quantitative assessment of the forest landscape: a landscape metrics analysis	129
3.5	<i>Relevance of the Cell Neighborhood Size in Landscape Metrics Evaluation and Free or Open Source Software Implementations</i>	141
3.6	After Vaia storm	164
3.7	Future changes scenarios in the Paneveggio Pale di San Martino Park	166
3.8	Evaluation of ecosystem services	169
4	CONCLUSIONS	172
	BIBLIOGRAPHY	175

## LIST OF FIGURES

---

Figure 1	The position of Provincia Autonoma di Trento in Italy (from Open-StreetMap [85]).	4
Figure 2	Protected areas in PAT, from [88].	5
Figure 3	Examples of habitat diversity in PAT.	6
Figure 4	Main forest types for extension in PAT, from [35].	7
Figure 5	Changes in forest governance over the past 40 years, from [43].	7
Figure 6	Non-native specie distribution in PAT, from [43].	8
Figure 7	Changes in the population of PAT distributed for altitude belts, elaborated from [55].	8
Figure 8	The legend from the 1859 cadastral maps and a sample from the same maps.	10
Figure 9	The legend from the 1936 cadastral maps and a sample from the same maps.	11
Figure 10	An example of each dataset: the Town of Trento in different years.	12
Figure 11	An excerpt from the maps of potential forest areas in PAT, from [87].	13
Figure 12	The map of lidar detected forest area.	14
Figure 13	The map of the historical agricultural land.	14
Figure 14	The map of the forest types.	15
Figure 15	The digital elevation model and its reclassification used in the current work.	16
Figure 16	An example of texture calculated with different window sizes on the 1954 orthorectified images using correlation.	39
Figure 17	Mean values for the first set of 1954 training areas.	40
Figure 18	Mean values for the second set of 1954 training areas.	41
Figure 19	Standard deviation for the first set of 1954 training areas.	41
Figure 20	Standard deviation for the second set of 1954 training areas.	42
Figure 21	Mean values for the first set of 1994 training areas.	42
Figure 22	Mean value for the first set of 1994 training areas.	43
Figure 23	Standard deviation value for the first set of 1994 training areas	43
Figure 24	Standard deviation value for the first set of 1994 training areas	44
Figure 25	Mean values value for the 2006 training areas.	44
Figure 26	Mean value for the first 2015 training areas.	45
Figure 27	Standard deviation value for the 2006 training areas.	45
Figure 28	Standard deviation for the 2015 training areas.	46
Figure 29	Geometric parameters for the 1954 training areas	46
Figure 30	Geometric parameters for the 1994 training areas	47
Figure 31	Geometric parameters for the 2006 training areas	47
Figure 32	Geometric parameters for the 2015 training areas	48
Figure 33	An example of different histogram from the 1954 dataset.	51
Figure 34	An example of different histogram from the 1994 dataset.	52
Figure 35	The position of Paneveggio Pale di san Martino natural park.	55
Figure 36	The position of the test areas in the Paneveggio Pale di San Martino natural park.	59
Figure 37	Calibration of prob_seed	59
Figure 38	Calibration of prob_germ	60

Figure 39	Index for the orthorectified images in the current work from the 1954 dataset.	66	
Figure 40	Rectified photo 19w_3856	67	
Figure 41	The measurements of the fiducial markers.	68	
Figure 42	Graphic User Interface for the <i>g.gui.image2target</i> command.	68	
Figure 43	The distribution of the RMSE for the orthorectified images.	69	
Figure 44	Example on the effect of placing a wrong GCP: on the right the wrong GCP is included in the orthorectification and it increase the values of the single ORMSE for every GCP, respect to the image on the bottom, where deselecting the wrong GCP drastically decrease the ORMSE.	70	
Figure 45	Comparison on the same area of orthophoto 13e_4954 with different DEM used for rectification (1x1m resolution on the left, 10x10m on the right).	71	
Figure 46	Comparison on the same area of orthophoto 13e_4954 with different DEM used for rectification: the left figure (DEM resolution of 1x1m, on the right a DEM resolution of 10x10m) guarantees a better interpretation of the highlighted area.	71	
Figure 47	Examples of unreliable natural features for placing GCPs: mountain peaks are easily recognizable, in the 2015 (on the right) image, while they are not in the 1954 imagery. For a unbiased comparison both images are shown in greyscale.	79	
Figure 48	Examples of reliable anthropic features for placing GCPs: malghe and stables are easily recognizable in both the imagery set.	79	
Figure 49	Examples of unreliable anthropic features for placing GCPs: some road junction changed drastically shape making it difficult to exactly recognize the same feature usable as GCP.	79	
Figure 50	The shape of hairpin bends, another anthropic feature usable for GCPs.	80	
Figure 51	The results of the digitization of the 1859 H.C.P.A.T.	81	
Figure 52	A simulation of the distribution of the trees in 1859.	82	
Figure 53	Forest types in wooded pastures	83	
Figure 54	Forest types in wooded pastures	84	
Figure 55	The calculus of the number of patches metric for the forest for the different years, divided by altitude ranges.	118	
Figure 56	The forest area extracted from the Land use classification process.	123	
Figure 57	The forest area extracted from the 1859 and 1936 dataset adapted to the forest of 1954.	123	
Figure 58	The evolution of forest coverage trough the years.	124	
Figure 59	The evolution of forest coverage trough the years divided per altitude ranges.	125	
Figure 60	The evolution of forest coverage trough the years around the treeline limit	126	
Figure 61	The evolution of forest coverage trough the years in the potential and excluded areas	127	
Figure 62	The evolution of forest class coverage	128	
Figure 63	Evolution of the aggregation index for the total forest in PAT	129	
Figure 64	Evolution of the mean patch area for the total forest in PAT	130	
Figure 65	Evolution of the largest patch index for the total forest in PAT	130	
Figure 66	Evolution of the number of patches for the total forest in PAT	130	
Figure 67	Evolution of the patch density for the total forest in PAT	131	
Figure 68	Evolution of the total edge for the total forest in PAT	131	



Figure 69	Evolution of the aggregation index divided by altitude ranges	132
Figure 70	Evolution of the mean patch area divided by altitude ranges	133
Figure 71	Evolution of the largest patch index divided by altitude ranges	133
Figure 72	Evolution of the number of patches divided by altitude ranges	133
Figure 73	Evolution of the patch density divided by altitude ranges	134
Figure 74	Evolution of the Simpson Index for forest divided by altitude ranges	134
Figure 75	Evolution of the Shannon Index for forest divided by altitude ranges	134
Figure 76	Evolution of the aggregation index divided by forest classes	137
Figure 77	Evolution of the mean patch area divided by forest classes	138
Figure 78	Evolution of the largest patch index divided by forest classes	138
Figure 79	Evolution of the number of patches divided by forest classes	139
Figure 80	Evolution of the patch density divided by forest classes	139
Figure 81	Evolution of the Simpson Index for forest divided by forest classes	139
Figure 82	Evolution of the Shannon Index for forest divided by forest classes	140
Figure 83	Cell neighbourhood shape, (a) Von Neumann, (b) Moore.	141
Figure 84	Areas interested by windthrow after Vaia storm (in red) over the 2015 forest map.	164
Figure 85	The windthrow after Vaia storm in Passo Lavazzé, from [42]	165
Figure 86	The prediction of the evolution of forest in Paneveggio Pale di San Martino using ABM.	166
Figure 87	The prediction of the evolution of forest in Paneveggio Pale di San Martino using Markov Chain.	167
Figure 88	The value for the RUSLE formula evaluated in the different year of the current analysis.	170

## LIST OF TABLES

---

Table 1	A list of the papers published in the current work.	3	
Table 2	A resume of the main features of each map used in the current study.		12
Table 3	Denomination of forest classes in PAT	15	
Table 4	Values of threshold and minimum area size for each set of aerial imagery.	35	
Table 5	Classification scheme used for the LULC analysis. The id column identifies the category number assigned to each LULC class.	37	
Table 6	Number of training areas for each year of aerial imagery.	38	
Table 7	Number of training areas divided for land use for each dataset	38	
Table 8	Computational time of the texture on a same area changing the window size.	40	
Table 9	Evaluation of the classification accuracy for the image 80151 in the 2015 dataset, which contains the training areas.	50	
Table 10	Evaluation of the classification accuracy for the image 80152 in the 2015 dataset, which does not contains the training areas.	50	
Table 11	Evaluation of the classification accuracy for the image 80151 in the 1994 dataset, which contains the training areas	50	
Table 12	Evaluation of the classification accuracy for the image 80152 in the 1994 dataset, which does not contains the training areas	50	
Table 13	Computational time (in days) for each set of images used in the study		51
Table 14	Distribution of random stratified points	53	
Table 15	Forest classes extensions in Paneveggio Pale di San Martino Area	55	
Table 16	Threshold for probability of seedling and accuracy of the model over the test years.	60	
Table 17	Threshold for probability of germination and accuracy of the model over the test years.	60	
Table 18	Best values for the threshold for probability of seedling and germinating.	60	
Table 19	Accuracy of the model in simulating the evolution of Paneveggio Forest from 1954, comparison between the real and simulated situation in years 1994, 2006, 2015	61	
Table 20	List of the orthorectified images	65	
Table 21	Features of the rectified images	66	
Table 22	Variation of the precision in photograph 13e_4954 for different choices of DEM and scanning of 600 DPI.	70	
Table 23	Images chosen for statistical analysis and relate ORMSE	72	
Table 24	First set of ICPs for the image 20c_4378.	73	
Table 25	Second set of ICPs for the image 20c_4378.	73	
Table 26	Measurements on the ICPs for the image 20c_4378.	74	
Table 27	Measurements on the ICPs for the image 13e_3207	75	
Table 28	Measurements on the ICPs for the image 20c_4384	76	
Table 29	Measurements on the ICPs for the image 23_1738	77	
Table 30	Measurements on the ICPs for the image 12e_5096	78	
Table 31	Forest types in wooded pastures	83	
Table 32	Mean altitude of the wooded pasture for different areas of PAT		84
Table 33	Accuracy level of the OBIA procedure for forest class	118	

Table 34	Evaluation of the threshold value of the false positives for the forest class in the 18_3821 image.	119
Table 35	Evaluation of the threshold value of the false positives for the forest class in the 21_4575 image.	120
Table 36	Evaluation of the threshold value of the false negatives for the forest class in the 18_3281 image	120
Table 37	Evaluation of the threshold value of the fake negatives for the forest class in the 21_4575 image	121
Table 38	Accuracy level of the OBIA procedure for forest class, after the correction of the 1954 and 1994 dataset.	122
Table 39	Total forest coverage for the PAT over the years	124
Table 40	Forest coverage (in ha) divided for altitude ranges	125
Table 41	Forest coverage (in ha) for the altitude belts over the treeline limit	126
Table 42	Forest coverage (in ha) for the potential or excluded potential areas	126
Table 43	Forest coverage for forest classes	128
Table 44	Evolution of the landscape metrics calculated for the whole PAT forest	129
Table 45	Calculated landscape metrics for the forest divided by altitude belts.	132
Table 46	Calculated landscape metrics for the forest divided in forest classes.	135
Table 47	Calculated landscape metrics for the forest divided in forest classes.	136
Table 48	Calculated landscape metrics for the forest divided by forest classes.	137
Table 49	Areas interested by Vaia storm.	164
Table 50	A possible evolution of forest extension in Paneveggio Pale di San Martino Park, a comparison between two methods for future changes scenarios prediction	166
Table 51	A comparison between metrics calculated on both the simulated landscape using MC and ABM	167
Table 52	Value of avoided soil erosion in areas occupied by forest across the dataset.	170

## INTRODUCTION

---

The Alps represent one of the most relevant mountain landscape in Europe. Those slopes, peaks, valleys which were formed and shaped from water and winds host a wide variety of habitats. The inhabitants of these mountains directly and indirectly have shaped the landscape, and especially in recent decades new tendencies have emerged in the interaction between human presence and the mountain landscape [71]. Abandonment of pastures and traditional agricultural grounds, because of their low productivity and distance from urban centres, resulted in a progressive afforestation of meadows and open areas [71]. This is noticeable as during the 20th century the historic shape of the landscape, an alternation between compact patches of forests, meadows and barren land, has been slowly replaced by homogeneous forests [101].

The issue of landscape changes has been studied with great concern by different authors, because it has been perceived mostly with a negative connotation [5].

Landscape changes involved a wide variety of different land uses, agricultural, urban and of course forest. It has been studied from different approaches, based on the perception of the people [110, 113] or remote sensed parameters [2, 24, 32, 34] or holistic approach [110] in order to describe the effects that landscape changes had on the populations. In particular, forest landscape has been extensively studied in mountainous environment because it provides habitats [6], timber [11], protection from erosion [69] and cultural value [110]. In this framework, it is crucial to know and understand the past forest dynamics which led to the current situation [20], in order to understand the pattern of forest dynamics in a given study area, and how these dynamics will shape the future of the landscape. A general tendency that has been observed across the Alps is forest expansion [95, 96, 106]. This process has led to a loss of habitat diversity, and the issue was so relevant that it has been addressed by the European Union, which published different reports and took relevant habitats conservation efforts [1] such as the Pan-European Biological and Landscape Diversity Strategy, the Bern Convention, the European Landscape Convention and the Birds and Habitats Directives.

The aims of the current work are to reconstruct the past forest landscape using a wide range of data between those available, to study the evolution of forest landscape across the study area, evaluate how this change affected the ecosystem, to create a future change scenario and to study changes in ecosystem services through time in the area of Provincia Autonoma di Trento, (PAT from now on). The study area had a great availability of valuable input data to include in the analysis: maps and aerial imagery [37, 38, 30, 31, 108, 87, 91] and other sources of historical data about forests [35, 83, 18]. PAT has natural traits similar to those found in different areas of the alpine arc. From lake Garda, which mild the climate of nearby valleys to the cold valleys in the North, PAT hosts a wide variety of different habitats.

During the last century PAT experienced profound changes that affected the local population and environment: World War I was fought all along the borders of this region, and forests were heavily exploited by all the armies involved, also World War 2 affected some areas likewise. After the wars, especially in the second half of the Twentieth Century the mountains were progressively abandoned by the local population. The agricultural and pastoral activities were abandoned in favour of other occupations [110], a dynamic that emerged across different mountain areas in Europe [71]. It is reasonable to assume that under the hypothesis of a similar human-forest interaction and without strong external factors, the changes that happened in PAT could represent the trend of a number of other areas in the alpine arc. For these reasons, PAT was chosen as an ideal "study area" for tackling the issue of forest

landscape changes in the alpine region. The different climatic situations and habitats in PAT, can well-represent a region as wide as the Italian alpine arc. Moreover, a great variety of input data regarding the PAT territory are available for studies on the subject of forest cover variations.

The data used in this work, came from different sources such as the Historical Cadastral maps of the Austro-Hungarian Empire [91], the Italian Kingdom forest maps from 1936 [31], and the more recent aerial imagery, that were integrated with forest classes maps [84] and other classifications of the forest area [87]. A remarkable aspect of the current work is the fine scale approach chosen: the maps used to characterize forest cover changes have a very fine resolution, which does not exceed the 2x2m pixel dimension [41]. Instead of using fine temporal resolution data, such as Landsat or other satellite imageries, it was decided to have a coarser time resolution, using aerial imageries and favouring spatial resolution. This approach has been attempted from other authors for regions with an extension similar to the PAT, but their studies were either limited in spatial resolution [17, 29, 93, 28, 3], or limited in time [74, 9, 67]. Assessing this analysis across a great number of different data meant also setting up procedures for automatizing the treatment of imageries and assessing of output results. The output data about the forest coverage were analysed using Landscape metrics to assess the quality of the forest evolution in the timespan analysed. On this subject, an interesting issue was found bounded with the definition of a basic parameters of the landscape metrics and it has been extensively studied and discussed.

Forest maps were used to analyse and estimate some ecosystem services, in particular protection from soil erosion, the changes in biodiversity and the landscape value of the forests. Finally, a procedure for the analysis of future change scenarios was set up to study how afforestation will proceed in absence of external factors in a protected area of PAT. The procedure was developed using a particular procedure called Agent Based Models, which considers trees as thinking agents, able to choose where to expand the forest area.

Understanding forest dynamics, which is the "real world" that stands behind the numbers and maps reported in the current thesis is the main need behind the current work. Many different works have been carried out throughout the PAT territory to evaluate the changes in forest [108, 110, 22], as well the quality of the forest landscape [19, 84, 100, 112], although for locals, limited areas only. Yearly, The Servizio Fauna e Foreste of the PAT releases a series of reports assessing the quality of the forests [35, 43], and these works analyse only the previous year of activity in forest monitoring. This work shaped a procedure and created a collection of data able to analyse changes for a long timespan: by understanding the past history of forest, it would be possible to predict the fate of rural abandoned areas, as well as the evolution of forests in areas destroyed by catastrophic events, such as fires and windthrows, as it happened in 2018 [81]. The dynamics that could change the forest landscape are difficult to control, but where there is a strong link between forest and policy makers it is possible to intervene to protect those areas which presents frailty and a fragile balance between forest and human presence [20].

In the framework of this work the following papers were published, and presented in this thesis (table 1).

Orthorectification of a large dataset of historical aerial images: procedure and precision assessment in an open source environment	page 34
Fine spatial scale modelling of Trentino past forest landscape (Trentinoland): a case study of foss application	page 26
Object-Based Image Analysis for historic maps classification	page 94
New Tools for the Classification and Filtering of Historical Maps	page 118
Relevance of the Cell Neighborhood Size in Landscape Metrics Evaluation and Free or Open Source Software Implementations	page 164

Table 1: A list of the papers published in the current work.

## MATERIALS AND METHODS

---

### 2.1 A BRIEF DESCRIPTION OF THE STUDY AREA

#### 2.1.1 *Forest in PAT*

Provincia Autonoma di Trento (PAT) is a region in the north of Italy. From an European point of view PAT is located in the southern part of the Mid-European region. Its extension is approximately 6200 km<sup>2</sup>. As the lowest and highest point of PAT are respectively 64 m.a.m.s.l (Valle del Sarca) and 3769 m.a.m.s.l. (Monte Cevedale), the great gradient in altitude provides different micro climate zones that support a great variety of habitats, from the glaciers in the Adamello, Pale di San Martino, Marmolada and Cevedale groups to different forest types along the elevation gradient, until the fertile agricultural bottom valley [100]. In the southern part of the Province, the influence of Garda Lake creates the conditions for spontaneous mediterranean vegetation. The northernmost holm oak forests and olive tree fields are found here [100].



Figure 1: The position of Provincia Autonoma di Trento in Italy (from OpenStreetMap [85]).

The importance of the natural environment in PAT is recognized both by the Italian Government and the local Government. In fact, the institution of a national park (Parco Nazionale dello Stelvio), two regional parks (Parco Naturale Adamello Brenta, Parco naturale di Paneggio Pale di San Martino) and 142 habitats in the Natura2000 network are some of the initiatives undertaken to preserve the natural environment [88].

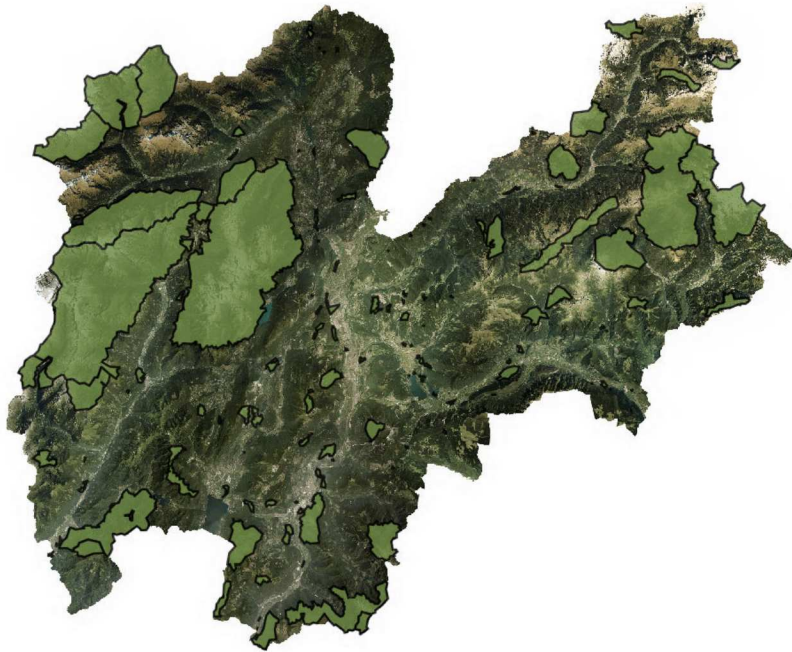


Figure 2: Protected areas in PAT, from [88].

Climate in PAT is classified as 'alpine', following Koppen classification [100]. The orography favours the presence of microclimates, which are responsible of extremes in temperatures (e.g. the record for the lowest temperature ever reached in Italy was set in Pale di San Martino group,  $-49,6^{\circ}\text{C}$ ) and favours the presence of endemic species.





(a) The lush vegetation in Val di Leno.



(b) The peak of Adamello



(c) Agricultural ground in Garda Lake



(d) Forests in Vigolana.

Figure 3: Examples of habitat diversity in PAT.

Forests cover approximately the 60% of PAT territory, and they are characterized by a majority of coniferous species, mostly Spruce, which covers an exceeding of 30% of the forest surface [35]. Beech is the predominant broadleaf specie, it covers the 14% of forest surface [35]. Beeches are coming back in PAT, after years of human intervention to favour resinous tree over broadleaves. The inversion of this tendency was the results of 50 years of governance aiming to recover and protect beeches, recognizing their importance. Fir tree are as well taking back some of the land they occupied before the forest management carried out by the Hasburgic empire improved the presence of Spruce in PAT [43].

**Distribution by area of the main forest types in PAT**

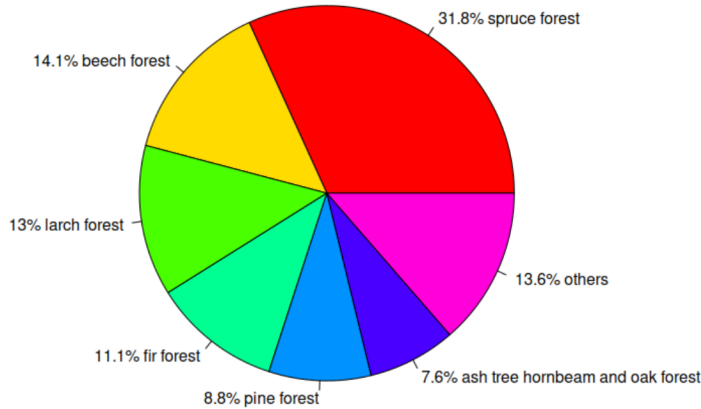


Figure 4: Main forest types for extension in PAT, from [35].

From the 1950's each public forest parcel and the main forest parcels owned by private and companies had to adopt forest management plans, which describe in details how trees in a single parcel should be managed [43]. These plans last 10 years [43]. Smaller parcels were grouped in an inventory of private forest land, and they had to adopt a management plan as well [35]. As a general trend the data collected by these plans underline an increase in wood mass and wood productivity, underlining the great economical and social value which forests have in PAT [35]. In figure 5 the percentage of forest governed as coppice and high forest is shown, and how this two methods varied throughout the last 40 years [43].

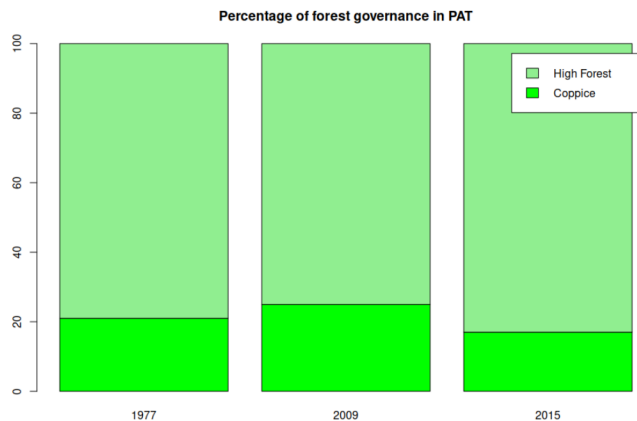


Figure 5: Changes in forest governance over the past 40 years, from [43].

Non-native species in PAT were introduced during the last century mainly for economical reason, for instance *Paulownia tomentosa* and *Ailantus Altissimus*, were introduced for their high biomass productivity [43]. However, the distribution of these species is uneven across the territory.

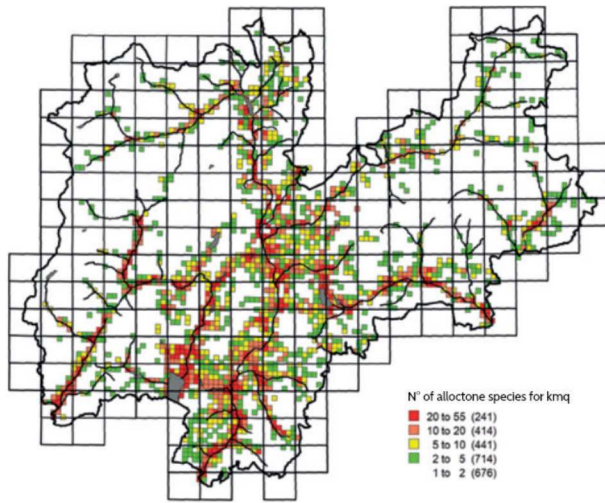


Figure 6: Non-native specie distribution in PAT, from [43].

Figure 6 shows how the majority of non-native specie are distributed in the valleys bottom, especially near the Garda Lake area, where the climate is milder.

### 2.1.2 Human population in PAT

The main cities in PAT are Trento, Rovereto, Riva del Garda and Borgo Valsugana. All these urban centres are located in valley bottoms, and while they increased in size, the towns and villages in the mid-mountains (500-1000 mamsl) and high mountains (above 1000 mamsl) altitude range decreased and were then abandoned, as shown in figure 7.

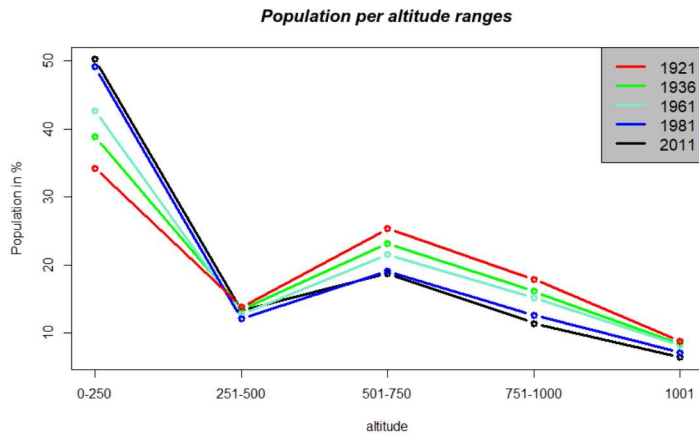


Figure 7: Changes in the population of PAT distributed for altitude belts, elaborated from [55].

## 2.2 DATA SOURCES

The study of the evolution of the forest in the current study began from the year 1859, year of the redaction of the oldest map considered in the analysis, until 2015, when the set of orthophotos of the entire of PAT territory became openly available.

The precision and accuracy of the analysis strongly depends on the good characteristics of the input images used for remote sensing [15], such as high resolution and a good readability [57]. This were the main criteria that drove the choice of the input maps used in the analysis. Despite Landsat imageries have good quality and are yearly distributed, so they could constitute a good source of input imageries, the choice for the current study was to use high resolution aerial imagery. These input may be coarser in the time resolution, but the spatial resolution and the precision achievable was considered a good trade-off in front of temporal low resolution. The mapset from the year 1859 and 1936 were thematic maps, which did not carry informations about the distribution of each single tree. Their importance resided in being a valuable and reliable source of information from an era when the aerial imagery were not yet diffused [31, 97].

The following sections presents in detail the input images used in the current work to analyse forests changes. The native datum used in the maps of the analysis are a ETRS89 datum with a UTM32N projection system for the 1859 mapset, a Bessel ellipsoid with a Samson-Flamsteed reference system for the 1936 maps, a RM40 and Gauss Boaga projection system for the 1994 aerial imagery, while all the other maps share a WGS84 datum with a UTM32N projection system.

### 2.2.1 The 1859 dataset

The '*Historical Cadastre Map for the Province of Trento*', or *Theresianischer Kataster* (H.C.P.A.T. in the text from now on) is a dataset constituted by hand-drawn map built for cadastral needs [97]. The surveys aimed to identify cadastral particles and propriety boundaries, but informations about the land use were included, in order to discriminate between natural and agricultural uses [98]. The survey for the PAT started in 1855, using a projection system centred at the Pfarrturm in Innsbruck [97]. The complete mapping of PAT includes 13 297 sheets each covering 288 ha [99]. In 2006 the Land Registry Service (Servizo Catasto) of the PAT completed the process of digitalization (scanning all the sheets at 230 DPI and 24 bits of color depth) and georeferenciation of all the maps, in the ETRS89/UTM 32 N (EPSG 25832) datum, using as Ground Control Points (GCPs) the four corner of each sheet [97]. The maps were made available for free download as JPEG images, under the Creative Commons Attribution 4.0 license. After this digital processing the resolution of the maps was around 0.32 cm per pixels, while the positional accuracy of the features is of 5-10 m. The maps and other informations can be found at on the HistoricalKat web site [99].

From these maps the presence of forest was extracted by manually digitizing the areas indicated by the legend as 'forests'. The results were saved in one vector file for each municipality of PAT. Additional land cover classes that were vectorized were pastures and meadows (as they were both private properties or public properties and they were registered carefully by the Austrian Government) and wooded pasture, areas where sparse trees grew in pastures. All these informations were stored in the H.C.P.A.T., coded with colours and icons, which were associated to each and every cadastral parcel. The legend in figure 8 reports on all the categories represented by the H.C.P.A.T.

The importance of the H.C.P.A.T. lies in the fact that it offers a view on the landscape of PAT as it was 160 years ago, coded as a 'modern' thematic map [125, 14]. In fact the Cadastral Service only projected the H.C.P.A.T. from a reference and coordinate system which was

coherent and was actually used for taxation, so had to represent all the land parcels precisely [97]. To remark the importance of these maps, they were drawn for the whole territory of the Austro-Hungarian empire sharing the same reference system and the same legend, nowadays they represent a valuable source of information for all the studies involving past landscape in the Mittel-European region [14]. A limit for the current study lies in the representation of the forests: single trees were not drawn, so without additional data (e.g. forestry commission reports) it is difficult to calculate how dense the forest cover was in a forest parcel.

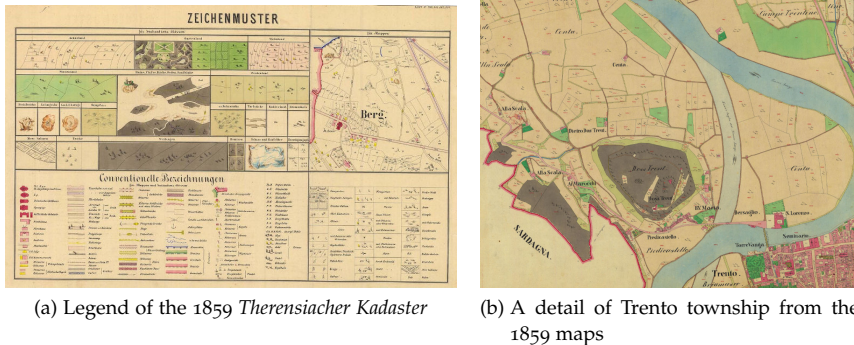


Figure 8: The legend from the 1859 cadastral maps and a sample from the same maps.

### 2.2.2 The 1936 dataset

The Italian Kingdom Forest map (IKFM from now on) was published in 1936 by the Milizia Italiana forestale, as results of forest and forests types surveys across the Italian territory [31]. This map is one of the first examples of forest types mapping: its legend reports a total of 25 different forest types [31]. The IKFM is based on the cartography produced by the Istituto Militare Italiano on a scale of 1:100 000 that used a peculiar datum, a Bessel ellipsoid with four different orientations (Genova, Monte Mario, Castanea delle Furie, Guardia Vecchia) and a Samson-Flamsteed reference system [31], and represented the whole Italian territory in 276 different sheets. To import these maps in a digital format, they were scanned with a resolution of 400 DPI in a TIFF format, and re-projected in a Gauss-Boaga reference system RM40 datum. Then forest areas and, within them, forest types were manually vectorized and classified by digitizing each forest patch [31].

The resulting digitalization is freely available under the Creative commons Attribution - Version 3.0 license and can be viewed and downloaded from a dedicated webgis [30].

The forest area for the PAT was extracted by converting the digitized map for the whole Italian territory in a raster tiff format and then using the PAT boundaries as a mask. The importance of this maps lays in indicating the presence of forests in the Italian territory in the middle Wars period. Although the historical importance of these maps, they hide some issues.

The first problem lies in the digitization of the process. These maps were drawn in a particular projection with 4 different orientations, and then they were translated in a Gauss-Boaga reference system, which introduced a random error on the coordinates of around 100m localized in some small areas.

Secondly the scale of 1:100 000 affects the representation of the smallest forested areas which were often neglected. Finally, some areas present an ambiguous representation because human error in the drawing process and cannot be interpreted as forested areas [31]. These maps were used in the present work for their historical importance, as they carry informa-

tions of the situation after the heavy exploitation of timber happened during Word War I [52].

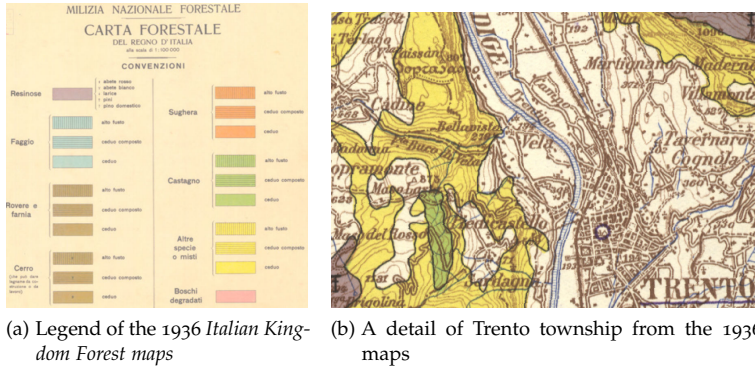


Figure 9: The legend from the 1936 cadastral maps and a sample from the same maps.

### 2.2.3 Aerial imagery of Provincia di Trento

The sets of aerial imagery that were chosen to reconstruct the past landscape of the study areas were extracted from the 1954, 1973, 1994, 2000, 2006 and 2015 aerial imagery datasets. An aerial image is a picture that represents the territory in a georeferenced format, with a cartographic projection. They differ from aerial photographs, as the first have a geographical projection, while the latter have a central plain projection. Being georeferenced means that each point in an aerial images has coordinates in a given reference system.

In particular, the imagery dataset of the 1954 was not in a geographical format, therefore needed orthorectification [44]. This process projected the 1954 imagery in WGS84, UTM32N reference system, shared with the 2006, 2015 imagery. The aerial imagery set from 1994 were in RM40, Gauss-Boaga projection, but it was decided that for the sake of having a unique, consistent reference system through the whole study was worth transforming their reference system in WGS84 UTM32N. The imagery sets of 1954, 1994 were in grey scale, with a resolution of respectively 2x2m pixels and 1x1m pixels. The imagery sets of 2006 and 2015 were in RGB with three different bands. Their resolution is respectively 0.5x0.5m per pixels and 0.2x0.2m per pixels. The maps of 1994, 2006 and 2015 are available on the PAT webgis in digital format, while the 1954 maps were kindly provided in their digital format, i.e. paper photographic prints. Then they were digitized and elaborated in the current work.

For the current analysis the images of the 1973 and 2000 dataset were preliminary excluded due to their quality being too poor in terms of exposure and quality of the images to extract accurate data about forest cover [108, 109]. The images of 1973 often presented scratches and marks on them, which affects the readability by users and makes it difficult for the software used to perform the requested elaborations. The images of 2000 had a poor exposure and contrast data, thus making it difficult for the software to perform a classification analysis with a satisfactory accuracy. Image 10 shows some extracts from each single dataset, while table 2 presents the main feature of the aerial imageries used and those discarded.

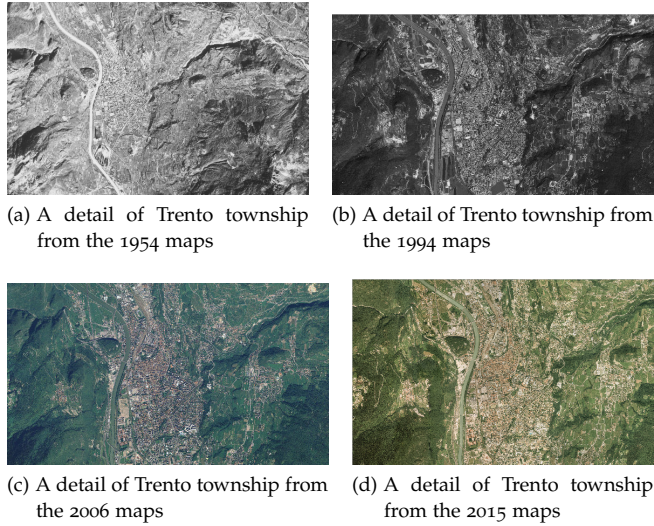


Figure 10: An example of each dataset: the Town of Trento in different years.

Year	1859	1936	1954	1973	1994	2000	2006	2015
Denomination	H.C.P.A.T	I.K.F.M	Volo GAI	Volo Italia	Volo Rossi EIRA	-	Volo Terraitaly	Volo AGEA
Scale	1:1440	1:100000	1:35000	1:10000	1:10000	1:5000	1:5000	1:5000
Type of map	Topographic map	Topographic map	Aerial images BN	Aerial images BN	Aerial images BN	Aerial images RGB	Aerial images RGB	Aerial images RGB
Pixel dimension	0.2x0.2	0.3x0.3	2x2m	1x1m	1x1m	0.5x0.5	0.5x0.5m	0.2x0.2m
Dimension [GB]	20	1	8	8	8	1000	100	730
Datum	ETRS89-UTM32N	Gauss Boaga RM40	WGS84-UTM32N	Gauss Boaga-RM40	Gauss Boaga-RM40	WGS84-UTM32N	WGS84-UTM32N	WGS84-UTM32N

Table 2: A resume of the main features of each map used in the current study.

#### 2.2.4 The potential treeline maps

In 1992 Piussi [87] proposed a map of the areas of potential forest areas for the whole PAT territory. This map classifies the PAT territory in 4 main classes:

- Forest areas: the presence of forests was indicated by on-fields survey and are considered forest areas those covered by the 30% of trees [87];
- Area of effective potential, areas where trees can potentially grow to cover the 30% of the territory despite some 'bad' orographic factors, such as presence of ancient landslides, bad exposure, potential presence of avalanches during winter [87];
- Area of potential exclusion, areas where trees cannot grow to cover the 30% of the territory because 'bad' orographic factors [87];
- Area over the potential limit, areas over tree limit [87].

These areas were identified by on-field survey and then reported on the topographical maps of 1980 of the IGMI in a 1:50 000 scale, projected in a RM40-Gauss Boaga reference system. In this representation the PAT territory is divided in 98 different maps, which were scanned, georeferenced and classified with a procedure similar to the LULC procedure described in sections 2.5 and 3.2.1. In this study were considered only tree-habit trees and not shrubs-like species, such as *Pinus mugo* and *Alnus alnobetula* (the latter also known as *Alnus Viridis*).

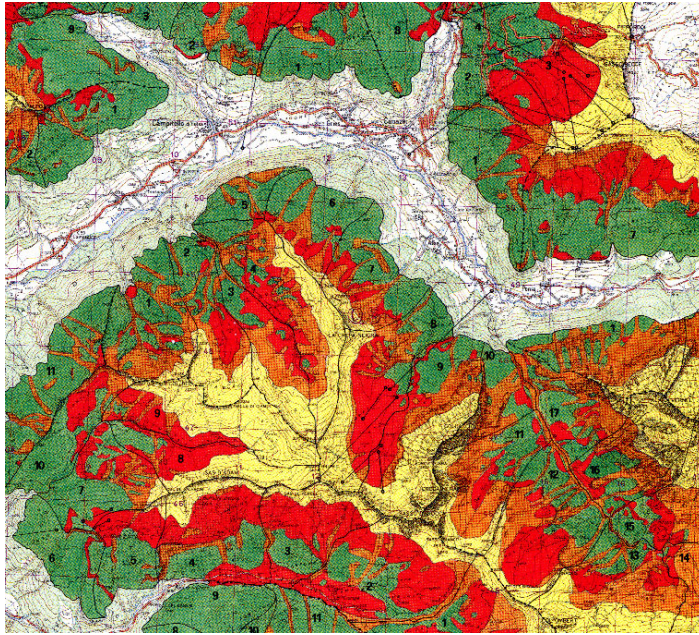


Figure 11: An excerpt from the maps of potential forest areas in PAT, from [87].

#### 2.2.5 Other maps for the analysis

A set of complementary maps was used as ground truth. One of these maps was the maps of forested areas calculated from the Lidar survey carried out in 2006 by the PAT [38], with a resolution of  $1 \times 1 \text{m}$  pixel. The areas of forest were calculated by subtracting the Digital Terrain Model from the Digital Elevation Model. The areas with a positive difference (forest areas) were then digitized in a vector map. This map was used to perform a comparison between the area of extracted forests from each dataset and the real extension of the forest in the period when it was maximal to perform visual assessment of the classification accuracy. Furthermore they were used to include areas of forest covered by shadows that otherwise would have been potentially excluded from the analysis.



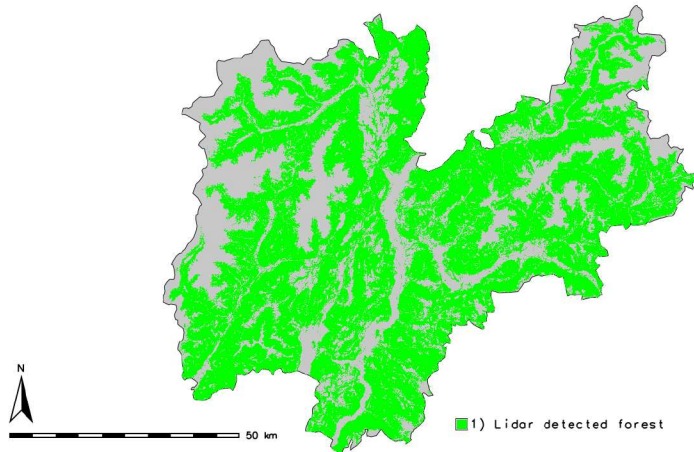


Figure 12: The map of lidar detected forest area.

The map of the historical agricultural land was used for the current work. These areas are defined as those valuable agricultural grounds, which were established during the XIX century, and since then never changed use. This map were produced by the cartographic service of the PAT, as vector map that indicates the cadastral parcels dedicated to some traditional cultures in PAT: apple, olive and grapes for instance [37]. This map was used as mask to discriminate false positives in the forest category, as tree in orchards were in some cases mistaken as forest during the elaboration of forest cover.

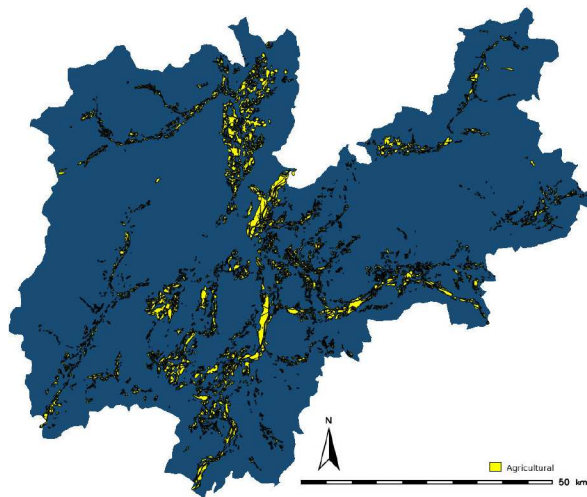


Figure 13: The map of the historical agricultural land.

The map of forest types, updated to the year 2009 was provided by The Servizio Fauna e Foreste of the PAT. This map is the result of a series of surveys carried out to register every forest parcel of public propriety into a specific forest type [39, 111, 19, 83, 84]. Under the hypothesis that forest types have not changed in the past 60 years this additional information was used to classify the forests in sub-categories depending on the forest type. This process was performed on the map of the forest extracted from the 1954, 1994, 2006, 2015 dataset. The forest types were grouped into 15 macro-categories, listed in table 3:

Forest class (English names)	Forest Class (Italian names)
Holm Oak forests	Lecceta
Oak forests	Querceta
Mixed Oak-Hornbeam forests	Querco-carpineto
Chestnut and Black Locust forests	Castagneto-Robinieto
Maple, Linden, Ash forests	Aceri-frassineto e Tiglieto
Pioneer formations	Formazioni transitorie
Scots Pine forests	Pineta
Beech forests	Faggeta
Fir forests	Abieteto
Mountain Pine forests	Mugheta
Alder forests	Ontaneta
Spruce forests	Pecceta
Larch forests	Lariceta
Larch and Stone Pine forests	Larici-Cembreta
Stone Pine forests	Cembreta

Table 3: Denomination of forest classes in PAT

Their distribution through PAT is shown in map 14.

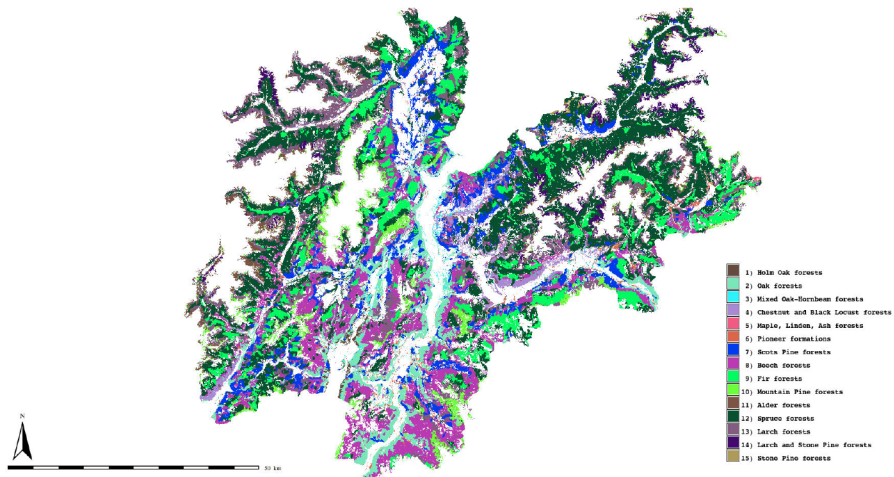


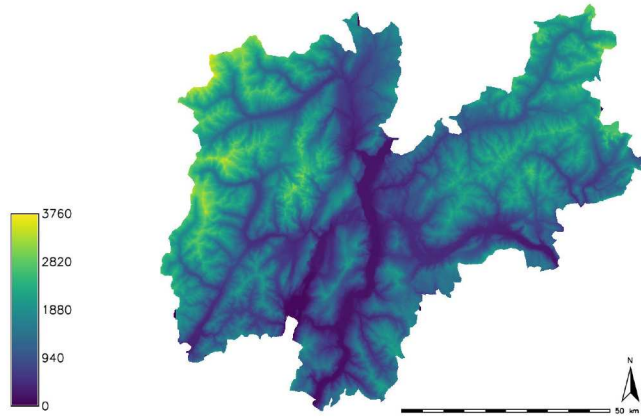
Figure 14: The map of the forest types.

Finally a Digital Elevation Model provided by the PAT [40] was used to divide the forest category in altitude ranges [56]:

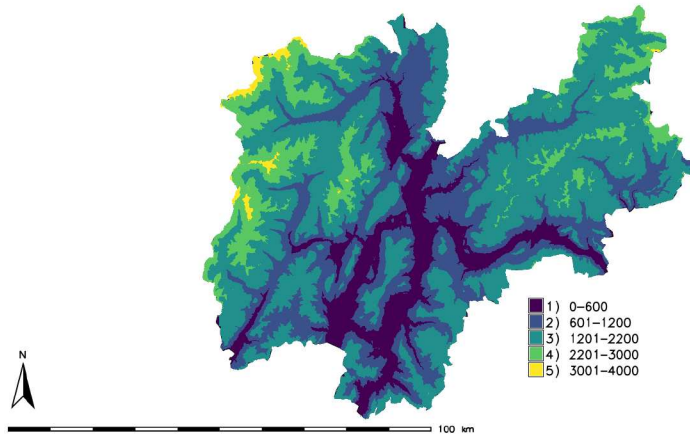
- 0-600 m a.m.s.l. valley bottom range;
- 601-1200 m a.m.s.l. low mountain range;
- 1201-2200 m a.m.s.l. mid mountain range;
- 2201-3000 m a.m.s.l. high mountain range;

- 3001-4000 m a.m.s.l. glacier range.

The DEM has a resolution of 1x1m per pixel, it was obtained from a Lidar survey carried out in 2008 [40].



(a) The digital elevation model for the PAT



(b) The reclassification of the digital elevation model

Figure 15: The digital elevation model and its reclassification used in the current work.

### 2.3 THE FREE AND OPEN SOURCE SOFTWARE USED FOR THE ANALYSIS

The harmonization of such diverse array of data, historical maps, aerial images, remote sensed information and historical archives, required a long time of this research project and the developments of different approaches according to the peculiarity of each source.

The datasets were achieved, processed and analysed using the Free and Open Source Software (FOSS) GIS GRASS and QGIS. The versatility of the wide variety of modules supplied from the FOSS GRASS and QGIS enabled to perform a diverse set of analysis and pre-processing (e.g.: orthorectification) on an heterogeneous dataset of input images.

The following paper describes the Open Source environment and introduces some of the procedures used to develop the current study [47, 82, 21]. This work was published in the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, volume 42 4W14, in 2019 [45].

## FINE SPATIAL SCALE MODELLING OF TRENTINO PAST FOREST LANDSCAPE (TRENTINOLAND): A CASE STUDY OF FOSS APPLICATION

S. Gobbi<sup>1,2,3\*</sup>, M.G. Cantiani<sup>1</sup>, D. Rocchini<sup>2,4,6</sup>, P. Zatelli<sup>1</sup>, C. Tattoni<sup>1</sup>, N. La Porta<sup>3,5</sup>, M. Ciolli<sup>1,3</sup>

<sup>1</sup> University of Trento, Dep. of Civil, Environmental and Mechanical engineering, Trento, Italy  
(stefano.gobbi, maria.cantiani, paolo.zatelli, clara.tattoni, marco.ciolli)@unitn.it

<sup>2</sup> Fondazione Edmund Mach, San Michele all'Adige (TN), Italy - duccio.rocchini@fmach.it

<sup>3</sup> The EFI Project Centre on Mountain Forests (MOUNTFOR), San Michele a/Adige, Trento, Italy

<sup>4</sup> Centro Agricoltura, Alimenti e Ambiente, San Michele all'Adige (TN), Italy

<sup>5</sup> IASMA Research and Innovation Centre, Fondazione Edmund Mach, San Michele a/Adige, Trento, Italy - nicola.laporta@fmach.it

<sup>6</sup> Centro Biologia Integrata (CIBIO), University of Trento

### Commission IV, WG IV/4

**KEY WORDS:** Remote sensing, mapping, orthophotos, GRASS GIS, QGIS, Trentino, Forestry, Ecology

### ABSTRACT:

Trentino is an Italian alpine region (about 6200  $Km^2$ ) with a forest coverage exceeding 60% of its whole surface. In the past, forest landscape has changed dramatically, especially in periods of forest over-exploitation.

Previous studies in some Trentino sub-regions (Val di Fassa, Paneveggio) have identified these changes and the current trend of forest growth at the expenses of open areas, such as pastures and grasslands, due to the abandonment of rural areas. This phenomenon leads to the rapid Alpine landscape change and profoundly affects the ecological features of mountain ecosystems. To be able to monitor and to take future actions about this trend it is fundamental to know in detail the historical situation of the progressive changes on the land use that occurred over Trentino.

The work aims to comprehensively reconstruct the forest cover of whole Trentino at high resolution (5m x 5m pixels) using a series of maps spanning a long period, consisting in historical maps, aerial images, remote sensed information and historical archives. The datasets were archived, processed and analyzed using the Free and Open Source Software (FOSS) GIS GRASS and QGIS. Historical maps include Atlas Tyrolensis (dated 1770), Theresianischer Kataster (dated 1859) and Italian Kingdom Forest Map (IKFM) of 1936. The aerial imagery dataset includes aerial images taken in 1954, which have been orthorectified during this research, and orthophotos available for years 1973, 1994, 2000, 2006, 2010 and 2016. Remote sensed information includes Landsat and recent Lidar data, while historical archives consist mostly in Forest Management Plans available since around 1950.

The versatility of the wide variety of modules supplied from the FOSS GRASS and QGIS enabled to perform a diverse set of analysis and pre-processing (e.g.:orthorectification) on a heterogeneous dataset of input images. We will focus on the different strategies and methodologies implemented in the FOSS GIS used to process the various types of geographic data, challenges for the future of the research and the fundamental role of the FOSS systems in this process.

Quantifying forest change in the time-span of our dataset can be used to perform further analysis on ecosystem services, such as protection from soil erosion, and on modification of biome diversity and to create future change scenarios.

## 1. INTRODUCTION

Provincia Autonoma di Trento (PAT or Trentino) is an alpine region of the north of Italy. Its surface is about 6200  $km^2$ , with a strong presence of forests, characterized by a variety of different types. In fact, the presence of the Alps shaped a whole variety of different habitats: while the lowest point of PAT is at 64 m.a.s.l. (Valle del Sarca) the highest point is Monte Cevedale (3769 m.a.s.l.) (Sitzia, 2009). The importance of the natural environment in Trentino is underlined by the presence of a national park (Parco Nazionale dello Stelvio), two regional parks (Parco dell'Adamello Brenta, Parco delle dolomiti di Paneveggio) and 142 habitats in the Natura2000 network.

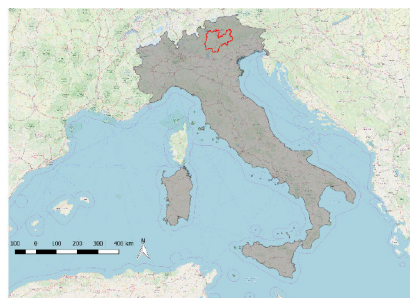


Figure 1. The position of Trento Province within the Italian territory

\*Corresponding author

Forest in PAT are characterized by a strong presence of coniferous species, mostly spruce (Servizio Fauna e Foreste, 2018).

### Forest types in Trentino

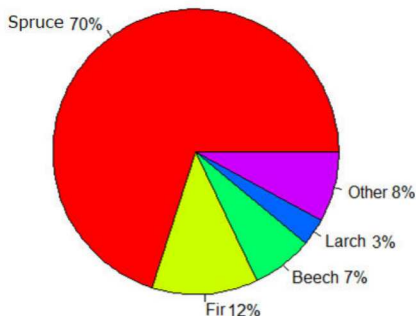


Figure 2. Principal forest types in PAT (Servizio Fauna e Foreste, 2018)

During the centuries human activities has shaped the Trentino landscape. While activities such as timber harvesting and agriculture, which required to be near the townships, shaped the valley bottoms, the traditional "malga" system for pasture shaped the high mountain meadows (MacDonald et al., 2000).

During World War I the position of Trentino was strategic for both the Austro-Hungarian Empire and the Italian Kingdom, so Trentino was the ground for long and weary battles. In this period both the involved armies exploited heavily the timber, contributing to the deforestation of the Alps (Tattoni et al., 2010). After World War II the socio-economical situation changed again: extreme poverty pushed people to leave villages in the valleys towards the two main towns (Trento and Rovereto), abandoning the traditional agricultural and pastoral activities, consequentially abandoning meadows and crop fields to a progressive afforestation (De Natale et al., 2005), (Sitzia et al., 2007), (Sitzia, 2009), (Tattoni et al., 2017).

All these changes favored some types of habitats and some tree and animal species, while other species were negatively influenced by the abandonment of human activities in the mountains (Tattoni et al., 2010). The problem is not just concerning the PAT region, but it is recognized at a European Level: the Pan-European biological and landscape diversity strategy, the Bern Convention, the European Landscape Convention, the Birds and Habitats Directives are some of the acts taken by the European Union to preserve high mountain landscapes (Cantiani et al., 2016).

In this framework it becomes crucial the knowledge of past forest landscape in PAT, in order to identify the evolution and the rate of change of some specific habitats through the investigated territory (Ciolli et al., 2012). A multi-temporal map analysis was carried out in the whole Trentino territory. In part 2 we will explain more in detail the features of the datasets involved, which cover a time span from 1859 to 2015. The Free and Open Source Software for Geography GRASS GIS has been used, exploiting its multi-purpose potential to process such a diverse and wide variety of data, as explained in parts 2 and 4. Focus of the current work will be given on the GRASS module used, while some already published results for limited areas of

Trentino will be presented, to show how this kind of study can address political and environmental decision over the conservation of specific habitats (Ministero dell'ambiente, 2010).

## 2. MATERIALS AND METHODS

The availability of aerial and satellite imagery in recent years and the development and free distribution of powerful GIS software made easier to perform remote sensing inquiries over a specific territory (Rocchini et al., 2012) (Tattoni et al., 2010) (Neteler, Mitasova, 2008). Techniques such as image classification are a consolidated method for studying multi-temporal forest evolution (Gaitanis et al., 2015) (Godone et al., 2014) (Gautam et al., 2004).

Here we will briefly present the cartographic data used and the algorithm implemented to carry out the analysis of PAT territory over the years.

### 2.1 The 1859 cadastral maps



Figure 3. An example of the 1859 maps: the depiction of the town of Trento.

The Austro-Hungarian cadastre was drawn by the Imperial edict of Franz Joseph I von Österreich in 1817. The part of PAT was completed in 1859 and it is divided in sheets (13300 sheets, approximately), in scale 1:2880 (Servizio Catasto della Provincia Autonoma di Trento, 2019a). In 2006 the procedure of digitalization and georeferentiation of the whole cadastre was completed. Each sheet is available in UTM32N-ETRS89 reference system and JPG format (Servizio Catasto della Provincia Autonoma di Trento, 2019b). The map represents a thematic map; areas with different colours represent different land use (figure 4):

1. grey for the forests;
2. light green for the private/public pastures;
3. green for the crops;
4. blank for the unproductive lands.

Water bodies, construction and roads are as well clearly recognizable and, if they are relevant, named. Information about the density and type of forests are missing.

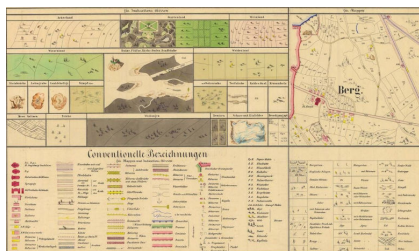


Figure 4. The legend for the 1859 cadastral maps (in German).

## 2.2 The 1936 Italian Kingdom Forest Map

The Italian Kingdom Forest Map (IKFM) was produced in 1936 by the Milizia Italiana Forestale, to map the presence of forests in the Italian territory. The cartographic base was the map in scale 1:100 000 produced by the Istituto Geografico Militare Italiano, now available in one of the Gauss-Boaga/Rome40 datum two zones (west and east, EPSG 3003 and 3004), depending on the map sheet location (Ferretti et al., 2018).

This map is considered the first example of forest and forest types mapping for the whole Italian Territory (Ferretti et al., 2018). In fact the legend reports a total of 25 different forest types (Ferretti et al., 2018).

The IKFM was digitalized to be preserved and consulted freely. Each sheet was scanned and georeferenced in a TIFF format. Each sheet has an approximate dimension of 112 MB, with a resolution of 400 ppi (pixel per inch); as for the georeferencing process it was decided to keep the original Gauss-Boaga/RM40 reference system (Ferretti et al., 2018). The map has been classified and vectorized by manually digitizing each forest patch (Ferretti et al., 2018). The resulting map is available under the Creative commons Attribution - Version 3.0 license and can be viewed and downloaded from a dedicated webgis (Ferretti et al., 2019).

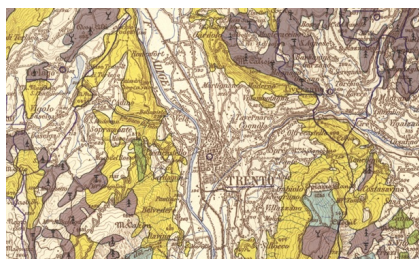


Figure 5. An example of the 1936 maps: the depiction of the town of Trento.

## 2.3 The single band aerial imagery of PAT



Figure 6. An example of the 1994 aerial images: the depiction of the town of Trento.

To perform the current study different sets of imagery were chosen, in particular the images from 1954, 1973 and 1994 that are coded in single band greyscale. The set of aerial photographs from 1954 (denominated "Volo GAI", Gruppo Aeronautico Italiano) in particular, has been orthorectified during this project, as explained in part 2.6. The table 1 shows details about resolution, datum and number of images used to depict PAT territory (Geoportale Provincia Autonoma di Trento, 2019a) (Geoportale Provincia Autonoma di Trento, 2019b).

Year	Pixel dimension	Images in set	Datum
1954	2x2m	229	WGS84- UTM32N
1973	1x1m	229	Gauss Boaga-RM40
1994	1x1m	229	Gauss Boaga-RM40

Table 1. Features of the greyscale sets of imagery of Trentino region, pixel dimension on the ground.

## 2.4 The multi-band aerial imagery of PAT

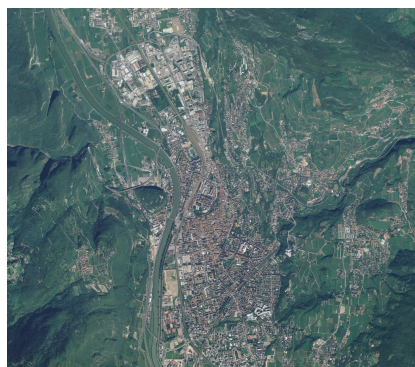


Figure 7. An example of the 2006 aerial images: the depiction of the town of Trento.

The multi band imagery of PAT used in the current work were taken in 2006 and 2015. Table 2 reports details about resolution,

datum and number of images used to represent PAT territory (Geoportale Provincia Autonoma di Trento, 2019c) (Geoportale Provincia Autonoma di Trento, 2019d).

Year	Pixel dimension	Images in set	Datum
2006	0.5x0.5m	229	WGS84- UTM32N
2015	0.2mx0.2m	839	WGS84- UTM32N

Table 2. Features of the multi-band sets of imagery of Trentino Territory, pixel dimension on the ground.

These newer images have a better geometric resolution and at least 3 bands (Red, Green and Blue). In particular the images of 2015 set include the additional near-infra-red band. The downside of this amount of additional information is the dimension on the computer memory required: while the dataset of 2006 has a dimension of approximate 100 GB, the dataset of 2015 has a dimension of approximate 780 GB.

## 2.5 GRASS GIS

GRASS GIS has been chosen to manage and perform the images processing. The software was first developed by the U.S. Army Construction Engineering Research Laboratories (USACERL, 1982-1995). Starting from 1999 to nowadays GRASS is released under a GNU license by the Open Source Geospatial Foundation (OSGeo) (Neteler, Mitasova, 2008). The architecture of GRASS is inspired by a modular structure, involving more than 350 different modules, which can be stacked together to perform complex analysis (Neteler et al., 2012). Single developers and researchers are welcome to study and develop new modules to fit their purposes (Preatoni et al., 2012), and they can be possibly saved in a on-line package repository, to be downloaded by other researchers all over the world. It is worth remembering that both the source code of GRASS and its modules are accessible: this fact is crucial to ensure robust analysis output (Rocchini et al., 2012) and its suitability for educational purposes (Ciolli et al., 2017). Moreover, GRASS can be scripted using Python programming language: this allows to stack different modules and apply identical procedures to a great number of input imagery (Van Rossum, 1995).

In the current work the versions of GRASS used are the 7.4 and the 6.4, as explained in part 2.6 (GRASS Development Team, 2017) (GRASS Development Team, 2018).

## 2.6 Orthorectification of the datasets of 1954

Orthorectification is the process of adapting a flat image to a rugged and curved surface, by adapting the reference and projection system (Gobbi et al., 2018). In this case-study the dataset of 1954 imagery required this pre-processing because only the original images are available.

This process is performed in three different steps:

1. internal orientation to evaluate the position of the image with respect to the camera frame;
2. external orientation to evaluate the position of the camera with respect to the external reference system (the chosen datum);
3. orthorectification to re-project the image.

To perform the first step it is necessary to measure and identify the position of 4 or more fiducial markers on the original photograph. To perform the second step is necessary to set a consistent number of Ground Control Points, or points whose coordinates are known in both the reference systems of the original image and the target reference system (Novak, 1992). Finally the re-projection is performed using a set of equations, called collinearity equations, which rectify the original image by shifting, rotating and scaling each of its pixel (Novak, 1992) (Gobbi et al., 2018) (Rocchini et al., 2012). A DEM which describes the geometry of the ground surface must be available.

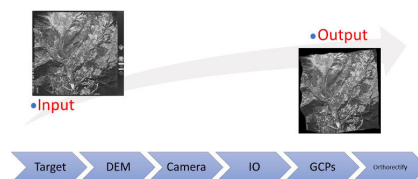


Figure 8. Flowchart of the modules and data in GRASS GIS for orthorectification.

In this study the versions used for orthorectification are the 6.4 and 7.4 versions. In fact the suite *i.ortho.photo* was removed from the versions between the 6.4 and the 7.2.

In GRASS 6.4 the process is divided in different steps:

1. *i.target* allows to chose a target reference system;
2. *i.ortho.elev* allows to set a Digital Elevation Model used to correct the position of points according to the orography;
3. *i.ortho.camera* allows to set the parameters of the camera which took the picture (e.g. the focal length);
4. *g.gui.photo2image* allows to input the position of the fiducial markers;
5. *g.gui.image2target* allows to input the position of the GCPs;
6. *i.ortho.rectify* performs the actual rectification.

After inserting manually the position of fiducial markers and GCPs it was possible to use Python to apply the orthorectification to a whole set of imagery inside a single mapset. It is possible to set a raster mask: this is useful if, as it is in this study, the original image has a frame, containing instruments reading, which must be removed in the orthorectified image. Due to the user-friendly interface, GRASS 7.4 was chosen to perform the input of the GCPs: this choice was considered less time consuming than the using the old 6.4 GRASS. On the other hand GRASS 7.4 did not allow the setting of a mask for excluding some parts of the image (the frame, in our case study) therefore it was chosen to run the *i.ortho.rectify* command on GRASS 6.4.

## 2.7 Landuse classification algorithms

To reconstruct the forest landscape it is necessary to perform a landuse classification for each dataset. This process generates a map where each "pixel" or "object" is grouped in a finite number of set, each one representing a type of land use. Typically



there are 5 macro-categories: urbanized, forest, agriculture, water, unproductive.

There are two families of algorithm: the maximum likelihood family and the OBIA family.

The maximum likelihood algorithms require as input at least one area for each class of landuse, where the software running the algorithm calculates a statistical distribution of the spectral response of the pixels within the area. Then each pixel of the image is classified inside different macro-class by checking which class its spectral response statistically belongs to (Bouman, Shapiro, 1992).

Object-based Image Analysis (OBIA) takes a different approach: instead than classifying each single pixel, OBIA creates groups of pixels, called segments, which are classified as single objects using machine learning. The parameters used by the machine learning are the statistical distribution of the pixels radiometric response inside the single object and the geometry of the same object (perimeter, area, compact circle, compact square, fractal index) (Clewley et al., 2014). Each segment of pixels is called "object" and it is characterized by pixels with a similar spectral response. The choice between the two algorithm should be driven by the resolution of the input image. If the objects (houses, trees, roads, water bodies) depicted in the image are larger than the pixel resolution an OBIA approach is more likely to give a nicer, cleaner land use output map (Burnett, Blaschke, 2003). If the objects depicted are smaller than the resolution (e.g. the resolution 30x30m of some Landsat images) a maximum likelihood algorithm is preferable, because it is possible that similar pixels will be grouped inside the same object while in reality they depict two different things.

In the current case study the resolution of the aerial imagery is always smaller than the represented objects (tables 1 and 2) and the OBIA approach was preferred.

In GRASS the procedure is not coded as a suite of command as it was for the orthorectification process, so it was necessary to create a procedure which differs for the multi-band and the single band imagery (Grippa et al., 2017).

Some common steps are required:

1. *i.segment* performs the segmentation of the imagery;
2. *r.texture* *r.texture* evaluates the textural differences, this provides an additional band which give more information about each single object;
3. *i.segment.stats* evaluates the radiometric features of each segment and store them in the table associated to the (vector) output map;
4. *v.class.mlR* uses machine learning to classify the object inside the chosen land-use categories;

Figures 10 and 9 show the flowchart with the input data for the case of single band imagery (1954, 1973, 1994) and colour imagery (2006, 2015). The textural measure was considered in both cases as additional image band, because where it is necessary to discriminate between forests and crop fields the spectral response is similar, but the textural measure within the same object differs, giving to the machine learning helpful data to discriminate the two situations (Haralick et al., 1973).

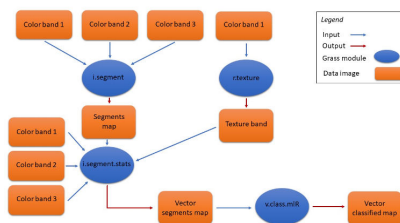


Figure 9. Flowchart of the modules and data in GRASS GIS for OBIA of multi-band images.

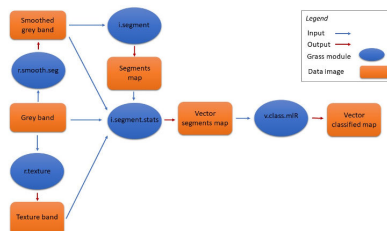


Figure 10. Flowchart of the modules and data in GRASS GIS for OBIA of single grey-band images.

The command *r.smooth.seg*, applied on the single band imagery performs a preliminary segmentation on the image by smoothing values of pixels within the same object and adding more contrast to the boundaries of the segments (Vitti, 2012). For this reason it was considered more reliable to apply this module to the image before using *i.segment*. The output from *i.smooth.seg* was used as input to the machine learning as well.

The procedure was scripted with Python and applied sequentially on each map inside the different datasets.

## 2.8 Manual classification of 1859 cadastral maps

The dataset of 1859 is composed by 13300 sheets, and applying to each one of them a OBIA procedure for classifying the maps would have costed too much in terms of computational time. Instead the maps were classified using a manual procedure. The procedure was performed in GRASS version 7.4 and QGIS version 2.18 (QGIS Development Team, 2019), and involved the use of the vectorizer instrument.

The vectorized areas where the ones indicated as "forest" (with attribute "1" in the attribute table), "pasture" (with attribute "2" in the attribute table) and "wooded pasture" (with attribute "3" in the attribute table). A few simple rules were followed:

1. if two or more neighbour cadastral parcel had the same landuse they were digitized as one single area;
2. assign a single identifier value to each area;
3. digitize following the black border of each area and include the border in the area;
4. exclude lakes and rivers from the digitalization process.

Year	Denomination	Scale	Type of map	Pixel dimension	Dimension [GB]	Datum
1859	<i>Theresiascher Kadastrte</i>	1:1440	Topographic map	0.2x0.2	20	ETRS89-UTM32N
1936	<i>Italian Kingdom Forest Map</i>	1:100000	Topographic map	0.32x0.32	1	Gauss Boaga RM40
1954	<i>Volo GAI</i>	1:35000	Aerial images BN	2x2m	8	WGS84- UTM32N
1973	<i>Volo Rossi-EIRA</i>	1:10000	Aerial images BN	1x1m	8	Gauss Boaga-RM40
1994	<i>Volo Italia</i>	1:10000	Aerial images BN	1x1m	8	Gauss Boaga-RM40
2006	<i>Volo Terraitaly</i>	1:5000	Aerial images RGB	0.5x0.5m	100	WGS84-UTM32N
2015	<i>Volo AGEA</i>	1:5000	Aerial images RGB	0.2x0.2m	730	WGS84-UTM32N

Table 3. The data collection for the current case-study.

### 3. RESULTS

**A complete dataset** The current work is leading to create to a complete collection of maps and aerial imagery of Provincia Autonoma di Trento, with a timespan of 160 years. In the perspective of study whose objective is to analyse landscape changes over the years having a complete dataset is a crucial results for the future of the work. The collected imagery covers the whole territory of Trentino, with a high-resolution and detailed scale, as shown in table 3.

**1859 maps** A partial result is the digitalization of the 1859 cadastral maps. By following the rules explained in section 2.8 areas of forests and open pastures were classified for the whole region of Trentino.

**Orthorectification of the 1954 dataset** A total amount of 92 aerial imagery of the dataset of 1954 aerial images were collected and orthorectified, integrating the imagery of (Cattani, 2015) (Tattoni et al., 2010) (Rocchini et al., 2012) (Maimeri, 2018) was possible to achieve a complete rectified set of imagery from the 1954 dataset for the whole PAT (Gobbi et al., 2018).

**Testing the OBIA procedure** The OBIA procedure has been calibrated and applied to the whole variety of aerial imagery in the dataset. This result is important because accomplish one of the aims of this project, i.e. to automatize the classification process as much as possible so that its application to large datasets covering the whole Trentino region is feasible (Tattoni et al., 2010) (Cattani, 2015) (Maimeri, 2018). The synergy between Python and GRASS rendered possible to automatize the OBIA in just 140 lines of code.

### 4. DISCUSSION AND CONCLUSIONS

**The use of GRASS and QGIS** In the course of this research the main tool has been GRASS. The availability of GRASS source code under the GPL was useful especially in the orthorectification phase, where it was crucial to understand how the errors in the GCPs were calculated, accessing the source code of the suite for orthorectification gave insight to the interpretation of the errors (Gobbi et al., 2018). As a matter of fact the two version of GRASS used (6.4 and 7.4) such errors were displayed with two different methods, and understand which one was more significant required accessing to the source code. The possibility to script GRASS with Python is a major advantage in managing large datasets (in terms of number of elements and dimension in GB, see table 3). Once the input parameters are set for a specific analysis, the same module can be sequentially applied to different images with the same parameters. It was possible to apply the same modules sequence in the OBIA procedure with different input parameters with minimum intervention from the user by scripting the algorithm. This approach

provides an efficient way of finding an optimum set of parameters of the land use classification procedure. The classification of the complete Trentino territory using OBIA is under way for all the available datasets with a scripted batch procedure. The effectiveness of this classification OBIA approach will be tested comparing the results with those obtained in the previous works (Cattani, 2015) (Tattoni et al., 2010) (Maimeri, 2018) with the OBIA approach.

Finally, the use of QGIS and its digitizing tools plugin prevented topological errors in the digitalization process of the 1859 dataset, assuring that the created vector files were viable for further analysis.

**The future of the research** In the current study, the reconstruction of the past forest landscape from landuse classification will be used for landscape and ecosystem services analysis (Ciolli et al., 2019). Following (Ciolli et al., 2012) and (Tattoni et al., 2010), the landuse maps will be used to evaluate landscape metrics using Fragstat or the LeCos suite in QGIS to evaluate parameters that can describe how the forest coverage has changed during the years. Metrics such as the mean forest patch area, for instance, can represent how the forests became more fragmented or compact. This information helps to give an overall view on the fluctuation of the total forest coverage and its density (Campagnaro et al., 2017) (Tattoni et al., 2017) and it will be crucial for mapping the hydro-geological risk (Ciolli et al., 2019).

The second field of application of the forest coverage data is the use of predictors such as Markov Chains to simulate how protected habitats will evolve if the "no human intervention" policy is applied, as specified in (Ciolli et al., 2012).

A total of 142 Sites of Community Importance have been recognized in Trentino: the fate of this areas can be easily protected by human intervention and this study can help to program the actions that can be performed to protect such areas.

### ACKNOWLEDGMENTS

We would gladly thank Luca Delucchi from Fondazione Edmund Mach for the help with GRASS, in particular with the module for the orthorectification. All the aerial imagery were provided by the Provincia Autonoma di Trento, while the cadaster historical map for the Province of Trento has been made available by the Servizio Catasto della Provincia Autonoma di Trento, under the Creative Commons Attribution 4.0 license. Finally we would like to thank the Ufficio Fauna e Foreste, in the person of Maurizio Miori for providing all the data about forest types in Trentino.

## REFERENCES

- Bouman, C., Shapiro, M., 1992. Multispectral Image Segmentation using a Multiscale Image Model. *Proceedings of IEEE International Conference on Acoustics, Speech and Sig. Proc.*, 565–568. <http://dx.doi.org/10.1109/ICASSP.1992.226150>.
- Burnett, C., Blaschke, T., 2003. A multi-scale segmentation/object relationship modelling methodology for landscape analysis. *Ecological Modelling*, 168, 233–249.
- Campagnaro, T., Frate, L., Carranza, M.L., Sitzia, T., 2017. Multi-scale analysis of alpine landscapes with different intensities of abandonment reveals similar spatial pattern changes: implication for habitat conservation. *Ecological Indicators*, 147–159. <https://doi.org/10.1016/j.ecolind.2016.11.017>.
- Cantiani, M.G., Geitner, C., Haida, D., Maino, F., Tattoni, C., Vettorato, D., Ciolli, M., 2016. Balancing economic development and environmental conservation for a new governance of Alpine areas. *Sustainability*, 802–820. <http://dx.doi.org/10.3390/su8080802>.
- Cattani, G., 2015. Metodologia per la valutazione della perdita di biodiversità dovuta al cambiamento del paesaggio agroforestale: il caso del tesino. Master's thesis, University of Trento.
- Ciolli, M., Bezzi, M., Comunello, G., Laitempergher, G., Gobbi, S., Tattoni, C., Cantiani, M.G., 2019. Integrating dendrochronology and geomatics to monitor natural hazards and landscape changes. *Applied Geomatics*, 11, 39–52. <https://doi.org/10.1007/s12518-018-0236-0>.
- Ciolli, M., Federici, B., Ferrando, I., Marzocchi, R., Sguerso, D., Tattoni, C., Vitti, A., Zatelli, P., 2017. FOSS tools and applications for education in geospatial sciences. *ISPRS International Journal of Geo-Information*, 6.
- Ciolli, M., Tattoni, C., Ferretti, F., 2012. *Understanding forest changes to support planning: A fine-scale Markov chain approach*. in F. Jordán, S. Jorgensen (edited by), *Models of the ecological hierarchy from molecules to the ecosphere*. Elsevier, Amsterdam. <http://dx.doi.org/10.1016/B978-0-444-59396-2.00021-3>.
- Clewley, D., Bunting, P., Shepherd, J., Gillingham, S., Flood, N., Dymond, J., Lucas, R., Armston, J., Moghadam, M., 2014. A Python-Based Open Source System for Geographic Object-Based Image Analysis (GEOBIA)utilizing raster attribute tables. *Remote sensing*, 6, 6111–6135. <https://doi.org/10.3390/rs6076111>.
- De Natale, F., Gasparini, P., A., Carriero, 2005. A study on tree colonization of abandoned land in the Italian Alps: extent and characteristics of new forest stands in Trentino. *Proceedings of the IUFRO Congress Sustainable forestry in theory and practice*.
- Ferretti, F., Sboarina, C., Tattoni, C., Vitti, A., Zatelli, P., Geri, F., Pompei, E., Ciolli, M., 2018. The 1936 Italian Kingdom Forest Map reviewed: a dataset for landscape and ecological research. *Annals of Silvicultural Research*, 42, 3–19. <https://doi.org/10.12899/asr-1411>.
- Ferretti, F., Sboarina, C., Tattoni, C., Vitti, A., Zatelli, P., Geri, F., Pompei, E., Ciolli, M., 2019. Carta forestale del regno di italia. <http://carta1936.dicam.unitn.it>. Accessed: 2019-04-29.
- Gaitanis, A., Kalogeropoulos, K., Detsis, V., Chalkias, C., 2015. Monitoring 60 Years of Land Cover Change in the Marathon area, Greece. *Land*, 4, 337–354. <https://doi.org/10.3390/land4020337>.
- Gautam, A.P., Shivakoti, G.P., Webb, E.L., 2004. Forest cover change, physiography, local economy, and institutions in a mountain watershed in Nepal. *Environment Management*, 1, 48–61. <https://doi.org/10.1007/s00267-003-0031-4>.
- Geoportale Provincia Autonoma di Trento, 2019a. Metadati volo 1973. tn:8cf5ffba-f738-4e60-a0e6-7669919eec24. Accessed: 2019-04-29.
- Geoportale Provincia Autonoma di Trento, 2019b. Metadati volo 1994. tn:629ccdc8-e046-411c-95b0-326cb0bdc656. Accessed: 2019-04-29.
- Geoportale Provincia Autonoma di Trento, 2019c. Metadati volo 2006. tn:a1dfd067-1f7d-40d1-b873-fc665a61af6b. Accessed: 2019-04-29.
- Geoportale Provincia Autonoma di Trento, 2019d. Metadati volo 2015. tn:f2e88f1b-05d9-4942-93ee-857a0a9e1f0b. Accessed: 2019-04-29.
- Gobbi, S., Maimeri, G., Tattoni, C., Cantiani, M.G., Rocchini, D., La Porta, N., Ciolli, M., 2018. Orthorectification of a large dataset of historical aerial images: procedure and precision assessment in an open source environment. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-4/W8, 53–59. <https://doi.org/10.5194/isprs-archives-XLII-4-W8-53-2018>.
- Godone, D., Gabarino, M., Sibona, E., 2014. Progressive fragmentation of a traditional Mediterranean landscape by hazelnut plantations: The impact of CAP over time in the Langhe region (NW Italy). *Land use Policy*, 48–61. <https://doi.org/10.1016/j.landusepol.2013.08.018>.
- GRASS Development Team, 2017. Geographic Resources Analysis Support System (GRASS) Software, Version 6.4. Open Source Geospatial Foundation. [grass.osgeo.org](http://grass.osgeo.org) (1 November 2017).
- GRASS Development Team, 2018. Geographic Resources Analysis Support System (GRASS) Software, Version 7.4. Open Source Geospatial Foundation. [grass.osgeo.org](http://grass.osgeo.org) (20 September 2018).
- Grippa, T., Lennert, M., Beaumont, B., Vanhuyse, S., Stephenne, N., Wolff, E., 2017. An Open-Source Semi-Automated Processing Chain for Urban Object-Based Classification. *Remote Sensing*, 4, 358–378. <https://doi.org/10.3390/rs9040358>.
- Haralick, R. M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. *IEEE transactions on systems, man and cybernetics*, 3, 610–621.
- MacDonald, D., Crabtree, J., Wiesinger, G., Dax, T., Stamou, N., Fleury, P., Lazpita, J., Gibon, A., 2000. Agricultural abandonment in mountain areas of Europe: environmental consequences and policy response. *Journal of Environmental Management*, 59, 47–69. <https://doi.org/10.1006/jema.1999.0335>.
- Maimeri, G., 2018. Variazione della copertura forestale in val di fassa tramite analisi gis multitemporale. Master's thesis, University of Trento.

- Ministero dell'ambiente, 2010. Elenco ufficiale delle aree protette (euap). Sixth update, approved the 27th of April 2010, published in the official Gazette (125).
- Neteler, M., Bowman, M.H., Landa, M., Metz, M., 2012. GRASS GIS: A multi-purpose open source GIS. *Environmental Modelling and Software*, 31, 142–130. <https://doi.org/10.1016/j.envsoft.2011.11.014>.
- Neteler, Makus, Mitasova, Helena, 2008. *Open Source GIS: A GRASS GIS Approach*. Springer, New York.
- Novak, Kurt, 1992. Rectification of Digital Imagery. *Photogrammetric Engineering and Remote Sensing*, 58, 339–344.
- Preatoni, D.G., Tattoni, C., Bisi, F., Masseroni, E., D'Acunto, D., Lunardi, S., Grimod, I., Martinoli, A., Tosi, G., 2012. Open source evaluation of kilometric indexes of abundance. *Ecological Informatics*, 35–40. <http://dx.doi.org/10.1016/j.ecoinf.2011.07.002>.
- QGIS Development Team, 2019. QGIS Geographic Information System, version 2.18. Open Source Geospatial Foundation.
- Rocchini, Duccio, Metz, Markus, Frigeri, Alessandro, Delucchi, Luca, Marcantonio, Matteo, Netler, Markus, 2012. Robust rectification of aerial photographs in an open source environment. *Computers and geosciences*, 39, 145–151. <https://doi.org/10.1016/j.cageo.2011.06.002>.
- Servizio Catasto della Provincia Autonoma di Trento, 2019a. History of the cadaster historical map for the Province of Trento. [http://www.catasto.provincia.tn.it/cenni\\_storici/pagina8.html](http://www.catasto.provincia.tn.it/cenni_storici/pagina8.html). Accessed: 2019-04-29.
- Servizio Catasto della Provincia Autonoma di Trento, 2019b. Mappe storiche di impianto (urmappe). <https://www.catastotn.it/mappeStoriche.html>. Accessed: 2019-04-29.
- Servizio Fauna e Foreste, 2018. *Schede dei tipi forestali, caratteristiche e indicazioni gestionali*. 1st edn, Provincia Autonoma di Trento, Trento, Via Trener 3.
- Sitzia, T., 2009. *Ecologia e gestione dei boschi di neoforazione nel paesaggio trentino*. Servizio Foreste e Fauna, Provincia Autonoma di Trento, Trento, Italy, pp. 301.
- Sitzia, T., Carriero, A., De Natale, F., Gasparini, P., Wolynski, A., Viola, F., 2007. Recent secondary woodlands in a regional sample of southern-alpine abandoned landscapes: implications for restoration ecology and silviculture. *Proceedings of the seventh IALE World Congress*, 2, 783–784.
- Tattoni, C, Ciolli, M, Ferretti, F, Cantiani, M.G., 2010. Monitoring spatial and temporal pattern of Paneveggio forest ( northern Italy ) from 1859 to 2006. *iForest*, 3, 72–80. <https://doi.org/10.3832/ifor0530-003>.
- Tattoni, C., Ianni, E., Geneletti, D., Zatelli, P., Ciolli, M., 2017. Landscape changes, traditional ecological knowledge and future scenarios in the Alps: A holistic ecological approach. *Science of the Total Environment*, 27–36. <http://dx.doi.org/10.1016/j.scitotenv.2016.11.075>.
- Van Rossum, G., 1995. *Python Library Reference*. CWI Report.
- Vitti, A., 2012. The Mumford-Shah variational model for image segmentation: An overview of the theory, implementation and use. *ISPRS Journal of Photogrammetry and Remote Sensing*, 69, 50–64. <https://doi.org/10.1016/j.isprsjprs.2012.02.005>.

## 2.4 PREPARATION OF THE IMAGES: ORTHORECTIFICATION OF 1954 DATASET

Raw aerial photographs are not suitable for analysis because they are not referenced in a geographical reference system. Image orthorectification must be performed to transform the images in maps available for the analysis. This procedure corrects the geometry according to the terrain surface described by a Digital Terrain Model (DTM) to create an image compatible with the cartographic projection in use.

The following paper describes in detail the process of orthorectification applied in GRASS GIS to transform the 1954 aerial photographs in aerial imagery with a geographical reference system [92]. Once the 1954 images set were orthorectified they were fit for the landscape analysis. Orthorectification is an important phase of this study, because resulted in a set of digital maps with a geographical reference. Before this work the images of 1954 for the PAT existed only in paper format and could not be used in a GIS. Since orthorectification is not a linear transformation the level of accuracy of the resulting rectified images will be discussed as well to provide information about the geographical reliability of the data. This is another crucial part of the analysis because for the future analysis of the forest area it is necessary having a spatially correct data. Other details about accuracy assessment will be added in section 3.1.

The following work was published in the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, volume 42 4W8, in 2018[44].

## ORTHORECTIFICATION OF A LARGE DATASET OF HISTORICAL AERIAL IMAGES: PROCEDURE AND PRECISION ASSESSMENT IN AN OPEN SOURCE ENVIRONMENT

S. Gobbi<sup>1,\*</sup>, G. Maimeri<sup>1</sup>, C. Tattoni<sup>1</sup>, M.G. Cantiani<sup>1</sup>, D. Rocchini<sup>1,2</sup>, N. La Porta<sup>2</sup>, M. Ciolli<sup>1</sup>, P. Zatelli<sup>1</sup>

<sup>1</sup> Università degli Studi di Trento, Dipartimento di Ingegneria Civile, Ambientale, Meccanica, Via Mesiano 77, Trento, Italy

<sup>2</sup> Fondazione Edmund Mach, via Edmund Mach 1, San Michele all'Adige (TN), Italy

### Commission IV, WG IV/4

#### ABSTRACT:

The availability of data time series spanning a long period is crucial for landscape change analysis. A suitable dataset, both in terms of time span and information content, must be available for the use with a GIS.

In Italy, one of the most important historical source of land cover analysis is the GAI (Gruppo Aereo Italiano) photogrammetric survey ("Volo GAI") commissioned in 1954 by the Italian national mapping agency, Istituto Geografico Militare Italiano (IGMI).

The survey covers the whole Italy, but so far only some Regions, namely Lombardia and Veneto, have carried out the image rectification and the successive analyses to map land cover and use.

This work describes the process of image orthorectification of the Volo GAI images for the Province of Trento (Provincia Autonoma di Trento).

Image orthorectification must be performed to transform the images in maps available for analysis. This procedure corrects the geometry according to the terrain surface described by a Digital Terrain Model (DTM) to create an image compatible with the cartographic projection in use.

To this end, the orthorectification modules available in GRASS GIS have been used, with the advantage of using the same GIS environment which will be used for the landscape analysis.

The dataset covering the whole Province contains almost 100 images, this paper presents the preliminary results of the orthorectification of a quarter of the images. A reduced dataset has been used to test the results obtained using different settings with respect to: digital image resolution, DTM resolution and number of Ground Control Points (GCPs) used for the external orientation.

These preliminary tests show that for the average quality of the Volo GAI images scan resolution beyond 600 DPI and DTM resolution above 10m do not provide significant improvements for orthorectification images. The minimum number of GCPs to guarantee the requested accuracy can vary from image to image, depending on the image quality and recognizable features position, but it is usually in the 15-20 points range.

## 1. INTRODUCTION

### 1.1 A brief introduction to photogrammetry

Historic data series of aerial photographs represent nowadays one of the most important source of geographical information (Rocchini et al., 2012) to perform multitemporal comparison of land use change and therefore detect landscape changes (Tattoni et al., 2010), (Ciolli et al., 2012), (Tattoni et al., 2017). However, aerial images present the same problem as topographic maps: the process of embedding Earth surface into a plan introduces distortions, because we are forcing a rugged and curved surface to fit on a two dimension flat plan; furthermore, photographs are sensible to the atmospheric conditions and light exposure (Cohen et al., 1996). Geometric errors alter the perceived position and size of some feature: tilt displacement is due to inclination of the focal plan with respect to the ground and relief displacement is due a change in elevation (Paine and Kiser, 2003). Figure 1 shows a clear example these distortions: the road in the non processed image (left) appears to be crooked, while is actually straight (processed image on the right) (OSSIM Development Team, 2015).

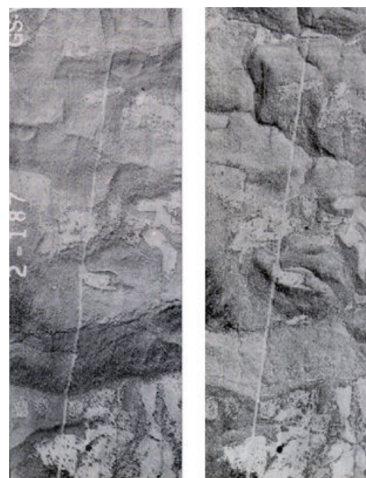


Figure 1. A comparison between a distorted image (left) and the same, rectified image (right). The road appears to be crooked on the left image, while in reality it is straight (right, processed image), courtesy from OSSIM Wiki. (OSSIM Development Team, 2015)

\*Corresponding author: stefano.gobbi@unitn.it

Radiometric errors are distortions in tone or colour of the photographs. They can be caused by vantage points, condition and calibration of the camera or by the post-processing of the image, during the filtering and film emulsion (Cohen et al., 1996) (Jensen, 1996). Photogrammetry is primary concerned in obtaining precise results from aerial photographs: in particular the process of compensating both geometric and radiometric errors is called orthorectification (Morgan et al., 2010). In this paper we will explain briefly elaboration of a large set of photographs, presenting a procedure implemented in GRASS GIS (Geographical Resources Analysis Support System, Geographic Information System) the Free and Open Source Software used for the procedure of orthorectification (GRASS Development Team, 2015). The frames used are part of the "Volo GAI", a photogrammetric survey commissioned in 1954. In particular we will focus on the survey carried out in Provincia Autonoma di Trento (PAT, a north eastern Italian region of 6.200 km<sup>2</sup>) that needed rectification before being used in GIS software.

## 2. MATERIALS AND METHODS

### 2.1 The dataset of 1954 Gruppo Aereo Italiano (GAI) aerial imagery

After World War II the Istituto Geografico Militare Italiano (IGMI) commissioned a series of flights over the whole Italy to survey the territory. The importance of this dataset lies in the fact that this is the first stereoscopic photographs depiction of the Italian region. The height of the flight was around 10.000 m a.s.l. for the mountainous areas, with a mean scale of 1:45.000 (Geoportale Regione Lombardia, 2018).

Each photograph is square shaped with a dimension of 230 x 230 mm and covers an area of about 140 km<sup>2</sup> (IGMI, 2016).

Currently IGMI still holds the rights for original data, but every Region of Italy acquired the digital version of these frames and it is possible for everyone interested to access the data (IGMI, 2016). Figure 1 shows an example of how frames are arranged. The fiducial markers of the frame, used for the internal image orientation, are highlighted in green. Parameters of the camera and of the flight era highlighted in blue: name of the camera, focal length of the camera lens, flight height, date and hour. The progressive number of the image in the survey is highlighted in red.

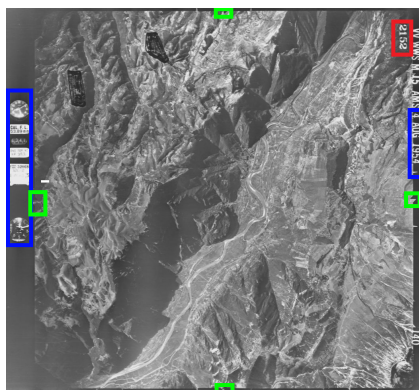


Figure 2. A non rectified photograph from 1954 GAI flight. For the meaning of highlighted areas we refer to paragraph 2.1.

The set of photographs used for the current work covers the entire Provincia Autonoma di Trento. All the photographs present very little or no cloud coverage. Shadowy areas are a major drawback: this is because lot of photographs were taken during the first hours of sunlight, when Sun is lower on the horizon and shadows are longer in mountainous areas.

The physical copy has been digitalized through scanner taking into account the final file size of the scanned image and its final mean ground resolution. The equation 1 was used to calculate ground resolution:

$$\frac{d}{f} = \frac{D}{H} \quad (1)$$

where  $d$  = pixel size [mm];  
 $f$  = focal length [mm];  
 $D$  = ground resolution [m]  
 $H$  = flight height [m a.s.l.].

### 2.2 Orthorectification

Orthorectification is the process of modifying the geometry of an image to make it compatible with a cartographic projection.

This procedure is performed in three steps:

**internal orientation** to evaluate the position of the image with respect to the camera;

**external orientation** to evaluate the position of the camera with respect to the external reference system;

**orthorectification** to re-project the image.

The first two steps use sets of equations whose parameters must be evaluated using external information, usually the coordinates of points recognizable on the image.

To perform the internal orientation the identification of the position of 4 or more fiducial markers on the original photograph is required.

A set of Ground Control Points (GCPs) is necessary to perform the external orientation. These GCPs are points of geometrical features on the image whose coordinates are known in a reference system. Novak (Novak, 1992) reviewed the most common used methods for orthorectification. Polynomial rectification consists in transforming the original image into the orthorectified one by means of polynomial equations:

$$\begin{aligned} x &= x'^T \mathbf{A} y' \\ y &= x'^T \mathbf{B} y' \end{aligned} \quad (2)$$

where  $x, y$  = original image coordinates;  
 $x', y'$  = rectified image coordinates;  
 $\mathbf{A}, \mathbf{B}$  = coefficient matrices of the polynomials.

The maximum order of these polynomials depends on how many GCPs are available.

Differential rectification is based upon a set of equations involving information about geometry of the camera sensor and the Digital Terrain Model (DTM) to correct for relief displacement. These equations, called collinearity equations, rectify the original image by shifting, rotating and scaling each of its pixel (Novak, 1992):

$$\begin{aligned} x &= x_p - c \frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} \\ y &= y_p - c \frac{r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} \end{aligned} \quad (3)$$

where  $c$  = focal length of the used camera;  
 $x, y$  = transformed image coordinates;  
 $X_0, Y_0, Z_0$  = coordinates of projection centre  
 $X, Y, Z$  = object coordinates of the original image;  
 $r_{xx}$  = orientation of the camera  
(components of the rotation matrix).

Usually the calibration certificate of the camera provides information about coordinates of the projection centre and focal length (Morgan et al., 2010).

The algorithm implemented in GRASS creates a grid on output orthorectified image and uses equations 3 to find the corresponding point on the original image. Every pixel of the grid is assigned a grey value (or color), using different interpolating methods (Novak, 1992):

1. starting from the original pixel location it transforms its coordinates and assigns the grey value to the nearest pixel (nearest neighbour);
2. starting from the original pixel location it transforms its coordinates and assigns the grey value interpolating the values of nearest pixels using bilinear methods;
3. starting from the original pixel location it transforms its coordinates and assigns the grey value interpolating the values of nearest pixels using cubic convolutions.

### 2.3 GRASS GIS procedure for orthorectification

The choice of a software based upon Free and Open Source philosophy (such as GRASS) allows the final user to access the source code behind every module (Neteler and Mitasova, 2008). It is worth noting that full access to the program code represents a major advantage with respect to a proprietary solution: the possibility of checking every step of an algorithm is useful to guarantee the appropriate robustness of the output (Rocchini et al., 2012). Furthermore, it is possible to modify the procedures and the algorithm according to the user requirements, provided that personnel with sufficient programming skills is available (Ciolli et al., 2017).

GRASS GIS version 7.4 includes a suite of commands for orthorectification similar to the one already implemented in the previous 6.4 version.

The main body of the suite is divided in modules, which can be invoked singularly or in sequence:

**l.ortho.target** This module sets a target location for the image to be rectified. The *target* location is a subset of maps in GRASS containing the maps used during the procedure (DTM and maps for GCPs) and implicitly defines the output ortho-image reference systems and projection.

**l.ortho.elev** This module allows the user to select a DTM for the ortho-rectification process. In the current study different DTMs with different resolutions (1m, 5m, 10m and 25m) have been used for testing purpose.

**l.ortho.camera** This module sets the parameters used for the internal orientation of the image. The focal length of the camera and the coordinates of the fiducial marks on the image must be provided. Since the calibration certificate of the camera is not available for the current dataset, the coordinates of have been determined measuring the distance between the marks on the physical copy of the photographs. These measures revealed a distance of 232 mm between east-west markers and 233 mm between north-south markers. The internal orientation is performed using a 6 parameters affine transformation, therefore the coordinates of at least 3 points must be provided.

**l.ortho.init** This module can be used to provide initial values (and their RMS) for the unknowns in the iterative least square adjustment for the evaluation of the external orientation parameters. The need to provide initial values of the 7 unknown parameters, representing 3 translations, 3 rotations and a scale factor, is due to the non linearity of the equations. Initial approximate values are automatically calculated. If no value is provided.

**g.gui.photo2image** This module provides a graphical interface to indicate the position of the fiducial markers on the digital image.

**g.gui.image2target** This module provides a graphical interface to locate GCPs on the image and provide their coordinates. It is possible to either manually enter the coordinates or locate the same point on a (geo-referenced) map. The height is provided by the DTM.

**l.ortho.rectify** This module performs the actual rectification of the image.

### 3. TESTS

A series of test have been carried out to assess the influence of the features of the input data and of the GCPs choice on the geometric proprieties of the output orthophoto.

The main aim of these tests is to determine whether increasing the resolution of the image and/or of the DTM provides a significant advantage in terms of geometric accuracy of the features on the ortho-rectified image. Given the large number of images, the possibility of using a reduced resolution reduces considerably the size of the dataset. The feasibility of the use of a medium resolution DTM makes possible to perform the orthorectification procedure even where a high resolution DTM is not available.

An example of the relationship between digital image resolution, file size and ground resolution is given in Table 1:



Scanning resolution [DPI]	600	1200
Mean space memory [MB]	40	120
Ground resolution [m]	2.1	1.4

Table 1. File size and ground resolution for 600 DPI and 1200 DPI resolutions 23x23cm images.

The analysis of the precision of the external image orientation is based on the root mean square error (RMSE) for each GCP (equation 4).

$$RMSE_i = \sqrt{\mu_i^2 + \nu_i^2} \quad (4)$$

where  $RMSE_i$  = Root Mean Square Error of the i-th GCP;  
 $\mu_i$  = residual of the i-th GCP in the east-west direction;  
 $\nu_i$  = residual of the i-th GCP in the north-south direction.

These RMSEs depend on the accuracy of the GCPs selection and on their correspondence the model used in the Least Square Adjustment for the parameters estimation (Rocchini et al., 2012).

An "overall RMSE" is defined as follows:

$$Overall\ RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N RMSE_i} \quad (5)$$

where  $RMSE_i$  = Root Mean Square Error of the i-th GCP;  
 $N$  = Total number of GCPs.

Tests on the influence of different image and DTM resolutions on the results have been carried out for image 4954, which is representative for the datasets in terms of coverage and height variation.

Using a 1200 DPI scan obviously improves the mean ground resolution of the non rectified image and the overall precision in the rectified image at the cost of larger disk space occupation (table 1), higher memory demand during the orthorectification process and longer processing times. Table 2 shows that with the same number and positions of GCPs the overall RMS is basically constant in terms of pixels, therefore it decreases with the same ratio as the resolution in meters.

image 4954	RMS [m]	RMS [pixels]	# GCPs
600 dpi, DTM 10 m	15	6.8	11
1200 dpi, DTM 10 m	8	7.3	11

Table 2. Comparison of overall RMS for image 4954 with different scan and DTM resolutions.

The DTM resolution effect on the overall RMS has been tested by performing the orthorectification procedure on the same image with the same resolution and the same GCPs, varying only the DTM resolution. Table 3 shows the resulting overall RMSEs using different DTMs on image 4954 with a 600 DPI resolution.

DTM resolution [m]	RMS [m]	RMS [pixels]	# GCPs
25	18.55	8.43	11
10	15	6.8	11
5	15	6.8	11
1	13.6	6.2	11

Table 3. Variation of the precision in photograph 4954 for different choices of DTM and scanning of 600 DPI.

Differences between using a 1m resolution DTM and a 10m resolution DTM hardly exceeds 1 pixel, so the optimal choice in terms of memory allocation and processing time for this set of photographs (at a resolution of 600 DPI) is 10m DTM.

These results are only a small example of the overall RMS for the large set of rectified images, but it is representative the vast majority of the results obtained.

#### 4. RESULTS

A grand total of 90 photographs from volo GAI have been rectified so far, using a 600 DPI resolution and a DTM of 10m resolution, following the indications emerged from the tests described in section 3.

The nearest neighbour method was chosen as interpolation method, because it is less computationally intense compared to the other two available methods (Novak, 1992). The processing has been carried out using the WGS84 reference system with UTM projection, zone 32 N.

Figure 3 shows the tile index for the orthorectified images. They cover the vast majority of Trentino territory, while images for some other areas, not covered by the tiles in figure 3, namely Monte Baldo, Paneveggio, Tesino and Val di Fassa have been already rectified in previous works (Rocchini et al., 2012) (Tattoni et al., 2010) (Cattani, 2015) (Maireri, 2018).

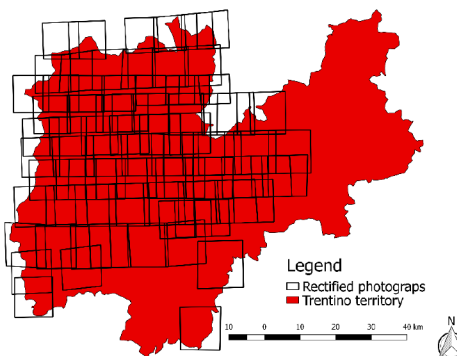


Figure 3. Tile coverage of the orthorectified photographs.

Figure 4 shows one rectified photograph (3853): the displacement effect of the altitude variations is clearly visible in the south (bottom) and north east (top right) part of the image. The image is also rotated counter-clockwise due to the difference of direction between the flight path and the datum axes.

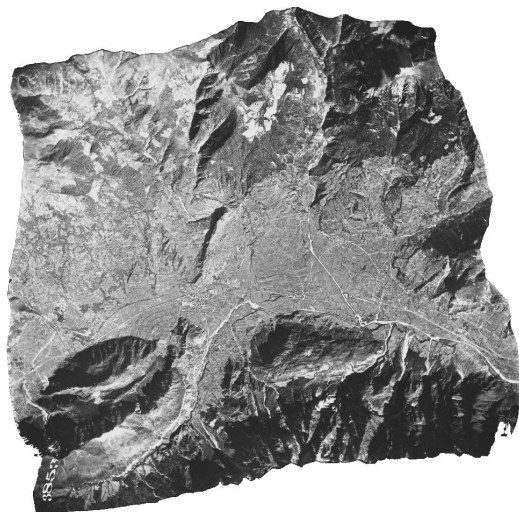


Figure 4. The result of orthorectification on photograph 3853.

The quality of the original images varies frame to frame, while some images easily allow the recognition of features such as trees or isolated structure (figure 5), in many images urban areas appear as a white blur, with poor detail (figure 6).



Figure 5. A pasture near Campitello (Val di Fassa): buildings, trees and open areas are clearly recognizable.

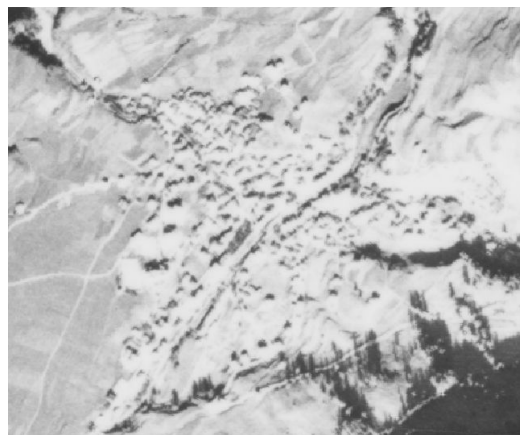


Figure 6. Depiction of the village of Campitello (Val di Fassa). It is not clear how buildings and streets are arranged.

Furthermore, another issue arises in areas with high slopes. There the DTM acts like a discontinuous function and the algorithm is not able to fit precisely the photograph to the DTM. This causes a stretch and a step pattern in the rectified image. Figure 7 shows a comparison between the step pattern obtained with different resolutions DTM: using a finer resolution causes a decrease in size of the pattern.

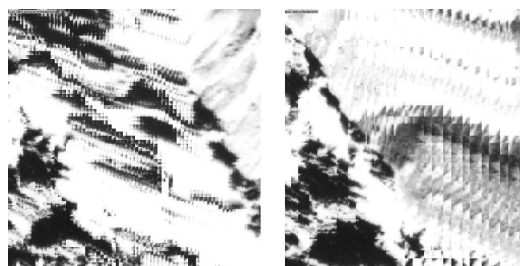


Figure 7. Effects of high slopes in rectified photographs. For the same area on the left a 1 m DTM was used, on the right a 10 m DTM was used.

## 5. DISCUSSION AND CONCLUSION

### 5.1 Precision assessment

Once image and DTM resolution is set, the only variable influencing the precision of orthorectification are number and positions of GCPs.

Their positioning strongly depends upon the availability of recognizable features on the territory depicted in photographs. Ideally, GCPs should be distributed homogeneously through the photograph to assure homogeneous results with good results for every part of the image. On this basis, Bernstein proposed that a sufficient number of GCPs should be 16 (Bernstein and Colwell, 1983). However, it is not always possible to place 16 GCPs homogeneously. For example, some photographs included in Volo

GAI depict the glacier zone of "Parco dell'Adamello-Brenta", which were completely occupied by glaciers in 1954 and have changed dramatically. Furthermore, the mean altitude of these areas is above 2000 m a.s.l. where there are few buildings, so features that can be used as GCPs are very rare and not homogeneously distributed.

In such cases orthorectification is still possible: GRASS GIS algorithm can orthorectify an image with 4 non-aligned points. The result needs to be carefully inspected: with *g.gui.mapswipe* tool in GRASS GIS it is possible to compare features (such as roads, lake shores and mountain ridges) between the processed photograph and a reference rectified image or map to check of differences. Although it is possible that they have changed shape through the years when using maps with different dates, an error in the rectification process is often recognizable, as figure 8 shows.

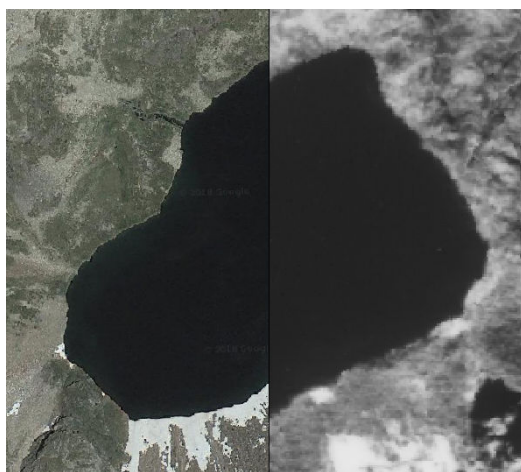


Figure 8. A comparison between the same lake in a 1954 rectified photograph with few GCPs (right) with an already rectified 2006 image (left). Note the displacement of the edges of this particular shape.

Volo GAI orthorectified photographs could be used in different studies such as landscape metrics analysis, glacier evolution studies and river morphological changes in Trentino.

In small areas of Trentino, namely Val di Fassa and Paneveggio authors Tattoni et al., and Maimeri have already conducted a landscape metric analysis with the use of 1954 photographs ((Tattoni et al., 2010), (Maimeri, 2018)). In both cases a land use classification was carried out to extract the forested and pasture areas; these were then resampled with a pixel resolution of 10m. In those studies the required precision was in the same order of magnitude as the RMSE of orthophotos. For other kinds of studies this precision could not be enough: before using 1954 Volo GAI rectified imagery it is important to manually check both precision and quality to avoid bias in the results.

## 5.2 Discrepancies between GRASS 7.4 and 6.4

GRASS 6.4 and 7.4 both implement apparently the same version of the orthorectification algorithm. However, using the same in-

puts (image, GCPs, fiducial markers and DEM) on the same computer can lead to two different results. Figure 9 is obtained using the same photograph (3862) and same inputs in GRASS 7.4 and 6.4. Using *r.mapcalc* the pixel by pixel grey level difference have been evaluated for the two images. The absolute differences between the two output orthophoto is shown in figure 9, in a grey scale representative with lower values in black and higher values in white.

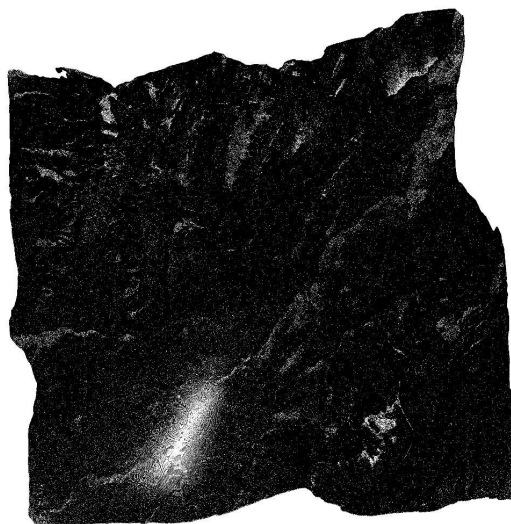


Figure 9. Discrepancies between the same photograph (3862) rectified in both 7.4 and 6.4 version of GRASS. In white are represented pixels with different grey values, in black pixels with the same grey value.

A correlation of the higher values with the slope values can be observed, but it is not clear why these discrepancies occur: further analysis are required on versions 7.4 and 6.4 of GRASS algorithm implemented in the *i.ortho.photo* module.

## 5.3 Conclusions

As stated in (Rocchini et al., 2012) differential rectification in a region characterized by a very diverse and ragged territory has proven its potential in georeferencing a large set of photographs, obtaining precision on RMSE around 10m for each photograph. In mountainous areas using just a polynomial rectification could lead to read tilt and displacement errors which will affect analysis based on remote sensing and image classification (Aronoff, 2005) (Morgan et al., 2010).

Results of section 3 show that an increase in image and DTM resolutions does not provide significant advantages for the orthorectification process with the 1954 Volo GAI dataset. Other limitations due to the original image quality play a more important role in the information that can be obtained from the orthorectified images, as seen in section 4.

GRASS GIS has proven to be a proper solution for rectification:

1. it is a free and open software, for many researchers it allows collaboration and research reproducibility;

2. it represents a way to save costs and still work with a powerful GIS tool capable of orthorectifying;
3. its code could be edited to give better performances in orthorectification;
4. it is user-friendly and troubleshooting is made easier by an active community of developers.

Future studies will investigate the differences between GRASS 6.4 and 7.4 results examining the source code to understand the different results. Further analysis will be conducted on the effects of the different available interpolating methods (nearest neighbor, bilinear and bicubic interpolation) during the orthorectification process.

However, GRASS GIS is not the only open source solution, as other FOSS systems, such as OSSIM (OSSIM Development Team, 2015), can also perform orthorectification.

## REFERENCES

- Aronoff, S., 2005. Remote sensing for GIS managers.
- Bernstein, R. and Colwell, R., 1983. *Manual of remote Sensing*. American Society of Orthorectification, New York.
- Cattani, G., 2015. *Metodologia per la valutazione della perdita di biodiversità dovuta al cambiamento del paesaggio agroforestale: il caso del tesino*. University of Trento, master's degree thesis, Trento, Italy.
- Ciolfi, M., Federici, B., Ferrando, Marzocchi, R., I., Sguerso, D., Tattoni, C., Vitti, A. and Zatelli, P., 2017. Foss tools and applications for education in geospatial sciences. *ISPRS international journal of Geo-Information*. art. no 225.
- Ciolfi, M., Tattoni, C. and Ferretti, F., 2012. Understanding forest changes to support planning: A fine-scale markov chain approach. *Developments in Environmental Modelling* 25, pp. 355–373. DOI: 10.1016/B978-0-444-59396-2.00021-3.
- Cohen, W., Kushla, J., Ripple, W. and S.L., G., 1996. An introduction to digital methods in remote sensing of forested ecosystems: focus on the pacific northwest, usa. *Environmental Management* 20, pp. 421–435.
- Geoportale Regione Lombardia, 2018. Metadata: Volo GAI 1954. Geoportale regione Lombardia, <http://www.geoportale.regione.lombardia.it/metadati> (2018).
- GRASS Development Team, 2015. Geographic Resources Analysis Support System (GRASS) Software, Version 7.4. Open Source Geospatial Foundation <http://grass.osgeo.org> (20 april 2018).
- IGMI, 2016. Product description: black and white/colours aerial imagery. Istituto Geografico Militare Italiano <https://www.igmi.org> (2016).
- Jensen, J., 1996. *Remote sensing of the environment: an Earth resource perspective*. Prentice-Hall, New York.
- Maimeri, G., 2018. *Variazione della copertura forestale in val di Fassa tramite analisi GIS multitemporale*. University of Trento, master's degree thesis, Trento, Italy.
- Morgan, J. L., Gergel, S. E. and Coops, N. C., 2010. Aerial Photography: A Rapidly Evolving Tool for Ecological Management. *BioScience* 60(1), pp. 47–59.
- Neteler, M. and Mitasova, H., 2008. *Open Source GIS: A GRASS GIS Approach*. Springer, New York.
- Novak, K., 1992. Rectification of Digital Imagery. *Photogrammetric Engineering and Remote Sensing* 58(3), pp. 339–344.
- OSSIM Development Team, 2015. Open Source Software Image Map (OSSIM, Version 1.8. Open Source Software Image Map <https://trac.osgeo.org/ossim> (04 april 2018).
- Paine, D. and Kiser, J., 2003. *Aerial photography and image interpretation*. 2 edn, Wiley, New York.
- Rocchini, D., Metz, M., Frigeri, A., Delucchi, L., Marcantonio, M. and Netler, M., 2012. Robust rectification of aerial photographs in an open source environment. *Computers and geosciences* 39, pp. 145–151.
- Tattoni, C., Ciolfi, M., Ferretti, F. and Cantiani, M. G., 2010. Monitoring spatial and temporal pattern of paneveggio forest (northern italy) from 1859 to 2006. *IFOREST* (3), pp. 72–80.
- Tattoni, C., Ianni, E., Geneletti, D., Zatelli, P. and Ciolfi, M., 2017. Landscape changes, traditional ecological knowledge and future scenarios in the Alps: A holistic ecological approach. *Science of the Total Environment* (579), pp. 27–36.

## 2.5 A PROCEDURE FOR IMAGES CLASSIFICATION

The aim of a classification procedure is to extract and represent the land cover in a raster map in a finite number of categories. Land cover is an attribute of a portion of landscape, that describe the use of that landscape. It can be a natural land cover (e.g. forest, meadows, water bodies, ecc..) or artificial (e.g.: houses, roads, parking lot, ecc...). In the current study, to reconstruct the past landscape in PAT it was necessary to perform a land use land cover (LULC from now on) analysis, using GRASS GIS and its dedicated modules. LULC can be achieved with two different classes of algorithms:

- pixel-wise, a family of algorithms that classifies every single pixel of the image in an LULC category, using its spectral signature [68];
- object oriented, an algorithm that creates "objects" (also called "segments") which are clusters of adjacent pixels with a similar spectral response, usually representing an entity, and then classifies them using their geometric and radiometric features [10].

Both these approaches exploit different algorithms and informations coming from the analysed images, but a starting point for each supervised algorithm are the training areas, or areas defined by the user where the land cover is known and defined. From these areas the information about the spectral signature of a land cover class are calculated and then used to classify all the land cover present in an analysed images [50]. A spectral signature is a set of statistical parameters (mean value, variance matrix and covariance) used to describe the value of a group of pixels (in our case the training areas), it can be the values on a single-band greyscale image, or it can be the different values of the single color band of a multiband image.

Both these approaches have proven extensively their viability in the literature ([2, 10, 15, 34]) so the choice of one over the other should be lead by the ratio between the image resolution and the objects depicted in the image (e.g.: houses, trees, roads) [36]. If the objects are composed by a set of pixels, in other words they are bigger than the pixel dimension, an object-oriented approach may present major advantages, such as a reduction of the error introduced in the boundaries areas between forest/open areas and the errors introduced by the semi-shadows (areas of non covering shadows, e.g.: all those darkened areas, but where the land use is still recognizable) in the pixels spectral response [36]. If the objects are depicted by one pixel, in other words the pixel dimension is bigger than the objects depicted in the image, a pixel-wise classification is preferable, because the segmentation could mix together pixels defining different land uses and therefore leading to a misclassification [36]. In this case study the dimension of the pixels was always in the order of magnitude of meters or less (see table 2) and an object-oriented approach was preferred.

### 2.5.1 Image segmentation

The first step in this approach is to define the segments, or single objects in an image [50, 115]. Objects represent entities in an image, e.g.: a group of trees, a crop field, a house, a road ecc... This process is called segmentation and it was performed using the module *i.segment* in GRASS GIS [80]. This module follows a segmentation algorithm which chooses random pixels inside the images, checks the neighbours pixels for the spectral response, if the difference in this response is less than a chosen threshold the pixels are considered parts of the same objects, if not *i.segment* generates a new object [94, 15]. The parameters that the user can set in input are the threshold value and the minimum area in pixel for each segment. If segments are less than the minimum area they are automatically merged with the bigger, more radiometric similar neighbouring segment. The threshold parameter is a normalized value ranging between 0 and 1: by lowering this value the algorithm becomes more sensitive to radiometric changes, until the algorithm considers each pixel as a single object. A special care must be taken in this step of the process: if the segments are too big they do not correspond the objects and an error about the shape of the objects is introduced in the classification process [58]. If the segment are too small their geometrical proprieties could not properly represent the geometry of objects belonging to the land cover class and the classification algorithm could not recognize the proper land class [57]. Assessment of the segmentation is usually visual by the user, but this approach is difficult and easily biased [58]. For this exact purpose a module named *i.segment.USPO* [65] was developed in GRASS GIS, which performs a series of segmentations on a small test area of the images varying the values of threshold and minimum size and evaluates the best combination of these parameters by calculating the intra-object variance weighted by object size, which measures the radiometric homogeneity of each object, and the spatial autocorrelation which measures the difference between neighbour objects in terms of spectral response [58].

For the current case study the module *i.segment.uspo* was applied on small regions inside an image for each set of images. The criteria used to chose these test areas was the complexity of the landscape they depict: situations where houses, roads and sparse trees are represented are a good benchmark for the segmentation algorithm. Once the module *i.segment.uspo* yields the best parameters a run of the *i.segment* module was performed on a whole image of the dataset to check the performance of the parameters. The segmentation was then checked visually to assess the accuracy of the segmentation. The table 4 reports the threshold values and the minimum size in pixels for each set of images. It is worth reminding that the images of 2015 had to be resampled to a 0.5x0.5m pixel resolution, because their resolution was too fine and the module *i.segment* generated too many segments for a single image and the processors of the calculator that run the module was not able to handle the output image.

Imagery set	Threshold	Minimum area size
1954	0.04	100
1994	0.06	125
2006	0.04	100
2015	0.04	100

Table 4: Values of threshold and minimum area size for each set of aerial imagery.

### 2.5.2 Classification of objects

The second step is the classification of the segments. This was performed by the application of a series of GRASS GIS modules on the input images to calculate some parameters on each object in the image, by then comparing this parameters with the ones of objects of known land use set by the user, also called training areas. The module *i.segment.stats* [79] transform the raster map of the segments into a vector map, where the associated table is made by a number of columns containing information about the:

- min;
- max;
- range;
- mean;
- mean of absolute values;
- standard deviation;
- variance;
- coefficient of variation;
- sum;
- sum of absolute values;
- first quartile;
- median;
- third quartile;
- 90th percentile

of each colour band of the image in each object.

The user can choose the set of parameter to include in the analysis. In the current study were considered the following parameters:

- mean;
- standard deviation;
- variance;
- first quartile;
- median;
- third quartile;
- 90th percentile.

This is considered to be the spectral signature of each object, to be compared with the spectral signature of the training areas. The module *i.segment.stats* calls upon the module *r.objects.geometry* [78] to calculate the geometric parameters of each object, which are additional information used by the OBIA procedure respect to the pixelwise procedure:

- area of the object;
- perimeter of the object;
- fractal dimension (FD) described as:  $FD = 2 \cdot \frac{\log(\text{perimeter})}{\log(\text{area} + 0.001)}$ ;
- compact square (CS) described as  $\text{compact\_sq} = 4 \cdot \frac{\sqrt{\text{area}}}{\text{perimeter}}$ ;
- compact circle (CC) described as  $\text{compact\_circ} = \frac{\text{perimeter}}{2 \cdot \sqrt{\pi \cdot \text{area}}}$ .

Even in this case the user can choose which parameters to use. In the current case study the chosen parameters were:

- fractal dimension;
- compact square;
- compact circle.

The module *i.segment.stats* evaluates the coordinate of the centroid of each object, but this data was not considered relevant and therefore excluded from the analysis. The classification scheme required the identification of a large number of training areas for each set of aerial imagery and for each land cover class:

Id	Land cover class 1954	Land cover class 1994	Land cover class 2006/2015
1	Forest	Forest	Forest
2	Forest in shadows	Forest in shadows	Forest in shadows
3	Pastures and meadows	Pastures and meadows	Pastures and meadows (and crops for 2015)
4	Crops (included orchards)	Crops	Crops (only for 2006)
5	Rock and barren land	Rock and barren land	Rock and barren land
6	Roads	Orchards	Orchards
7	Roofs (Red and Black roofs)	Roads	Roads
8	Shaded areas	Roofs (Red and Black roofs)	Red roofs
9	-	Shaded areas	Black Roofs
10	-	-	Shaded areas

Table 5: Classification scheme used for the LULC analysis. The id column identifies the category number assigned to each LULC class.

It is important to underline the central role played by the training areas, since they must well-represent the land class they refer to. Since land cover categories may have a high variability in spectral response, due to variable condition of the incident light, or time of the day, or tilting of the camera taking the aerial images, the training areas must represent all this variable conditions [50]. The users should choose a sufficient large amount of training areas for each land cover category [50]. In this process a particular care should be taken: if training



areas of different land class share a very similar spectral response the output will contain various misclassified areas. In that case it will be necessary, after assessing the result and realizing that the the accuracy of the classification is low, to re-define the training areas until the classification is good enough. The table 6 indicates the number of training areas used for each of the temporal set of the images used for the present study. A more detailed overview of the training areas number is reported in table 7.

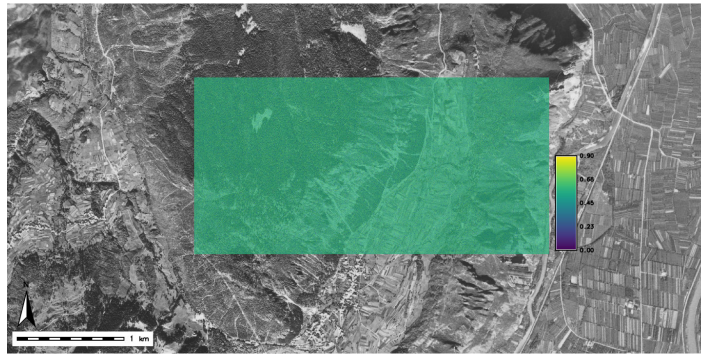
Year	Training areas
1954	2572
1994	1920
2006	1998
2015	1831

Table 6: Number of training areas for each year of aerial imagery.

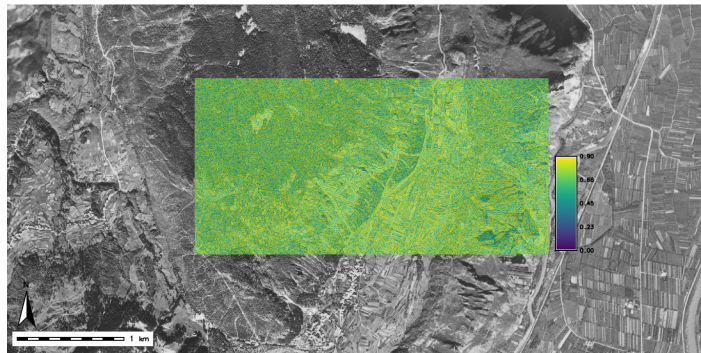
Set of training areas	Land use class									
	1	2	3	4	5	6	7	8	9	10
1954	667	145	289	355	343	182	150	441	-	-
1994	611	119	250	375	95	93	100	146	131	-
2006	375	274	209	199	205	249	135	167	130	55
2015	469	101	513	-	161	324	69	108	31	55

Table 7: Number of training areas divided for land use for each dataset

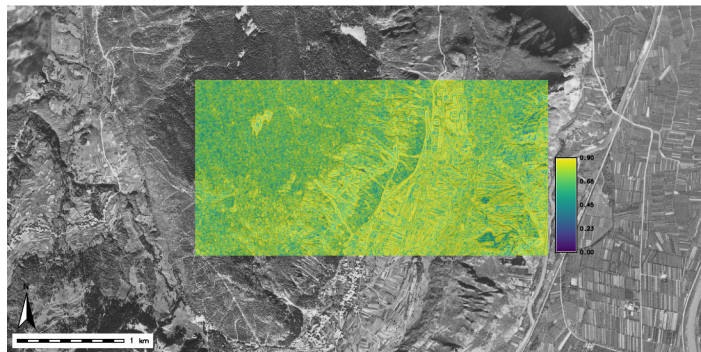
The texture of the image was used to provide additional informations to the classification scheme for the greyscale images. Texture is a feature of a specific land class, linked to their spatial features, because it depends on the presence of patterns on the images [51]. For instance texture can be used to discriminate between a natural forest and a man made cultivation to reveal the relationship between the pixel value of a single pixel and its neighbour: a man organized situation will have a range of values, that differs from a natural situation [13]. This added band was helpful in the case of single band image, where land classes have a small spectral difference. Texture was calculated for each greyscale image in the dataset used for the current case study using the module *r.texture* in GRASS GIS [4]. This module requires the user to define a window size where the relationship between the central pixel and the neighbours will be evaluated to define the texture. This is usually made by observation and by considering the scale of the phenomena that needs to be described. Texture depends on many factors: the coarseness of the input image, the natural level of contrast of the image itself and the object represented on the image.



(a) An excerpt from the texture band calculus: a 3 cell window was used.



(b) An excerpt from the texture band calculus: a 9 cell window was used.



(c) An excerpt from the texture band calculus: a 17 cell window was used.

Figure 16: An example of texture calculated with different window sizes on the 1954 orthorectified images using correlation.

As image 16 shows, a small window does not give a significant correlation of the reflectance level between a pixel and another inside the same object, giving a final effect of a raster mostly composed by a constant value. On the other hand using a window of an excessive size present two problems: pixels belonging to areas with different LULCs are mixed, thus lowering the ability to discriminate between them, and 2) computational time increases dramatically, as shown in table 8.

Window size	Computational time [s]
3	7
5	10
9	58
17	242

Table 8: Computational time of the texture on a same area changing the window size.

A moving window of size 9 pixels was deemed optimal to perform the calculation of the texture in the current work.

GRASS *r.texture* uses the common texture model based on the so-called grey level co-occurrence matrix as described by Haralick, [51]. This matrix is a two-dimensional histogram of grey levels for a pair of pixels which are separated by a fixed spatial relationship. The matrix approximates the joint probability distribution of a pair of pixels. Several texture measures are directly computed from the grey level co-occurrence matrix. GRASS *r.texture* module provides several texture evaluation methods: Sum Average, Entropy, Difference Entropy, Sum Entropy, Variance, Difference Variance, Sum Variance, Angular Second Moment, Inverse Difference Moment, Contrast, Correlation, Measures of Correlation and Maximal Correlation Coefficient [51]. Correlation has been used for all the images because it has proved the best to differentiate between LULC classes in this dataset.

### 2.5.3 Distribution of radiometric and geometric parameters for the training areas

The graphs in figures 17 - 28 show the distribution of the mean and standard deviation values for each band for each LULC class, evaluated from training areas. Different values correspond to different land use training areas, following the scheme explained in table 5. The texture value was scaled to fit the graphs scale.

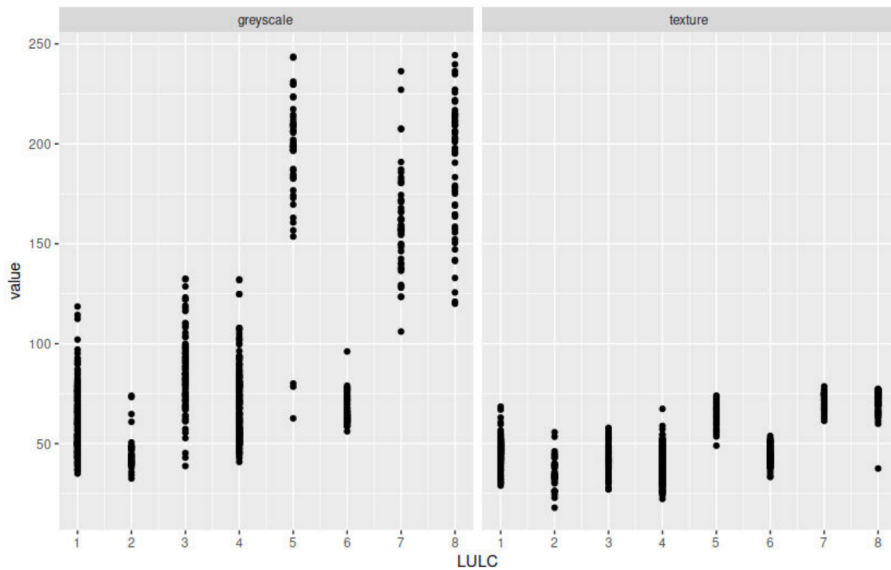


Figure 17: Mean values for the first set of 1954 training areas.

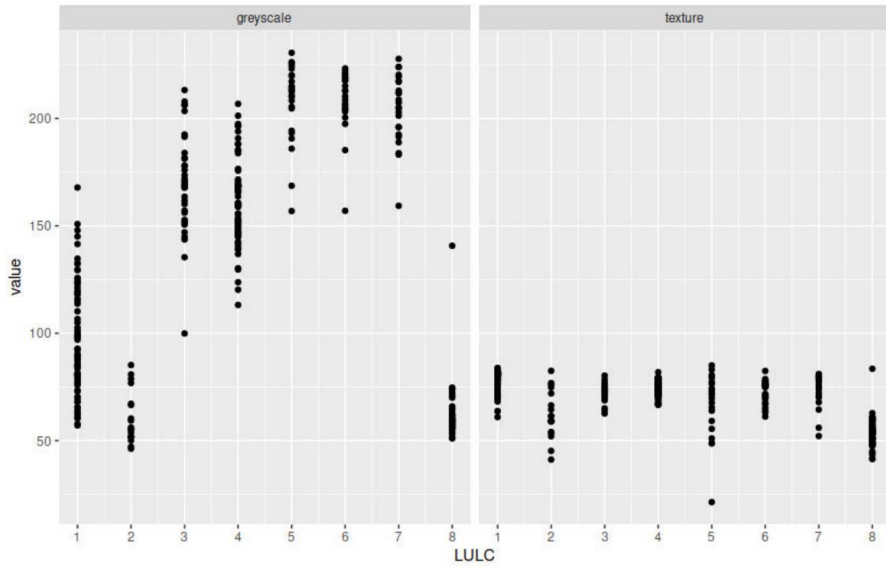


Figure 18: Mean values for the second set of 1954 training areas.

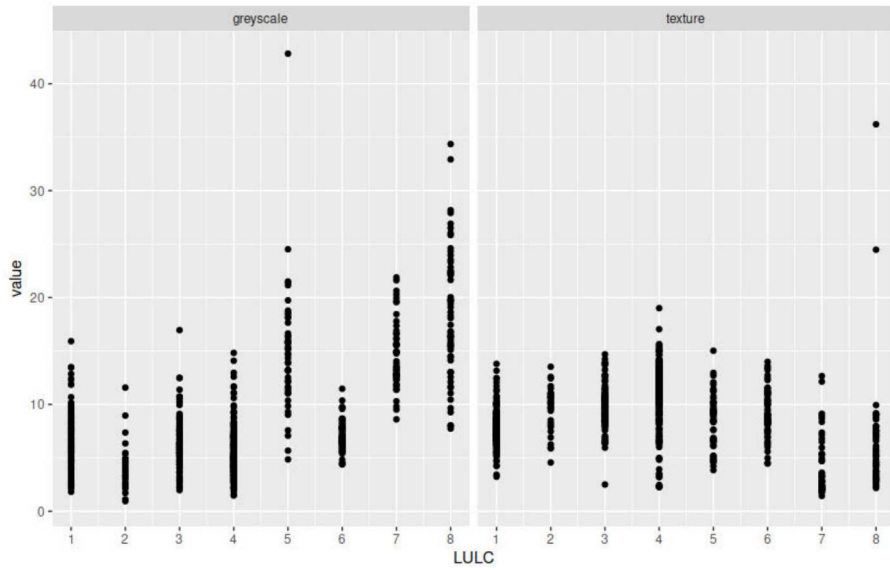


Figure 19: Standard deviation for the first set of 1954 training areas.

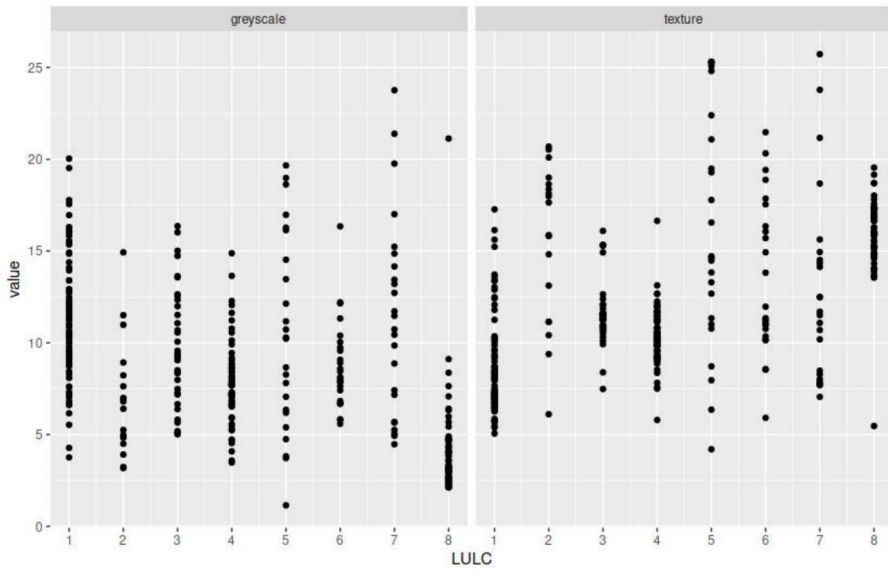


Figure 20: Standard deviation for the second set of 1954 training areas.

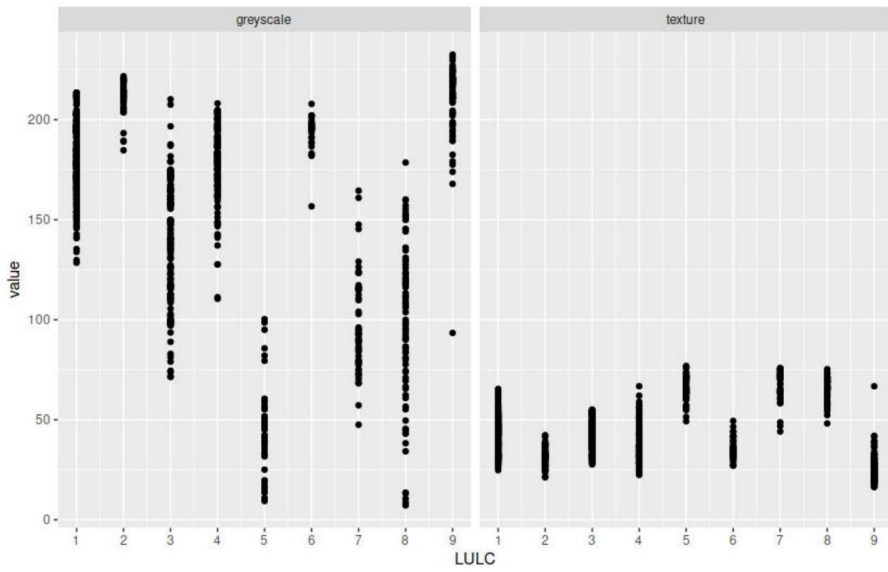


Figure 21: Mean values for the first set of 1994 training areas.

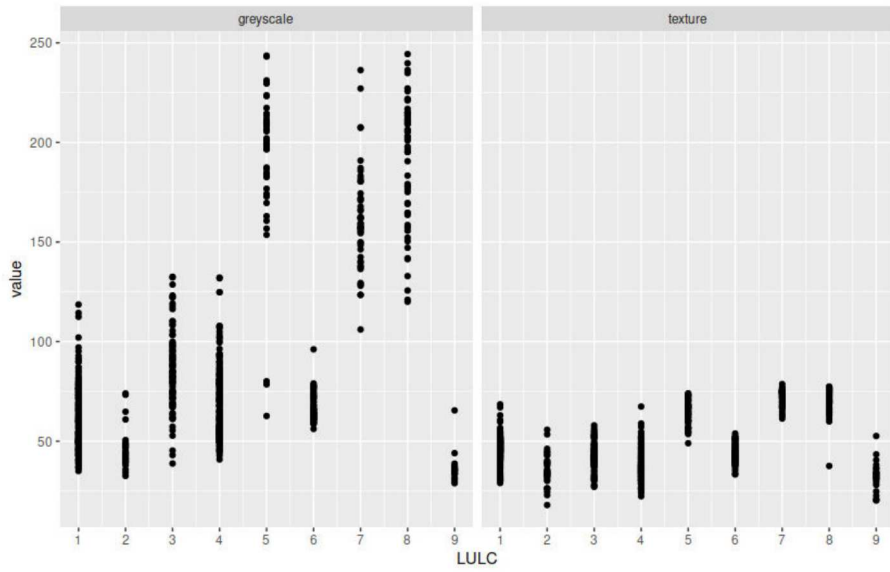


Figure 22: Mean value for the first set of 1994 training areas.

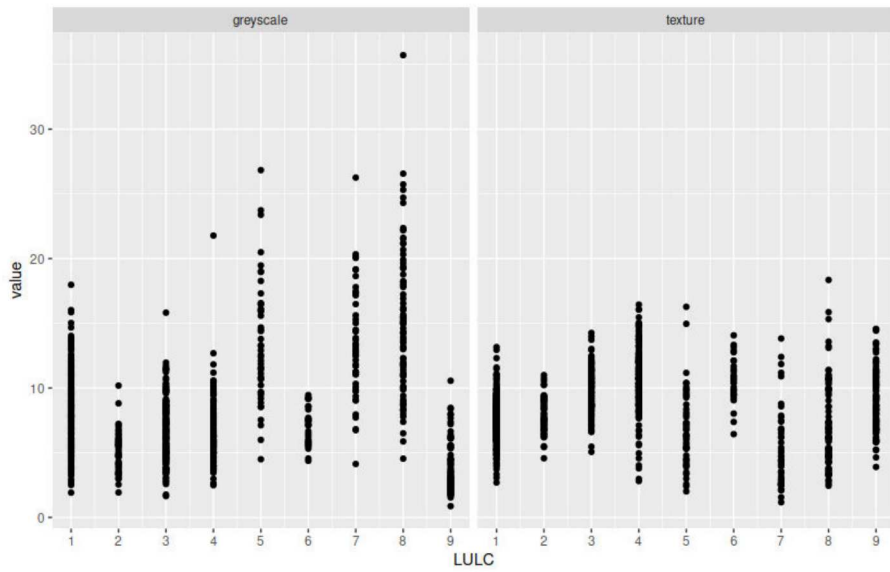


Figure 23: Standard deviation value for the first set of 1994 training areas

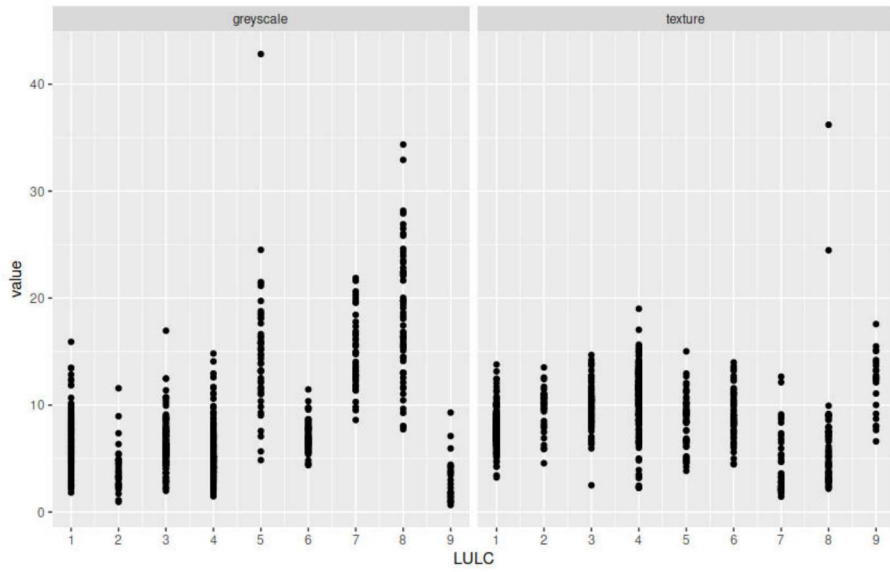


Figure 24: Standard deviation value for the first set of 1994 training areas

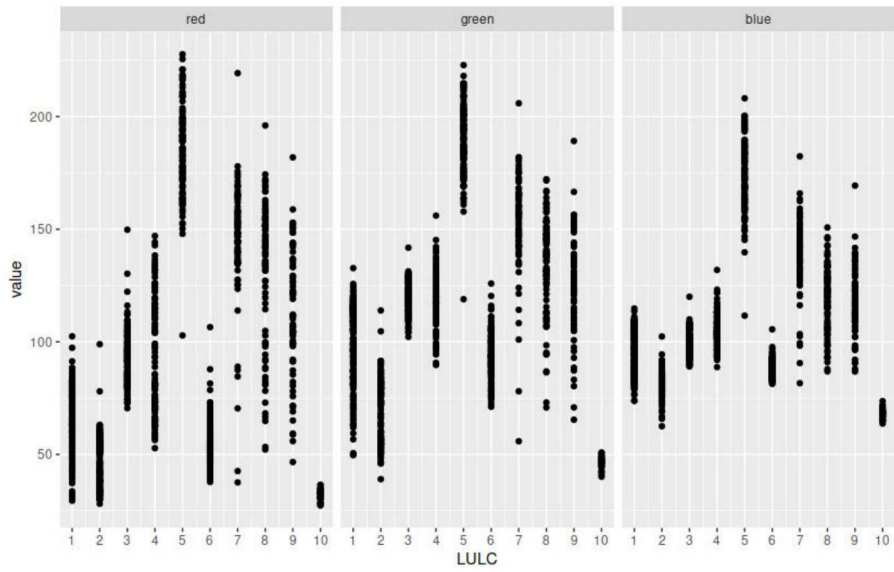


Figure 25: Mean values value for the 2006 training areas.

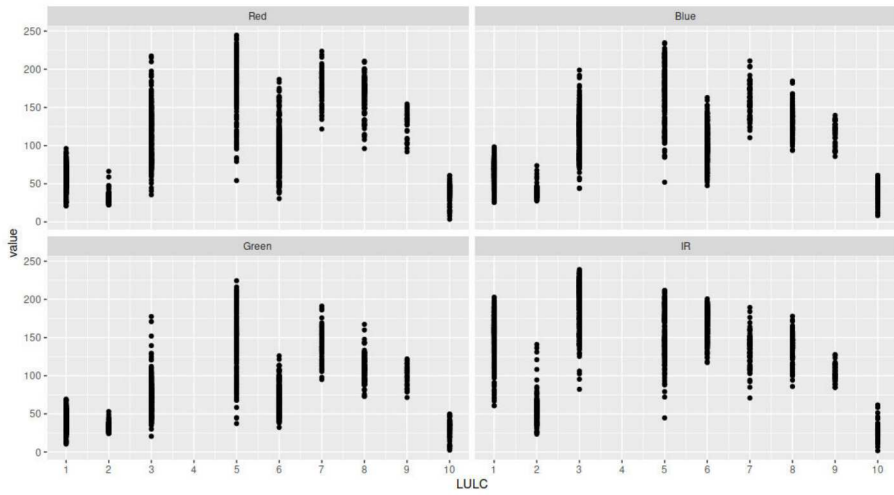


Figure 26: Mean value for the first 2015 training areas.

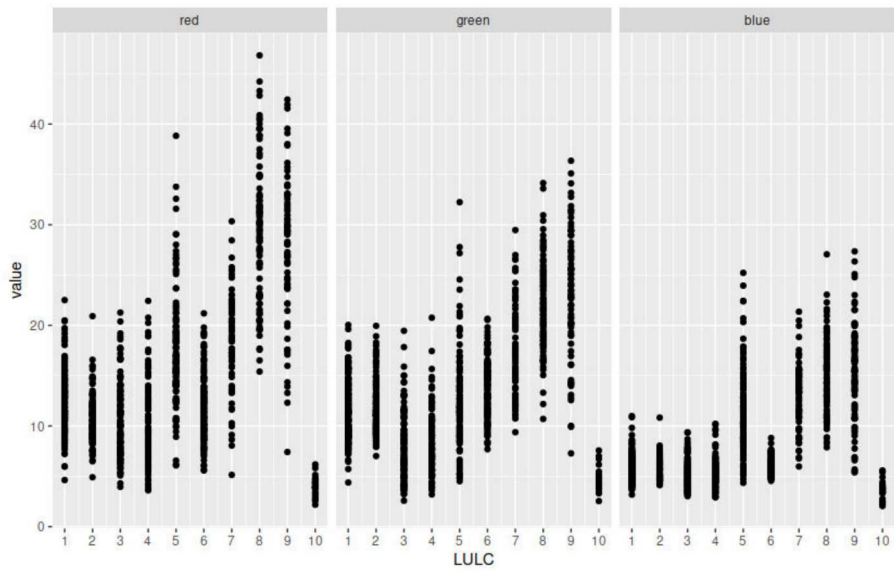


Figure 27: Standard deviation value for the 2006 training areas.



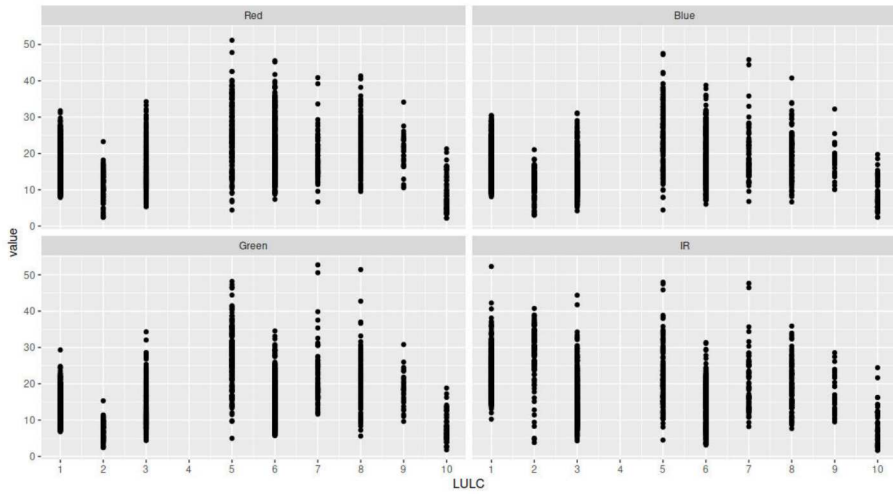


Figure 28: Standard deviation for the 2015 training areas.

The graphs in figures 29 - 32 show the distribution of the values of the geometrical parameters for the chosen training areas and for each set of imagery, grouped by land use category. The parameters used in this analysis were Compact Circle, Compact square and Fractal dimension, which are shown in these graphs.

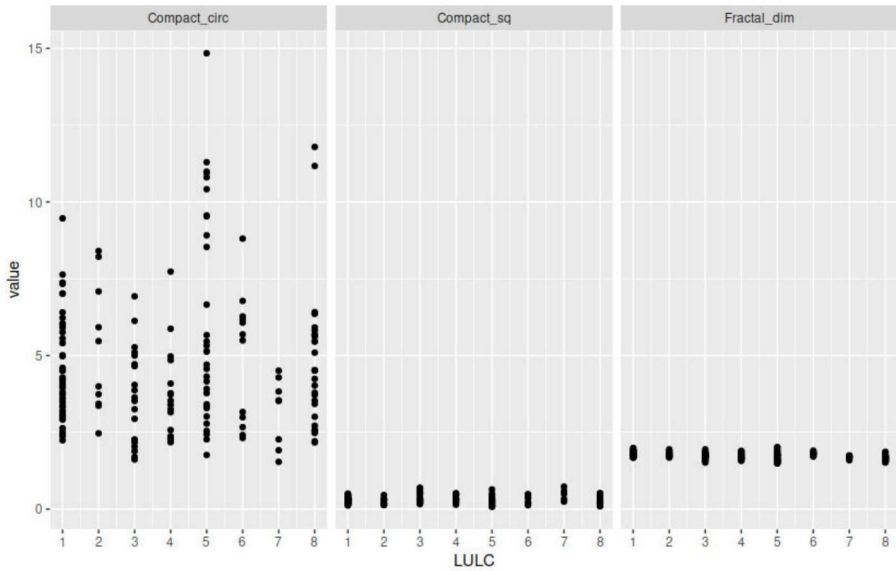


Figure 29: Geometric parameters for the 1954 training areas

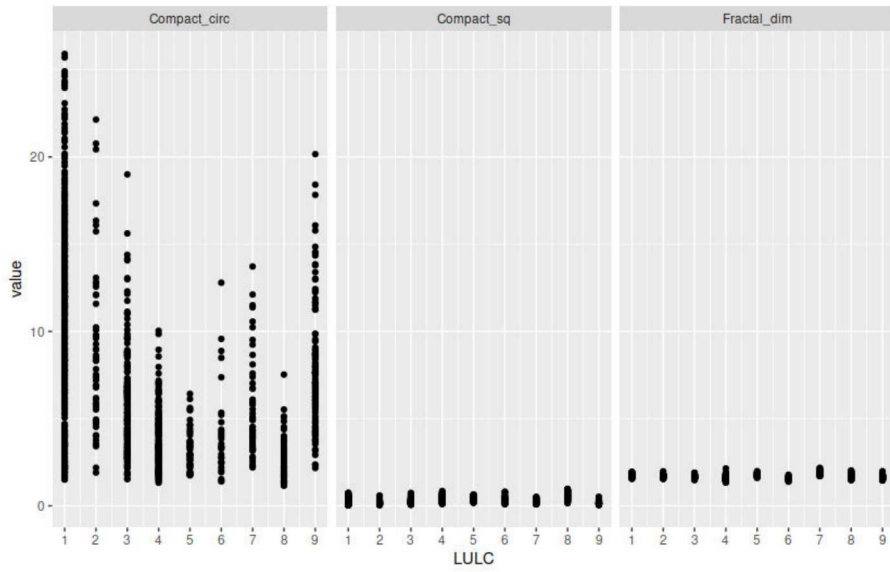


Figure 30: Geometric parameters for the 1994 training areas

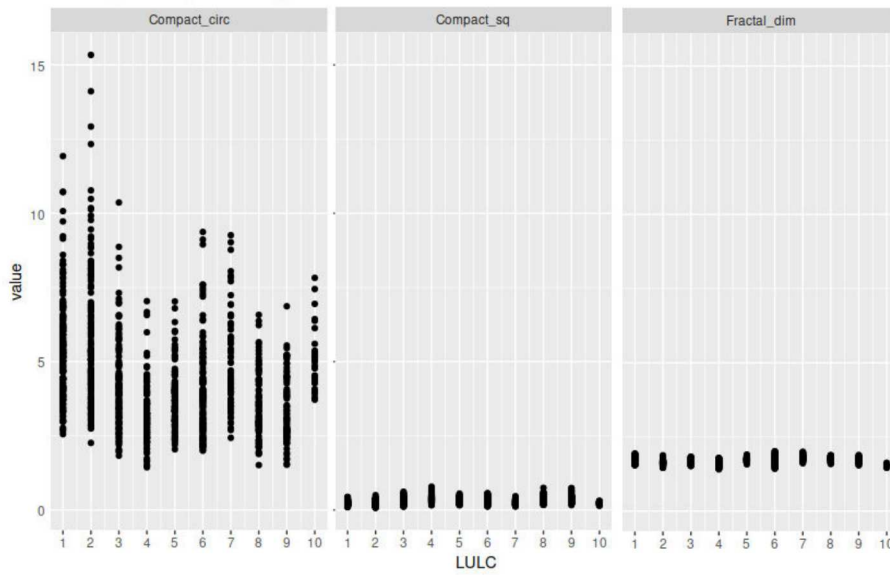


Figure 31: Geometric parameters for the 2006 training areas

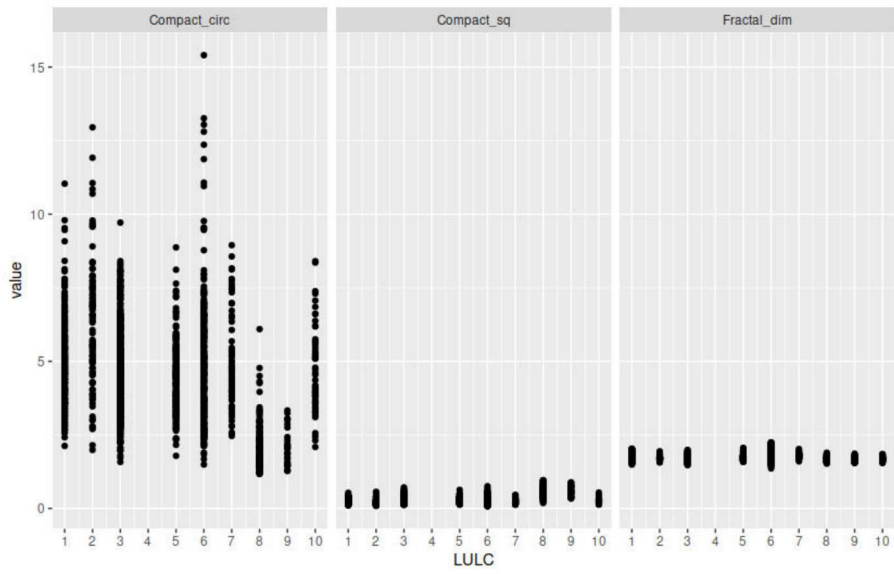


Figure 32: Geometric parameters for the 2015 training areas

#### 2.5.4 Classification of the imageries

The classification itself was performed using the GRASS module *v.class.mlR* [66], which combines GRASS and the R Caret library to perform machine learning classification [63]. This module allows the user to choose between a number of machine learning techniques that exploit the radiometric and geometric parameters set by the training areas to classify the objects in the vector maps of the segments in each image. The output of the classification consists in the LULC category assigned to each segment and an estimate of the probability of having achieved the correct classification. This information is recorded in two columns of table of the vector segment map.

The user can choose a single classifier rather than a number of different ones, in that case the module classifies the land cover of each object via a voting process, where "voters" are represented by the machine learning algorithms chosen by the user. The voting process combines the output of the single classifiers using different criteria. Two columns containing the assigned LULC class and its probability are added to the table for each voter.

Five different classification techniques have been used: k-nearest-neighbour classifier (k-NN), naive Bayesian classifier (NB), support vector machine (SVM), multi-layer perception (MLP) and decision tree (DT) [50]. Each classifier "votes" for the LULC class for each segment, then the module *v.class.mlR* calculates the LULC of each object by rules basing on the votes:

- Simple Majority Vote (SMV): all the classifiers have the same authority and the decision on the LULC class of an object is determined by a majority vote;
- Simple Weighted Vote (SWV): the weight of the vote of each classifier in the decision process is proportional to its accuracy in the identification of the class of the training areas;
- Best-Worst Weighted Vote (BWWV): as in the SWV process, all the classifiers are ranked according to their accuracy; the worst voter is excluded from the voting process;
- Quadratic Best-Worst Weighted Vote (QBWWV): as in BWWV, but the weights of the best voters are squared in order to give them more authority.

The voting systems were tested on a single greyscale image and on a single multiband image and it was identified a method which was best for all the datasets. For this particular scope, two images from the 2015 dataset, and the two corresponding images from the 1994 dataset were chosen, in the area of Monte Baldo, containing areas of forest and meadows.

On these images a number of training areas were extracted, and the different voters were ranked by the accuracy they reached. The accuracy was measured by the use of Cohen's K coefficient, with the module *r.kappa* in GRASS [121].

For the test images the training areas were selected only on one image and then the same training areas were used to classify the second image. This procedure simulated the situation of the classification, when all the orthophotos in a dataset have been classified using a single set of training areas. Tables 9 - 12 reports on the precision and accuracy assessment of the classification for the 1994 and 2015 images.

80151	kappa	accuracy
smv	0.96	98.78
swv	0.95	98.10
bwwv	0.95	98.10
qbwwv	0.95	98.10

Table 9: Evaluation of the classification accuracy for the image 80151 in the 2015 dataset, which contains the training areas.

80152	kappa	accuracy
smv	0.87	97.75
swv	0.89	97.76
bwwv	0.90	98.28
qbwwv	0.90	98.28

Table 10: Evaluation of the classification accuracy for the image 80152 in the 2015 dataset, which does not contains the training areas.

o80150	kappa	accuracy
smv	0.93	96.15
swv	0.93	95.62
bwwv	0.92	95.67
qbwwv	0.92	95.67

Table 11: Evaluation of the classification accuracy for the image 80151 in the 1994 dataset, which contains the training areas

o80160	kappa	accuracy
smv	0.86	96.35
swv	0.88	96.37
bwwv	0.90	97.4
qbwwv	0.90	97.4

Table 12: Evaluation of the classification accuracy for the image 80152 in the 1994 dataset, which does not contains the training areas

All the voting methods seem to grant an high grade of accuracy and precision (k is always greater than 0.9, with an accuracy of 97-99%) for the images which contains the training areas, while the bwwv and qbwwv voters yield a better kappa values and accuracy for the images which do not contains the training areas. For this reason the qbwwv voter which performs a quadratic weighting of the voter's authority was considered the best method for classifying the large temporal dataset of this research that might contain a great variation in features belonging to the same class.

### 2.5.5 The automatization of the procedure

Using the *grass\_scripts* library in the Python environment it was possible to semi-automatize the procedure for the classification of the imagery. The user intervention was minimized and limited to the input of the training areas and the set up of the parameters for the image segmentation.

Firstly the code made a list of all the rasters inside the current GRASS mapset, and, in the case of multi - band images it listed the different band of each single image, called "group". This list will be used as an array of input names to apply the GRASS modules for performing the LULC. Secondly, the code applied to each map or group the modules *i.segment*. This step was the longest in computational times, because the segmentation algorithm had to check each single cell of the image analysed and test the spectral difference with four neighbour cells. The module iterates the segmentation to cluster smaller segments into bigger ones, until all the segment in the image matches the condition set by the user of threshold value and minimum size. Segmentation was processed via a server machine with 32 GB of dedicated RAM, 8 intel i-7 processor running a Debian 10 OS, the time used for segmentation for each dataset are shown in table 13.

Dataset	Number of processed images	Comp. time for segmentation [d]
1954	91	10
1994	230	20
2006	230	20
2105	838	30
Potential treeline	98	2

Table 13: Computational time (in days) for each set of images used in the study

Thirdly, the code made a list of the segmented images and performs the vectorialization of the segmented images. The table of each vector maps contained a column for each parameter indicated in the paragraph 2.5.2, and a row for each segment. This step allowed the creation of training areas, which was the only part that required user intervention. To define these areas the images used as input need to have a normalized histogram for each single image set. In particular the dataset of the 1954 and 1994 presented specific situations of exposure highlighted by their histogram, therefore required the definition of different sets of training area to be properly classified.

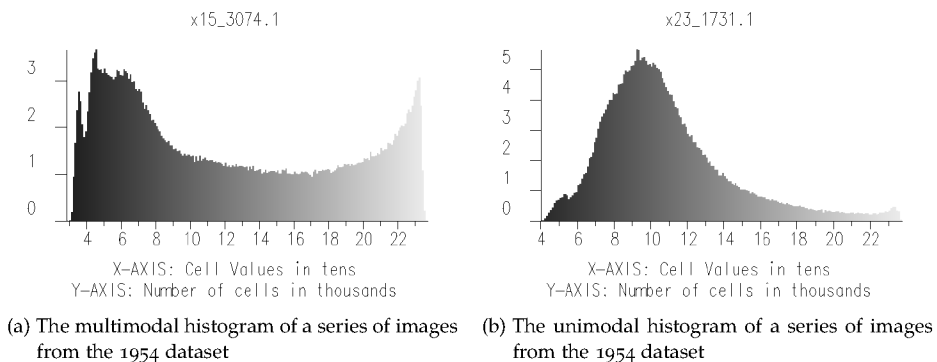


Figure 33: An example of different histogram from the 1954 dataset.

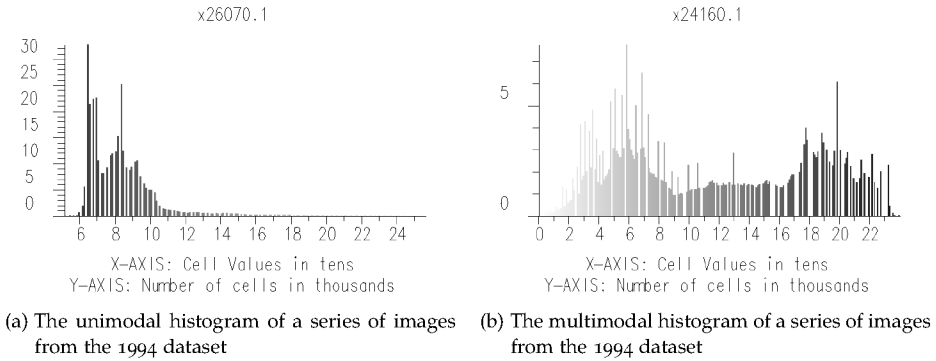


Figure 34: An example of different histogram from the 1994 dataset.

The results of the classification process were more a patch of different classification than the result of a single classification process.

To set up the training areas it is worth reminding that the segment carries an unique identification number, that was passed in a column of the table of the vectorized segmented map. To chose the training areas the user had to create a vector map made of point, and place a single point inside a single segment, where the user had to identify the LULC category according to table 5. This information was then stored in a column of the vector map. Once enough points for identifying training areas were placed for a set of imagery the second step performed by the Python code is to create a new column for the vector map of the training points, which is filled by the module *v.what.rast* [12] with the identification number of the segment that contains each points. Then, the module *v.db.join* [82] joins the tables of the maps of training points and the segmented areas, creating a new column, named LULC, that contains the values of the land categories, only for the segment identified by the same number as the training points. Finally the vector map of the training areas is created by the module *v.extract*, which select from the segmented vector map the areas which have a non-null value for the LULC column.

The final step of the Python code created for the classification process was the creation of a list of all the vector segmented maps and, using the training areas, the actual classification via the application of the *v.class.mlR* module to each map in the final list. The LULC value of each segment is stored in a column of the vector map of the segments, which was used to rasterize these maps, outputting a raster map of the classification result of each set of aerial imagery.

#### 2.5.6 The assessment of the results

The accuracy assessment of the results in such large region for such a great variety of data required an approach which could give standardized and comparable results. As the maps of the 1859 are concerned, the accuracy assessment was carried out while they were being vectorized by hand. All the parcels where a forest class was registered were vectorized with a 'forest' class, referring to the table 5. The maps of 1936 had a similar accuracy assessment, as they were processed by hand and all the parcels indicated as forests were registered with the 'forest' class in the table 5 [31].

As the aerial imagery are concerned, the variety of input data and the spatial scale which this study covers required a specific approach. A standard procedure for evaluating the quality of an automatic LULC process is to create a 'ground truth' map, which contains areas of known LULC using the same land cover classification scheme as the result of the LULC process [15].

These areas are compared with the LULC results to calculate the Cohen’s K coefficient and the overall accuracy, or the percentage of coincident land use class for the areas of the ground truth map respect to the results of the LULC. This passage can be performed with instrument such as the *r.kappa* module in GRASS GIS [121].

If the accuracy and the Cohen’K are sufficiently high for the purposes of the study the LULC procedure is considered satisfactory, otherwise the LULC has to be repeated changing the training areas or varying other parameters to achieve the required results.

This approach was discarded because at a regional scale and in a multitemporal analysis the construction of a ground truth map is not possible [90]. Maps such as the CORINE land cover could provide information for the ground truth, if they are from the same year as the classified dataset. CORINE lacks in resolution (100x100m pixels), which is coarser than the aerial imagery used in this study, making it difficult to interpret the results of *r.kappa*: the errors could depend on both the different resolution or a misclassification. Ground truth maps can be designed by hand but this approach was considered very little effective, because a human made ground truth map is prone to bias error, which lowers the robustness of the accuracy assessment procedure [74].

To overcome these problems the chosen approach was to design a simple script in Python, which was implemented in GRASS GIS to perform a semi-automatic accuracy analysis. The chosen approach is based on distributing random points across the PAT territory where to evaluate the LULC and the ground truth [74] and it is called *random stratified*. A ‘Strata’ is a portion of a population which members share an attribute, in our case the LULC. A number of points is randomly selected inside a single strata where evaluate the ground truth and compare it with the result of the classification. Stratified random sampling combines random sampling with some prior knowledge about the study area, creating random samples for each LULC class. The number of samples for each class is proportional to the ratio of its area and the total area of the region.

The procedure set up takes in input the maps of the LULC of a dataset, by using the module *r.stats* it calculates the percentage of the cover of a single land class respect to the total area of the landscape. A total of 750 points were distributed all across PAT territory divided proportionally to the area of a tested land use class respect the total area of PAT territory. The module *v.what.rast* [12] stores in a field of the table of the generated random points vector map the land class where the single point is placed, and by a human check point by point the ground truth is assigned. The classes for the ground truth are the same of the classification scheme for each dataset.

For the current analysis the tested land cover categories for the accuracy of the classification were the forest and shaded forest, the distribution of the points in the forest/non-forest stata is shown in table 14.

Year	Points in forest class	Points outside forest class
1954	337	413
1994	412	338
2006	427	323
2015	427	323

Table 14: Distribution of random stratified points



### 2.6.1 *Agent Based Models*

Agent Based Models (ABM) are used in complex theory to model system-level patterns that emerges from interactions between individual entities, also called Agents, and the environment [116]. The main difference between this model and system dynamics models is that interactions are governed by rules rather than equations [49]. In forestry these models are called 'forest gap' modelling [119]. The purpose of ABM is to explain the evolution of an environment through the behaviour of single individuals, so that each parameter that describes a population in an environment such as density, death to birth rate, size and so on are the output of the specific model rather than the input [122]. To be performant ABM thus require the understanding of the behaviour of a single individual. A basic concept is the 'emergence', identified by three criteria [49]:

- Emergent properties are not simply the sum of the properties of the individuals;
- Emergent properties are of a different type than the properties of the individuals (e.g. the spatial distribution of individuals is a system property of a type that none of the system's individuals has);
- Emergent properties are often counterintuitive and cannot easily be predicted by looking only at the individuals.

The fact that an emergence is not easily predictable should be perceived as an added value to the ABM procedure by unveiling how system-level properties emerge from individual behaviour [119].

An *Agent* is a discrete and autonomous entity that acts independently [119]. Even if all agents follow the same set of rules their actions will be influenced in reaction of their surroundings, such as environment or as other neighbourhood agents. Environment in ABM provides a spacial context, while '*rules*' explain the interaction between agents. These rules can be derived from empirical data or from theory and can be influenced by probabilistic or stochastic processes [49].

Finally ABM are regulated by discrete time: an action is undertaken by every agent once in each timestep which makes the system evolve in a different state. A timestep should well-represent the event we are trying to simulate.

### 2.6.2 *Implementation of ABM in the current Case Study*

The aim of the ABM procedure implemented for this case study was to make predictions on the possible evolution of the forest in PAT for the next century. The simulation for the possible future evolution in the current study used some simplification: the evolution of the forests depends on many factors, the health of the trees, the humidity, the type of trees presents and seed they produce, the presence and density of seed feeder and so on, and in this study they were considered negligible in the overall balance of the forest expansion [118]. Moreover, the effect of climate change on treeline dynamics were not considered.

To balance complexity and not making the error of oversimplify the model, this should be applied to forests or areas where there is a predominant specie, or predominant species that share similar reproduction dynamics. In the current work the procedure was applied to the Paneveggio Pale di San Martino natural park, located in the north-east of PAT. The forest in this area is composed mainly by spruce with related resinous species (see talbe 15), which made acceptable the simplifications adopted.

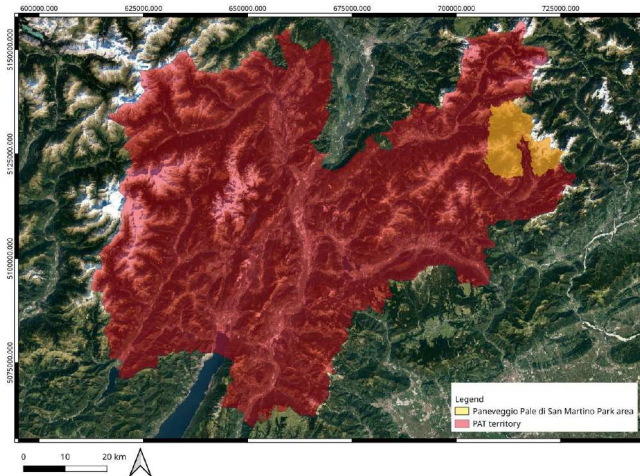


Figure 35: The position of Paneveggio Pale di san Martino natural park.

Forest class	Extension - ha
Hop-Hornbeam forest	0.214
Hornbeam and Oak forest	3.116
Maple Linden and Ash tree foresst	64.412
Pioneer formations	74.753
Pine forest	10.172
Beech forest	65.429
Fir forest	1548.641
Mountain pine forest	315.343
Alder tree forest	659.450
Spruce forest	8866.844
Larch forest	2073.853
Larch stone pine forest	1244.555
Stone pine forest	30.292

Table 15: Forest classes extensions in Paneveggio Pale di San Martino Area

The procedure required the definition of trees behaviour with statistical methods, which needed a calibration phase. This was achieved by studying how the distribution of the forest evolved in the past years, by comparing the 'real' situation with the 'simulated' situation. As real situation it was assumed the forest coverage extracted from the LULC output of the imagery dataset. The starting year was the 1954, because is the year of the oldest imagery set which gave information about the tree distribution. Different simulations were set up with varied input parameters, assuming as initial condition the 1954 forest coverage. The simulated situations of the years 1994, 2006 and 2015 (years of the other imagery set used in the current work) were used as check, to see which parameters best fitted the simulation of forest evolution in the study area.

The aim of the agents (trees) in the simulation is to distribute seeds, which will generate newborn trees that have to age until sexual maturity to start spreading other seeds. Envi-

ronment is defined by a grid, which contains information about the presence of trees in the cell themselves, altitude and the land use, to exclude all the areas where trees cannot grow (e.g.: rocky areas, rivers, roads ecc..). The seed distribution is controlled by a probabilistic distribution, function of the distance of the evaluated cell from the nearest sexually mature trees [118]:

$$\text{Prob}_{\text{seed}}(\text{dist}) = a \cdot e^{-0.5(\frac{\text{dist}}{c_1})^{d_1}} + b \cdot e^{-0.5(\frac{\text{dist}}{c_2})^{d_2}} \quad (1)$$

where:

- $\text{prob}_{\text{seed}}$  is the probability of the presence of a seed in a given cell;
- $\text{dist}$  is the distance between the current evaluated cell and the nearest cell full of trees;
- $a, b$  are normalization constant;
- $c_1, c_2$  are distance parameters;
- $d_1, d_2$  are curve shape parameters.

This distribution is the composition of two overlaying distributions. The first one takes into account the seeds dispersal due to the proximity of sexual mature trees to the evaluated cells, the second one takes into account seeds dispersal due to the winds. In fact,  $c_2 > c_1$ , because seeds transported by the wind reach farther distances, and  $a, b$  control the percentage of seeds which are spread by the wind with respect to those which fall nearby. The curve shape parameters control the shape of the distribution [118]. If the probability of distribution is greater than a fixed threshold value in a cell, then the cell is filled with a seed which will germinate. This assumption excludes other methods of seedling, such as distribution of the seeds by birds, insects or human intervention. The threshold value has to be calibrated: this is made by using as input the map of the forest extracted from the 1954 aerial imagery dataset, and confronting the output of the model in the years 1994, 2006 and 2015 to check the threshold value which yield the more matching situation with the real forest. Once a seed is planted in a cell a check is made to evaluate if the altitude allows the seed to survives and the newborn tree to grows [118].

$$\text{Prob}_{\text{germ}}(\text{dist}) = 1 - \frac{1}{1 + e^{-\frac{i}{c}(\text{elev} - \frac{i}{c})}} \quad (2)$$

where:

- $\text{prob}_{\text{germ}}$  is the probability of germination in a seeded cell;
- $\text{elev}$  is the elevation of a cell
- $i$  is the inflection point, or the altitude where trees stops growing;
- $c$  is a normalizing constant

The parameters of these equations where set following indication in the literature:

- $a=b=0,5$  [118];
- distance parameters  $c_1= 15$  m [107, 23];
- distance parameters  $c_2= 110$  m [107, 23];

- $d_1 = 2$  [118];
- $d_2 = 2$ [118];
- Inflection point,  $i = 2500m$  [107];
- Normalising constant,  $c = 155$  [107].

These probability functions are evaluated for each cell free from trees in the environment. If the threshold value is less than the probability of seedling the cell is seeded, and a new tree will be born if the threshold value is less than the probability of germination. The threshold values have to be calibrated, as previously stated, to find the best couple of values which are able to simulate the situation in the study area.

After being calibrated the model can be applied to build future changes scenarios.

### 2.6.3 *Software used: GAMA*

The model here described was implemented in the software GAMA.

GAMA is a modelling and simulation development environment for building spatially explicit simulations [104]. It was developed and it is maintained by several teams under the umbrella of the Institut de Recherche pour le Développement (IRD/SU) international research unit which coordinated groups of research from France (University of Toulouse, University of Paris Sud, University of Rouen) and from Vietnam (University of Can Tho and University of Hanoi). GAMA offers several utilities and advantages [104]:

- As *integrated platform* it allows to switch easily between the 'code' point of view to the 'simulation' output, so that the user can see real time the effects of the changes in the code on the results;
- It integrates GIS and ABM, so that each entity in a simulation can have a geographical set of attributes (coordinates, altitude, ecc...) and more importantly it can handle all the variety of input data defined as standard geographical files defined in the GDAL library [25];
- It generates real-time statistics on the state of the system, making easier to evaluate the robustness of a simulation running;
- It is very flexible in the definition of behaviour and variables;
- In GAMA it is possible to define different neighbour rules, Von Neumann (the 4 orthogonal nearest cells are considered neighbourhood), Moore (the 8 nearest cells are considered neighbourhood) and hexagonal.

GAMA is based on GAML language, an agent-oriented language, which means that everything that is considered active (i.e.: entities in a models, models themselves, ecc...) can be represented as an agent. In other words the interaction between each agent can be setted up to compute complex and realistic simulations [48]. Thus, GAMA can support high-level model compositionality leveraging multi-modeling or multi-paradigm modeling as particular cases of composition. Agents in GAML are specified by 'species', which are provided with attributes (what they know), actions (what they can do) and reflexes (what they will actually do).

Environment is represented as 'topology' and is represented as a gridded discrete domain. The topology itself can have different attributes such as an explicit spatial coordinates, defined by a datum, different values organized in layers, and topological features. To write a

model in GAMA a user should define all the agents, their behaviour, which includes defining reflexes, actions and attributes and then drop them in the environment [48]. Timesteps have not a real value, instead the delta between present time and future time must be decided by the user. The schedule of the actions of the agents is defined by the user.

Models developed in the GAMA environment are applied in several different fields:

- Ecological spatial simulations [118, 117, 116];
- Models of traffic and transportation [70, 61];
- Risk mitigation [102, 103];
- Crisis management [73, 64, 73];

In the current work, the topology was defined in a grid of 10x10m pixels resolution. The resolution was downgraded with respect the input data (which had at least 2x2m resolution) because of the computational resources GAMA requires. The limiting factor is the quantity of RAM required which increase exponentially with the number of cells in a grid. The first layer in the topology is the DEM, used to define the orography of the area and the euclidean distance between each cell. The second layer identifies the areas where trees cannot grow: in this layer the informations about the rocky area, the major roads and the mayor water bodies were included. The third layer describes the initial condition of the forest, as the situation of the forest in the year 1954, to perform the calibration of the parameters, and as the situation of 2015 to perform the actual simulation. This layer contains values of 0 if trees are not already grown, and 1 if the trees are present in the cell.

The code works as follows: as the simulation begins all the cells are examined, and the cells containing trees are selected for acting, or spread seed. The reflex of seeding has effect in the cells of 0 value, where the probabilities of seedling and sprouting of the seeds as function of the altitude of the cell are evaluated. The last two reflexes in the current timestep are the ageing of the trees and their eventual death. The age of the trees were distributed randomly, with a gaussian distribution of mean 100 years, and a standard deviation of 20 years, the maximum age reachable by a tree is 250 years and the minimum age for the production of seed is 10 years. After a tree is dead the cell in the forest layer is set to 0.

#### 2.6.4 *The calibration of the model*

The ABM model requires a long time to be run, therefore the calibration was not run on the whole Paneveggio Pale di San Martino natural park study area, instead two different areas were chosen to perform a process of cross-calibration.

The parameters of the two probability distributions were defined by the users, then the accuracy of the simulation was measured as percentage of correct simulated cells by the model. By using as input the situation of 1954 and comparing the real situation in the years 1994, 2006 and 2015 with the simulated situation the correct cells were those where the forest is present in both the simulation and in the reference maps for the same year. The parameters that determined the greater number of correct simulated cells were chosen to perform the simulation of the forest growth in future years. The robustness of the code was then tested by performing a simulation with the calibrated parameters on the total area of study, using as initial situation the year 1954 and by comparing the real and simulated situation in the years 1994, 2006 and 2015.

The test areas were chosen in order to first calibrate the probability of seedling in an area below the treeline, in order to exclude the variable of the probability of germination. Once the probability of seedling was calibrated secondly the threshold for the probability of germinating was tested in an area around the treeline.

In figure 36 the two areas chosen for the calibration on the 1954 aerial imagery and forest map are identified.

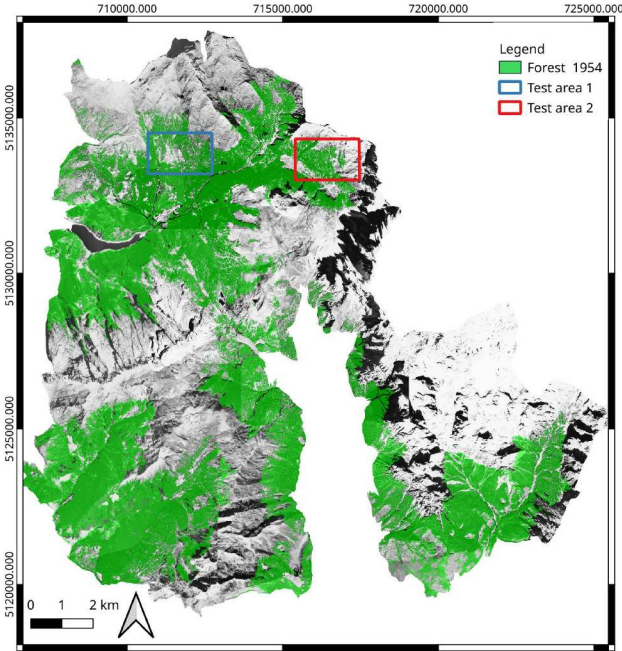


Figure 36: The position of the test areas in the Paneveggio Pale di San Martino natural park.

The results of the calibration are shown in the graphs in figures 37 and 38.

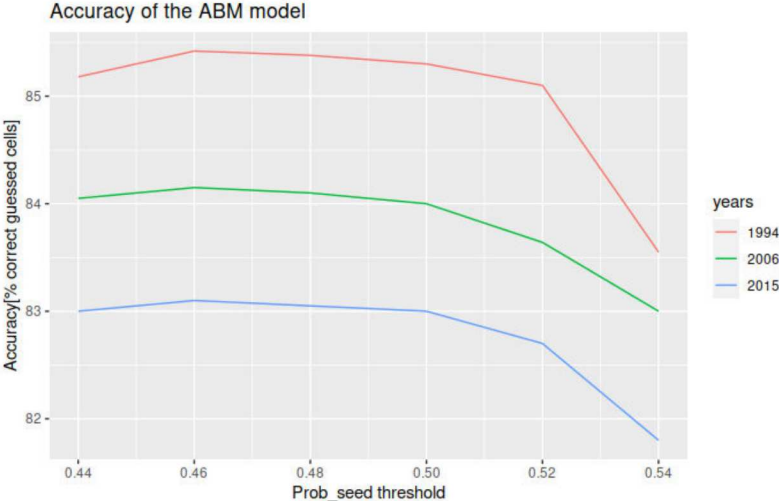


Figure 37: Calibration of prob\_seed

Threshold for prob_seed	Accuracy for 1994 [%]	Accuracy for 2006 [%]	Accuracy for 2015 [%]
0.44	85.18	84.05	83
0.46	85.42	84.15	83.1
0.48	85.38	84.1	83.05
0.5	85.3	84	83
0.52	85.1	83.64	82.7
0.54	83.55	83	81.8

Table 16: Threshold for probability of seedling and accuracy of the model over the test years.

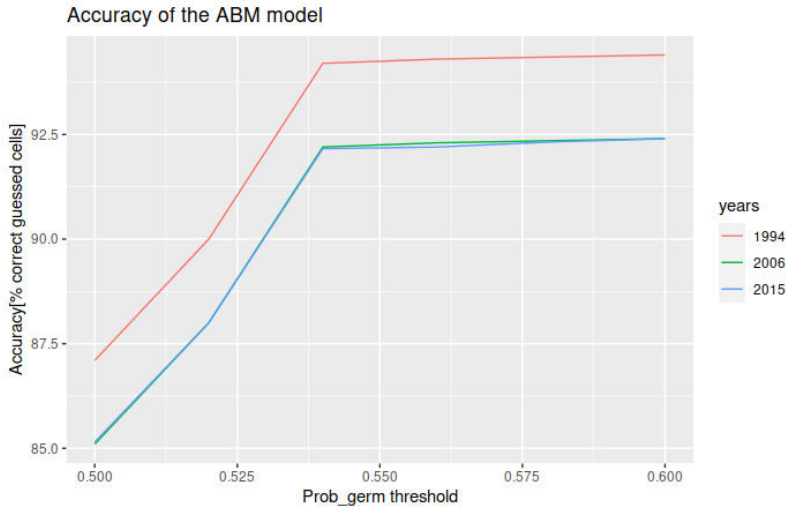


Figure 38: Calibration of prob\_germ

Threshold for prob_germ	Accuracy for 1994 [%]	Accuracy for 2006 [%]	Accuracy for 2015 [%]
0.5	87.1	85.1	85.15
0.52	90	88	88
0.54	94.2	92.2	92.16
0.56	94.3	92.3	92.2
0.58	94.35	92.35	92.32
0.6	94.4	92.4	92.4

Table 17: Threshold for probability of germination and accuracy of the model over the test years.

Tested probability	Optimal threshold for accuracy
seedling	0.46
germination	0.54

Table 18: Best values for the threshold for probability of seedling and germinating.

These values, reported in table 18 were then applied to a simulation from 1954 to 2015 extended to the whole area of Paneveggio Pale di San Martino Park, and a comparison between the real and simulated situations was performed. The accuracy was again measured

as percentage of correct guessed cells from the simulation with respect to the real situation in the same year, table 19 shows the result.

Year comparison	Accuracy of the model [%]
1994	88.0
2006	90.0
2105	91.0

Table 19: Accuracy of the model in simulating the evolution of Paneveggio Forest from 1954, comparison between the real and simulated situation in years 1994, 2006, 2015

The threshold then tested were used to perform a possible future scenario of Paneveggio forest starting from the year 2015 to reach the year 2100. The possible future changes scenarios were then compared against the future changes scenarios proposed in [20] built with a Markov chain procedure.



## 2.7 THE *landscapemetrics* PACKAGE IN THE *r* ENVIRONMENT

Landscape metrics are indexes used to describe the feature and composition of a landscape, as is described in mathematical terms, typically from a raster representing land uses [76]. These indexes aim to study the composition of a landscape or, as in the current study, to compare the evolution over time of the landscape itself or to compare different areas [53]. Metrics are based on the concept of patches, which are homogeneous areas of a single land use category (e.g.: a forest) and they can quantify the composition of the map taking into account the spatial distribution of the patches or not, depending on the specific landscape feature and corresponding metrics. Metric can be divided for the level at which they are computed: patch, class or landscape. At the patch level landscape metrics describe land cover statistics such as area, perimeter and number of patches, while at a class level they are used to describe the distribution of the class itself. Finally, at a landscape level metrics can describe the landscape in its integrity.

For the purpose of this work, a limited number of landscape metrics has been selected among the many available, that could represent and quantify the fragmentation of the forests. It was possible to calculate these metrics on the whole PAT territory without splitting the forest maps in smaller areas. The metrics were calculated at a class level.

**PATCH NUMBER (NP)** is the number of patches of each type; it is an adimensional metrics. It is defined as [75]:

$$NP = N_i \quad (3)$$

Where:

- $N_i$  is the number of patches in the  $i$  – th class.

**AGGREGATION INDEX (AI)** equals the number of adjacent patches involving the corresponding class, divided by the maximum possible number of like adjacencies, multiplied by 100 (to convert to a percentage). It is useful to quantify fragmentation of the landscape. It is defined as [75]:

$$AI = \left[ \frac{g_{ii}}{\max(g_{ii})} \right] (100) \quad (4)$$

where:

- $g_{ii}$  is the number of *like adjacencies* between pixels of patch class  $i$  based on the *single-count* method;
- $\max(g_{ii})$  is the maximum number of *like adjacencies* between pixels of patch class  $i$  based on the *single-count* method;
- 100 is a constant used to express the index as percentage.

**MEAN PATCH AREA (AREA MN)** is the average area of all the patches of a given type. It is measured in square meters. When used in combination with NP, Area MN gives information about how the patches of a given land cover class are growing or merging over time. It is defined as [75]:

$$AreaMN = \frac{A}{NP} \quad (5)$$

Where:

- A is the total area of the patches in [m<sup>2</sup>];
- NP is the Number of Patches.

PATCH DENSITY (PD) describes the fragmentation of a class, however, does not necessarily contain information about the configuration or composition of the class. It is measured as number of patches per 100 hectares and expressed as [75]:

$$PD = \frac{n_i}{A} * 10000 * 100 \quad (6)$$

where:

- n<sub>i</sub> is the number of patches for a single class;
- A is the total area of the landscape in [m<sup>2</sup>];
- 10000 is a normalising constant;
- 100 is a normalising constant;

LARGEST PATCH INDEX (LPI) is the percentage of the landscape covered by the largest patch in the landscape. It is a simple measure of dominance. It is measured in percentage and expressed as [75]:

$$LPI = \frac{\max(a_{ij})}{A} * 100 \quad (7)$$

where:

- max(a<sub>ij</sub>) is the area of the j-th largest patch for the i-th class;
- A is the total area of the landscape in [m<sup>2</sup>];
- 100 is a normalising constant;

TOTAL EDGE (TE) equals the sum of the lengths (m) of all edge segments in the landscape of a given land use. It is expressed as [75]:

$$TE = E_i \quad (8)$$

where:

- E<sub>i</sub> is the edge length for the i-th class expressed in meters;

SHANNON DIVERSITY INDEX (SHANNON) is an index of the probability of two random pixels belonging to the same class. It is a widely used metric in biodiversity and ecology and takes both the number of classes and the abundance of each class into account. It is expressed as [75]:

$$SHANNON = - \sum_{i=1}^m P_i \cdot \ln(P_i) \quad (9)$$

where:

- P<sub>i</sub> is the proportion of the i-th class.

SIMPSON DIVERSITY INDEX (SIMPSON) is an index of the probability of two random pixels belonging to the same class, as the Shannon Diversity index does, but it is less sensible to less frequent landuse classes. It is expressed as [75]:

$$\text{SHANNON} = 1 - \sum_{i=1}^m P_i^2 \quad (10)$$

where:

- $P_i$  is the proportion of the  $i$ -th class.

### 2.7.1 *The Landscapemetrics package in R*

The R environment is widely used in statistical treatment of data, manipulation and display [89]. Its popularity is due to the fact that it is a flexible Free and Open source software which allows its users to perform different types of analysis. The R environment offers an effective data handling and storage facility, a long series of operators to perform operations on complex types of variables (matrices and arrays), various systems for data display and graphs manipulations [89]. The software is organized in 'packages', which are libraries where there are defined functions dedicated to specific aims [89].

For instance, a GDAL package was developed to manipulate the geo-referenced data, as tiff rasters or vectorial data inside the R environment. The package *Landscapemetrics* offers a complete and robust set of functions to calculate landscape metrics in R.

In fact, the *Lecos* plugin in QGIS, as well as the *r.le* suite in GRASS do not allow to calculate all the metrics available, among which are some used in the current work.

*Landscapemetrics* [53] is based on *Fragstat* [76], a software widely used in the Literature for calculation of landscape metrics. The robust basis where the package *landscapemetrics* is rooted, with the flexibility and the efficiency of the R environment were exploited to implement a simple procedure to calculate the landscape metrics in section 2.7 on the various landscapes extracted from the LULC procedures from the various set of aerial imagery. A fundamental, basic parameter for the evaluation of landscape metrics is the Cell Neighbourhood (CN). It can be of two different type:

- Von Neumann, or the neighbourhood is considered as the 4 orthogonal cells of a central one;
- Moore, or the cell neighbourhood is considered as the 8 closest cells of a central one.

The shape of CN establishes how pixels are connected to form a single patches. If a pixel shares with its neighbours the same class then this will be considered a single patch. Choosing between these CN shapes obviously influences the results of a landscape metric analysis [126]. For this work an 8 CN rule was chosen.

## RESULTS AND DISCUSSION

The following section aims to present all the results reached by the current study. The first result to be displayed are the preliminary orthorectification of the 1954 dataset, a process that allowed the geographical collocation of this images set. The results of the classification procedure are then presented, discussing as well the procedure used to assess the quality of the classification. Finally, the results of the assessment of some ecosystem services and of the future changes scenarios for the Paneveggio Pale di San Martino natural Park are showcased.

### 3.1 THE 1954 ORTHORECTIFICATION OF THE DATASET

The orthorectification process of the images of 1954 resulted in a set of 91 rectified maps. These maps were chosen in a set of 200 different images, selecting the ones which better depict the features of the territory, and minimize the overlaying between the neighbour images. The images were identified by a progressive 2 digit number which indicates the flight name in the longitudinal direction, while the following 4 numbers identify the frame in the flight. Table 20 reports on the images chosen in the current work.

Flight number	Frame number
12e	4506, 4508, 4510, 5091, 5094, 5096
13e	3205, 3207, 3209, 3211, 3213, 3215
14e	3181, 3183, 3185, 3187, 3189, 3191, 3193, 4955, 4957,
15	3057, 3059, 3061, 3063, 3065, 3067, 3068, 3070, 3072, 3074
16C	4452, 4453, 4456, 4458
17C	4641, 4643, 4645, 4646, 4648, 4650, 4652, 4654, 4656
18C	4296, 4298, 4300, 4302
18	3278, 3279, 3281, 3283, 3286, 3288, 3290, 3292, 3294
19c	4362
19w	3853, 3855, 3856, 3862, 3864, 3866, 3868
20C	4372, 4374, 4376, 4378, 4380, 4388, 4389
21	4563, 4564, 4567, 4569, 4571, 4573, 4575, 4576, 4578, 4580
22	1678, 1680, 1681
22a	4601, 4602
23	1727, 1736, 1738
25	2195

Table 20: List of the orthorectified images

The image 39 shows the position of each rectified image respect to the PAT territory.

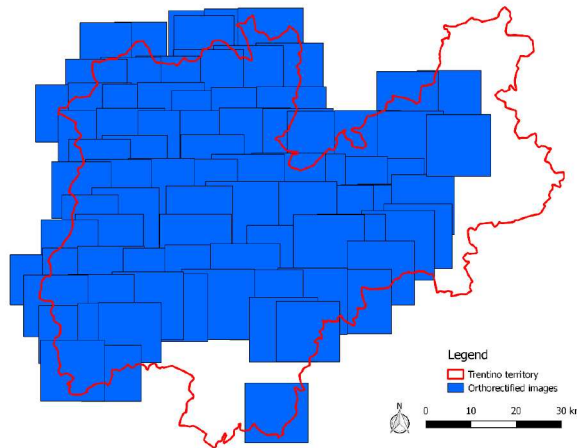


Figure 39: Index for the orthorectified images in the current work from the 1954 dataset.

As figure 39 reports, the imagery of the 1954 dataset that were rectified in the current work did not covers the entirety of PAT territory. A series of orthorectified imagery for some regions, the Monte Baldo Area, Val di Fassa, Paneveggio Pale di San Martino Park and Tesino municipality were available and kindly provided for the current work by their respective authors [72, 108, 92, 16]. These imageries and the imageries orthorectified in the current work concur to cover the entirety of PAT territory and represent the basis of the 1954 dataset. GCPs were placed by hand for every single images using as reference images the 2006 imageries set. The ground resolution of these images is 0.5x0.5m per pixel, which is higher than the resolution of the orthorectified images therefore GCP's coordinates can be considered accurate. In table 21 are summarized the mean features of the rectified images.

Ground resolution [ $m \times pixel$ ]	Covered area [ $km^2$ ]	Dimension [MB]	Scanning resolution [DPI]
2.1	140	40	600

Table 21: Features of the rectified images

Image 40 is a single rectified image: the tilt displacement of the pixel (due to a high difference in altitude between a pixel and its neighbour) is clearly visible in the central west area, where a steep valley is present, and the slopes are high. The slight rotation of the image axis depends on the direction of the flight respect to the North direction.

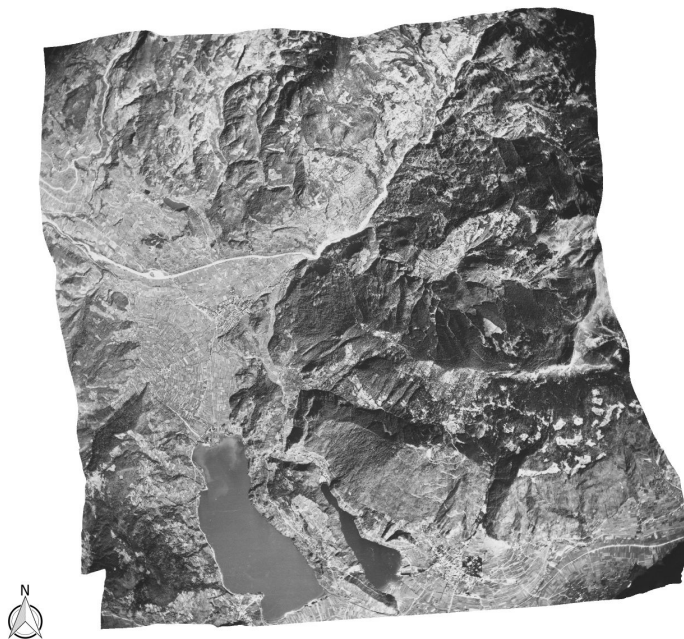


Figure 40: Rectified photo 19w\_3856

### 3.1.1 *Qualitative assessment of the precision*

The assessment of the precision of the coordinates in the resulting orthophotos begins with a precise position of the fiducial markers, which are used to bound the physical dimension of the image to the resulting ground resolution. On the premises that the calibration certificate of the camera is not available for the current dataset, the coordinates of the fiducial markers have been determined measuring as precisely as possible the distance between the marks on the physical copy of the photographs, as figure 41 shows.

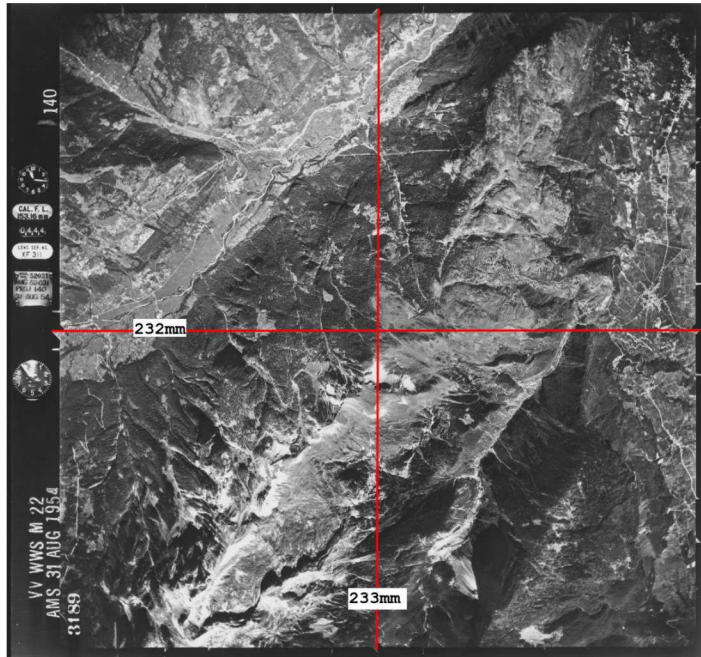


Figure 41: The measurements of the fiducial markers.

To control and provide a feedback to the user if the position of input GCPs is correct the command `g.gui.image2target` [77] outputs a series of parameters, as figure 42 shows.

#	error			image		target		elev.
	east	north	target	east	north	east	north	
1	4.0	-5.7	6.9	1935.2	1357.9	644470.9	5099670.1	443.7
2	-4.4	-2.4	5.0	1968.8	1304.9	644577.5	5099544.8	457.7
3	0.3	-3.5	3.5	2301.5	930.6	645495.3	5098694.0	527.2
4	1.1	-4.3	4.4	2367.9	536.1	645755.3	5097809.4	686.0
5	-2.4	4.6	5.2	1909.6	383.6	644631.4	5097202.6	485.8
6	4.3	6.2	7.5	1384.8	219.5	643381.5	5096759.8	580.4
7	-7.0	9.7	12.0	1401.0	1231.3	643199.0	5099264.7	519.8
8	-2.5	-4.5	5.2	1122.5	1117.2	642556.2	5098969.7	581.7
9	-4.9	-4.1	6.4	996.5	1065.2	642286.5	5098845.3	639.2
10	-3.0	8.5	9.0	2255.3	2122.9	645157.3	5101726.4	762.7
11	-6.9	7.4	10.1	2119.4	2131.9	644828.3	5101720.6	752.3
12	-2.1	4.2	4.7	1977.4	2080.9	644485.8	5101566.4	731.8
13	23.6	-16.6	28.9	1664.3	1970.3	643692.0	5101225.7	617.8

Overall rms error: 10.52  
 DONE PRINT FILE Double click on point to be included/excluded

Figure 42: Graphic User Interface for the `g.gui.image2target` command.

The last five columns shows the coordinates of the GCPs both in the XY image reference system and in the geographical reference system chosen for the orthorectification. The first three columns represent the "control parameter" for the orthorectification. On each GCPs, in fact, is performed an evaluation of the Root Mean Square Error, defined as [92]:

$$RMSE_i = \sqrt{\mu_i^2 + \nu_i^2} \quad (11)$$

where:

- $RMSE_i$  = Root Mean Square Error of the i-th GCP;

- $\mu_i$  = residual of the  $i$ -th GCP in the east-west direction;
- $\nu_i$  = residual of the  $i$ -th GCP in the north-south direction.

These RMSEs depend on the accuracy of the GCPs selection, but a for a deeper understanding, the distribution of the RMSE is discussed in section 3.1.2. If a GCP is incorrectly placed it will largely increase the RMSE of every other point: for instance as figure 44 shows RMSE on an incorrectly placed GCP was computed as double in value as the error on other GCP. It is important to underline how this evaluation of the discrepancies is possible only for the GCPs and not for every point of the image because at this stage of the orthorectification process GCPs are the only point of known coordinates both on the image and in the target datum.

The overall RMSE is defined as follows [92]:

$$\text{Overall RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N \text{RMSE}_i} \quad (12)$$

where:

- $\text{RMSE}_i$  is the Root Mean Square Error of the  $i$ -th GCP;
- $N$  is the Total number of GCPs.

In figure 43 is displayed the distribution of the RMSE for each image orthorectified.

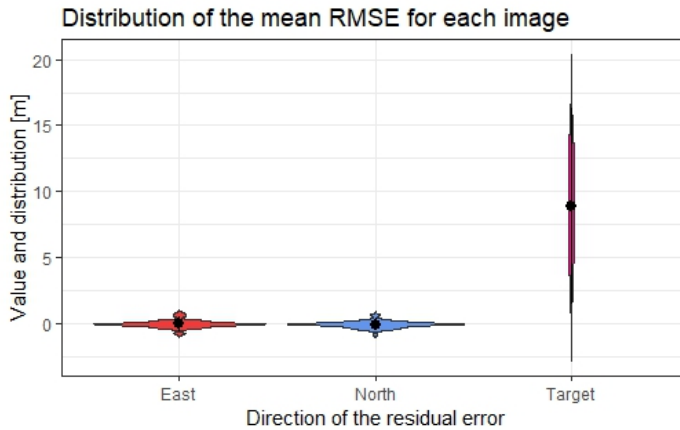


Figure 43: The distribution of the RMSE for the orthorectified images.

The errors are distributed on a wide range of values, with a mean value of around 10m, that represent an index of the overall positional accuracy reached after the orthorectification for the 1954 dataset. This wide distribution of values indicates how the value of RMSE on GCPs depends on a wide variety of factors.



#	error			image		target		elev.
	east	north	target	east	north	east	north	
1	4.0	-5.7	6.9	1935.2	1357.9	644470.9	5099670.1	443.7
2	-4.4	-2.4	5.0	1968.8	1304.9	644577.5	5099544.8	457.7
3	0.3	-3.5	3.5	2301.5	930.6	645495.3	5098694.0	527.2
4	1.1	-4.3	4.4	2367.9	536.1	645755.3	5097809.4	686.0
5	-2.4	4.6	5.2	1909.6	383.6	644631.4	5097202.6	485.8
6	4.3	6.2	7.5	1384.8	219.5	643381.5	5096759.8	580.4
7	-7.0	9.7	12.0	1401.0	1231.3	643199.0	5099264.7	519.8
8	-2.5	-4.5	5.2	1122.5	1117.2	642556.2	5098969.7	581.7
9	-4.9	-4.1	6.4	996.5	1065.2	642286.5	5098845.3	639.2
10	-3.0	8.5	9.0	2255.3	2122.9	645157.3	5101726.4	762.7
11	-6.9	7.4	10.1	2119.4	2131.9	644828.3	5101720.6	752.3
12	-2.1	4.2	4.7	1977.4	2080.9	644485.8	5101566.4	731.8
13	<b>23.6</b>	<b>-16.6</b>	<b>28.9</b>	1664.3	1970.3	643692.0	5101225.7	617.8

Overall rms error: 10.52

DONE PRINT FILE Double click on point to be included/excluded

#	error			image		target		elev.
	east	north	target	east	north	east	north	
1	<b>5.9</b>	<b>-6.5</b>	<b>8.8</b>	1935.2	1357.9	644470.9	5099670.1	443.7
2	-2.9	-3.0	4.2	1968.8	1304.9	644577.5	5099544.8	457.7
3	-1.8	-2.1	2.8	2301.5	930.6	645495.3	5098694.0	527.2
4	-2.1	-1.6	2.6	2367.9	536.1	645755.3	5097809.4	686.0
5	-4.4	3.5	5.7	1909.6	383.6	644631.4	5097202.6	485.8
6	<b>3.1</b>	<b>2.8</b>	<b>4.2</b>	1384.8	219.5	643381.5	5096759.8	580.4
7	<b>-2.4</b>	<b>5.7</b>	<b>9.0</b>	1401.0	1231.3	643199.0	5099264.7	519.8
8	2.6	-5.7	6.3	1122.5	1117.2	642556.2	5098969.7	581.7
9	0.4	-5.2	5.2	996.5	1065.2	642286.5	5098845.3	639.2
10	0.1	4.7	4.7	2255.3	2122.9	645157.3	5101726.4	762.7
11	-2.3	3.5	4.2	2119.4	2131.9	644828.3	5101720.6	752.3
12	3.8	0.7	3.8	1977.4	2080.9	644485.8	5101566.4	731.8
13		not used		1664.3	1970.3	643692.0	5101225.7	617.8

Overall rms error: 5.48

DONE PRINT FILE Double click on point to be included/excluded

Figure 44: Example on the effect of placing a wrong GCP: on the right the wrong GCP is included in the orthorectification and it increase the values of the single ORMSE for every GCP, respect to the image on the bottom, where deselecting the wrong GCP drastically decrease the ORMSE.

The concept of overall RMSE is useful to define a mean error on the position of some recognizable features of the terrain, but it is to be underlined how it depends only on mathematical computation on the image: there is no explicit dependence on the terrain geometry in the RMSE and overall RMSE equations.

It is possible to lower the RMSE on single GCPs by using a finer resolution DEM, as it is shown in table 22. Test on the same image (13e\_4945) underline how using the same GCPs and changing the DEM resolution is necessary, but not sufficient to improve the RMSE value.

DEM resolution [m]	RMSE [m]	RMSE [pixels]	# GCPs
25	18.55	8.43	11
10	15	6.8	11
5	15	6.8	11
1	13.6	6.2	11

Table 22: Variation of the precision in photograph 13e\_4954 for different choices of DEM and scanning of 600 DPI.

Improving the resolution of the DEM is worth as well for improving the final quality of the image, especially in those areas where there are steep slopes. The algorithm tries to fit the image to the elevation model, resulting in step shaped stretches when the slopes are too steep and no GCPs are defined in the neighbourhood (figure 45).

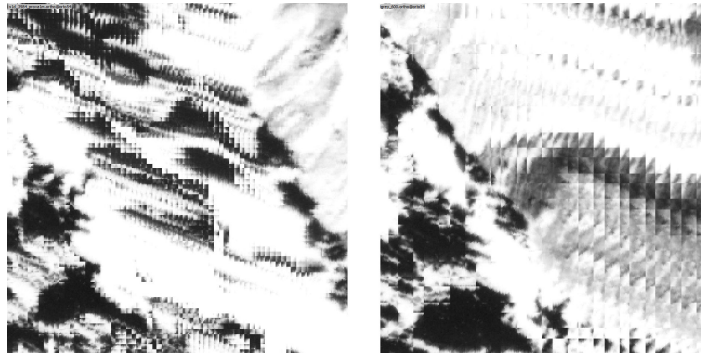


Figure 45: Comparison on the same area of orthophoto 13e\_4954 with different DEM used for rectification (1x1m resolution on the left, 10x10m on the right).

The *step stretch* is reduced by using a finer resolution DEM: the same area as in figure 45 at a wider zoom, it is possible to notice an improvement of the visual interpretation of the terrain.

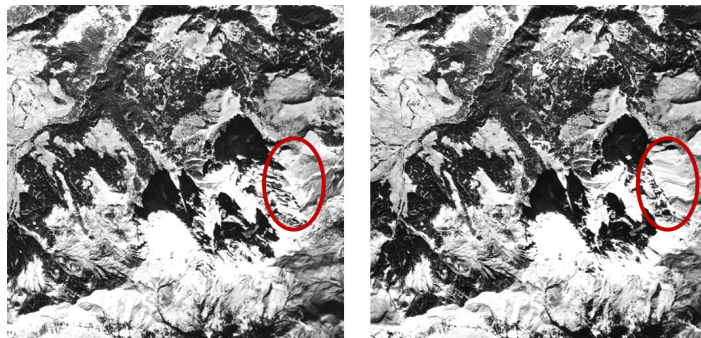


Figure 46: Comparison on the same area of orthophoto 13e\_4954 with different DEM used for rectification: the left figure (DEM resolution of 1x1m, on the right a DEM resolution of 10x10m) guarantees a better interpretation of the highlighted area.

### 3.1.2 Statistical assessment of the precision

The positional accuracy of the pixels in the image which is being orthorectified is sensitive to errors in the placement of GCPs, specific configuration of fiducial markers in relation to GCPs and local slope at the coordinates of the GCPs [62]. The assessment of the quality of the results of an orthorectification procedure is usually done by evaluating the standard deviations of the internal and external orientation steps [62]. The values are evaluated with respect to values expected in similar configurations in terms of geometry, mainly flight mean height and focal length, image quality and GCPs coordinates accuracy [62]. During the internal and external orientation procedure, residuals for each point can be used to identify less accurate points and deselect them so that they are not used.

While these parameters can provide an indication of the precision of the overall orthorectification process, the final user is interested in the geometric precision of the resulting orthorectified image. Ideally, the cumulative effect of errors in the procedure should result in a standard deviation in the ground coordinates smaller than the orthophoto resolution. The actual procedure follows an inverse path: the grid of the coordinates of the centres of the

orthophoto pixels is set and the corresponding coordinates on the image are determined. A more relevant assessment of the accuracy of an orthophoto is carried out by comparing the coordinates of points on the image with the coordinates of the same points from external sources [54]. Such points are called independent control points (ICPs) and their coordinates are obtained in the same way the coordinates of GCPs are determined. In the present study this analysis was performed by comparing the rectified images to the topographical maps (1:10 000) of PAT. The command *g.gui.GCP* was used to set the ICPs. Their coordinates in both the source (orthophotos) and target (topographical maps) were stored in a vector file and then their differences was calculated using *db.execute*. Ideally the same feature depicted by both maps should have no difference in coordinates the value 0, but due to the orthorectification process this difference is variable. The local slope around the ICPs was calculated, as well as the local altitude, and their distance from the image centre. These quantities were related to the difference between the source and destination coordinates, with a Pearson's R. Tables 24, 27, 28, 29 report the results obtained by such analysis, performed on a set of 5 images.

For the statistical assessment of the precision, the analysis on the ICPs were carried out on 5 different images. The choice was based on the image with the highest (20c\_4384) and lowest (20c\_4378) ORMSE. Another parameter taken into account was the normalized ORMSE. The result of this analysis revealed that image 23\_1738 has the lowest ORMSE-GCPs ratio and image 12e\_5096 has the highest ratio (see table 23). The number of the ICPs was chosen in order to be equal to the number of GCPs to have the same sample size from both the elaboration and avoid problems in estimate statistical mean value.

Image index	ORMSE [m]	ORMSE per GCP [m]
20c_4378	1.28	0.64
13e_3207	9.68	0.44
20c_4384	17.02	1.06
23_1738	7.69	0.31
12e_5096	7.08	1.95

Table 23: Images chosen for statistical analysis and relate ORMSE

Tables 24 and 25 underline an important aspect of the accuracy analysis. The first set of ICPs were placed using as source image the photo 20c\_4378, rectified using only 4 points. The ORMSE was calculated as 1.28m (table 23) in GRASS GIS, but the overall assessment of the precision revealed a displacement on the ICPs much greater than the ORMSE. When placing few points the calculus of the ORMSE should not be considered enough reliable for the assessment of the precision in the final result. The second set of ICPs (table 25) were placed using as source the neighbours images of 20c\_4378, which overlapped the image for more than the 70% of its area. Their GCPs set was larger and the orthorectification resulted more precise.

RE E [m]	RE N [m]	TRE [m]
-43.93	-14.29	46.19
-24.34	-1.94	24.42
8.97	29.44	30.77
-70.11	-30.95	76.63
18.93	16.97	25.42
16.98	7.51	18.57
29.52	20.24	35.79
Mean TRE		36.83

Table 24: First set of ICPs for the image 20c\_4378.

RE E [m]	RE N [m]	TRE [m]
16.07	12.55	20.39
-10.29	-5.40	11.62
6.36	8.44	10.56
-3.3	-6.95	7.69
12.00	5.97	13.40
7.4	15.26	16.95
1.05	1.23	1.6
Mean TRE		10.27

Table 25: Second set of ICPs for the image 20c\_4378.

The tables 26 - 30 aims to study the correlation between residual errors and some features of the landscape (slope and elevation) and the distance from the camera centre in each ICPs. The acronyms in the header of tables 24 - 30 refer to:

- RE E= Residual Error in East direction;
- RE N= Residual Error in North direction;
- Total Residual Error, calculated as square root of the sum of the square of RE N and RE E;
- DCC= Distance from Camera Center, distance of the ICPs from the centre of the image.

RE E [m]	RE N [m]	TRE [m]	slope [°]	Altitude [m]	DCC [m]
-1.3	-5.3	5.4	7.2	528.5	5281.9
-2.6	-13.7	13.9	7.6	581.8	4740.5
6.7	24.9	25.8	3.3	651.6	4780.3
-10.1	13.7	17.1	6.6	621.8	5014.1
-0.7	25.9	25.9	4.6	652.6	4808.2
1.5	-11.2	11.3	16.9	640.4	5140.2
0.6	-1.5	1.6	13.7	729.9	4563.6
-4.3	-6.5	7.8	0.9	526.4	4301.2
-5.2	21.6	22.3	7.2	632.7	5039.1
21.6	13.8	25.6	5.0	677.6	3443.3
18.4	5.9	19.4	3.0	717.3	2560.9
-3.8	7.5	8.4	4.8	987.3	2419.2
Mean TRE		15.4			
Pearson's $r$ (TRE vs Slope, Altitude, DCC)			-0,4071	-0,0653	-0,0360

Table 26: Measurements on the ICPs for the image 20c\_4378.

RE E [m]	RE N [m]	TRE [m]	slope [°]	Altitude [m]	DCC [m]
15.74	-1.35	15.80	6.49	647.94	1743.51
-1.05	-3.62	3.77	2.13	670.52	2655.00
3.92	2.44	4.62	4.35	668.58	1965.42
6.63	2.86	7.22	3.85	653.75	1728.14
-0.28	-6.54	6.54	9.77	677.83	1821.39
3.93	3.28	5.12	13.38	634.71	3094.93
12.69	-14.62	19.36	16.79	1029.32	2812.22
12.79	-2.37	13.01	3.53	710.54	1659.66
11.92	-7.27	13.96	11.43	665.48	1131.24
-3.64	1.63	3.99	5.15	708.24	2283.39
-12.55	-4.07	13.19	11.59	593.01	2615.37
-24.72	-0.07	24.72	3.72	621.59	2653.50
3.42	-6.04	6.94	4.00	701.19	1847.67
0.25	-12.44	12.44	15.53	678.42	1097.12
-7.87	-15.87	17.71	16.00	724.39	2761.96
0.74	5.77	5.82	22.76	686.09	2538.26
-3.70	-11.69	12.26	9.77	722.08	3344.65
3.76	-4.11	5.57	3.48	711.24	1706.41
-14.16	-16.77	21.95	22.26	609.15	3674.14
-0.83	-10.40	10.44	16.64	571.90	2712.67
3.40	8.11	8.79	8.83	601.44	2564.11
-8.76	-1.14	8.84	6.74	698.99	2403.06
-9.35	-12.63	15.71	8.85	673.07	2857.87
2.34	-3.32	4.06	3.62	600.06	1053.17
Mean TRE		10.1			
Pearson's $r$ (TRE vs Slope, Altitude, DCC)			0.3536	0.1752	0.3444

Table 27: Measurements on the ICPs for the image 13e\_3207

RE E [m]	RE N [m]	TRE [m]	slope [°]	Altitude [m]	DCC [m]
-4.03	-0.79	4.11	2.40	183.86	3545.98
-4.54	0.14	4.54	6.28	310.04	2363.54
-12.00	14.00	18.44	0.29	184.05	5165.27
-10.97	26.29	28.49	0.65	200.05	6944.85
-10.09	18.49	21.07	0.03	181.81	5616.61
6.22	-14.25	15.55	16.84	1133.55	4078.72
-2.26	-2.77	3.57	3.95	486.38	2175.63
3.73	12.88	13.41	8.61	1027.95	3855.71
8.42	10.84	13.72	2.65	693.19	4001.53
-4.91	-1.71	5.21	13.53	680.70	434.71
-7.21	5.09	8.82	32.39	917.48	1691.15
-0.46	-5.24	5.26	5.47	1211.46	2918.13
4.84	8.63	9.90	0.70	690.12	2222.94
5.86	10.33	11.88	1.37	450.27	5916.51
18.62	4.89	19.25	17.87	552.64	5037.21
1.77	0.17	1.77	14.85	527.54	5182.02
Mean TRE		11.56			
Pearson's $r$ (TRE vs Slope, Altitude, DCC)			-0.2053	-0.2438	0.6888

Table 28: Measurements on the ICPs for the image 20c\_4384

RE E [m]	RE N [m]	TRE [m]	slope [°]	Altitude [m]	DCC [m]
18.04	15.08	23.52	21.86	997.16	2651.10
12.70	12.39	17.74	16.63	814.50	621994.42
12.13	14.85	19.17	12.84	824.11	622089.98
10.02	-1.55	10.13	16.82	742.66	623583.15
-0.37	-0.16	0.41	0.11	385.90	622042.87
5.66	6.30	8.47	1.99	379.91	620551.69
-2.91	2.21	3.66	24.22	1327.17	617634.46
0.89	2.37	2.53	14.47	1470.01	617955.66
3.28	3.14	4.54	0.54	380.52	621136.10
15.11	19.59	24.74	19.12	1384.24	624073.61
24.16	33.98	41.70	20.96	1470.56	623112.77
22.91	32.30	39.60	17.00	1504.08	623283.25
19.04	29.06	34.74	11.96	1156.42	621614.79
1.89	0.41	1.94	1.49	387.56	622121.91
0.98	5.88	5.96	0.58	383.43	621151.15
4.23	3.44	5.46	13.46	1322.92	619132.42
-0.61	5.42	5.46	2.47	379.77	620317.02
-0.60	7.19	7.22	0.56	378.28	619507.42
8.15	8.72	11.94	2.78	388.79	621562.24
8.97	5.87	10.72	20.35	473.62	623472.48
-0.50	-3.61	3.64	27.33	583.09	621209.35
-1.05	-4.72	4.84	14.39	1396.76	619956.59
1.50	-0.19	1.51	17.36	1578.30	618130.58
-1.77	-0.99	2.03	24.65	1202.15	620328.05
21.53	10.55	23.98	15.71	1121.77	623819.08
1.61	9.19	9.32	0.18	376.36	619819.62
0.95	1.43	1.72	26.58	1095.35	619208.54
Mean TRE		12.09			
Pearson's $r$ (TRE vs Slope. Altitude. DCC)			0.2194	0.3675	-0.1811

Table 29: Measurements on the ICPs for the image 23\_1738



RE E [m]	RE N [m]	TRE [m]	slope [°]	Altitude [m]	DCC [m]
6.86	0.91	6.92	22.04	3039.98	2928.96
-2.41	-1.78	3.00	19.52	3156.16	2615.44
-0.77	2.344	2.47	27.97	3118.79	2675.04
-0.22	0.03	0.23	23.57	3327.36	3725.93
hline Mean TRE		3.15			
Pearson's $r$ (TRE vs Slope. Altitude. DCC)			-0.3315	-0.9087	-0.4947

Table 30: Measurements on the ICPs for the image 12e\_5096

The image 12e\_5096 depicts a high mountain area with few recognizable features, and for this reason only 4 points were chosen as GCPs. For the same reason ICPs and GCPs are coincident, so the analysis of the displacement could be statistically dependent from this concurrence.

From the calculated correlation coefficients it is not possible to tell if there is a strong dependence between orography and displacement. In fact, table 30 shows a strong negative correlation between the altitude of the ICPs and the displacement, in tables 28 and 24 the same correlation is shown, but weaker. Tables 27 and 29 on the contrary show a weak positive correlation. The same result emerges from the correlation between the local slope in the ICPs and their distance from the image centres. However, it is worth noting how the Pearson's  $r$  of slope/displacement and altitude/displacement are concordant in sign.

The accuracy of the orthorectified image depends on a set of different factors, such as the mutual position of the GCPs and respect to the fiducial markers, the digital elevation model set for the orthorectification.

A process of orthorectification as precise as possible should be performed by placing the higher number possible of GCPs, then by a process of elimination only the combination of GCPs that yields the lowest RMSE possible should be used for the orthorectification.

During the current study different land features were chosen as GCPs. The less reliable features were the one linked to landscape shapes, for instance rocks in vantage points, mountain peaks and feature linked to hydrography. Although it may seem like these points will not change their position in relatively small time span, their visibility greatly change depending on the light condition, on the time of the day and year the aerial picture was taken. Features linked to the hydrography, such as river bends, gulfs in lakes and so on are mainly shaped by the water level, even though they are clearly recognizable in different pictures of the same area.

Man made structures are more reliable, because they are clearly recognizable in different pictures independently from the light condition, because their regular shape and geometry. For the current work the position of (mountain huts) was every time set as a GCPs, because their elongated shape inside open areas are clearly recognizable. Roads intersections provides another reliable source of GCPs, but it is worth reminding that in the last century road intersections were massively reshaped, especially for main roads, passing from direct 90° crossing to roundabouts or mild angle intersection, making difficult to identify the same point on the starting picture and target images. However, secondary roads often maintained the shape of their intersections and it was possible to use them as reliable GCPs. Inner parts of hairpin bends can be used as GCPs, with the added value of providing valuable GCPs on steep sides.

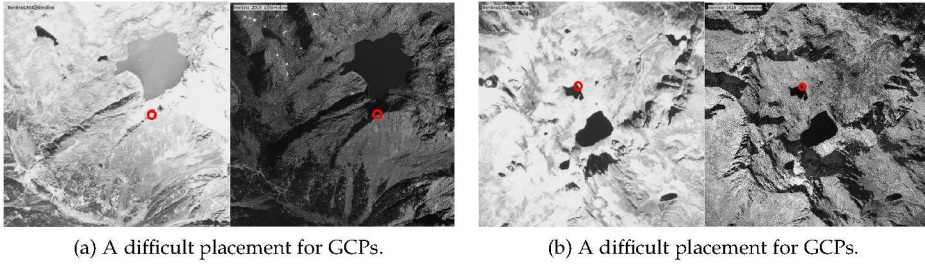


Figure 47: Examples of unreliable natural features for placing GCPs: mountain peaks are easily recognizable, in the 2015 (on the right) image, while they are not in the 1954 imagery. For an unbiased comparison both images are shown in greyscale.

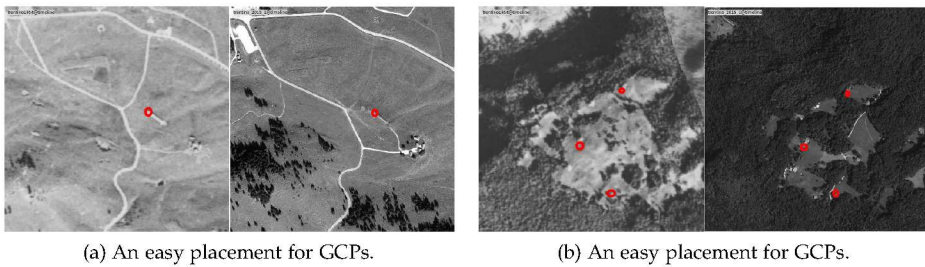


Figure 48: Examples of reliable anthropic features for placing GCPs: malghe and stables are easily recognizable in both the imagery set.

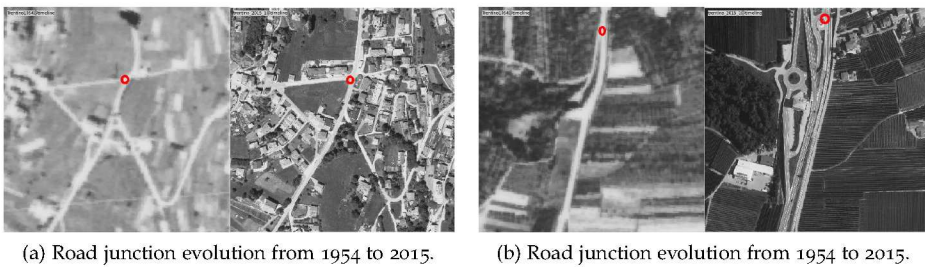


Figure 49: Examples of unreliable anthropic features for placing GCPs: some road junction changed drastically shape making it difficult to exactly recognize the same feature usable as GCP.



Figure 50: The shape of hairpin bends, another anthropic feature usable for GCPs.

### 3.2 THE DIGITIZATION OF 1859 DATASETS

The digitization of the 1859 dataset aimed to provide a map of the areas indicated as forests in the cadastral maps, to give an indication of the shape of the forests in PAT of 160 years ago.

The 1859 dataset was manually digitized for the areas indicated as 'forest', 'meadows and pastures', and 'wooded pastures' by the redactor of the 1859 land registry.

Wooded pastures are a particular type of land-use where sparse trees (usually beeches or larches) were distributed over meadows to let cattle rests in the shadows. The results for all the PAT territory are visible in the figure 51.

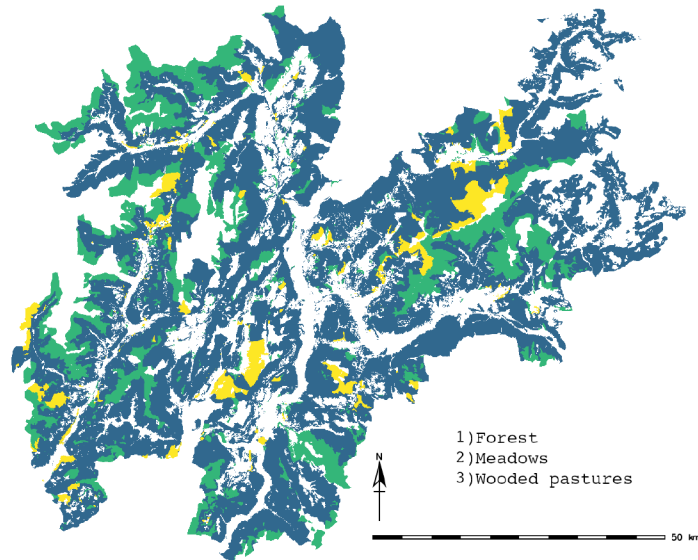


Figure 51: The results of the digitization of the 1859 H.C.P.A.T.

The areas were digitised manually by tracing the boundary of the forest as precisely as possible above the boundary of the forest areas drawn on the cadastral maps. Neighbour cadastral parcels which were registered as forests were considered a single area of forest, to make the process faster. The maps of forest area were built by displaying one municipality at a time in GRASS and digitizing one municipality at a time.

After the manual digitization was finished the resulting maps were topologically validated to avoid invalid geometries, which cannot be further analysed. The topological errors were then manually corrected. Neighbour digitized areas share a single common boundary line, there is no blank space between them. This was made to make the digitization of the landscape as continuous as possible.

It is worth reminding that these data are substantially different from the results of the forest coverage obtained from the aerial images, because the presence of forest is indicated as a compact area, where we have no indication on the actual distribution of the trees. For a qualitative comparison between the situation of the 1859 and the rest of the maps it was made a strong hypothesis: the distribution of the trees in areas of forest for the 1859 maps was considered the same as the 1954 trees distribution, that is the closest mapset in time to the H.C.P.A.T.. The same hypothesis was made for the I.K.F.M., the other thematic maps

considered in the analysis. This was achieved in GRASS GIS by using the *r.mapcalc* module with the following function:

$$\text{if}(\text{forest}_{1859} = \text{true AND forest}_{1954} = \text{true}) \text{ then trees}_{1859} = \text{true} \quad (13)$$

where:

- $\text{forest}_{1859}$  is the map of the forest extracted for the 1859 maps;
- $\text{forest}_{1954}$  is the map of the forest extracted for the 1954 aerial images;
- $\text{trees}_{1859}$  is the resulting map of forest distribution of the 1859 mapset.

Using this method the trees of 1859 were distributed as the trees sensed in 1954, only in those areas recognized as forest in 1859.

In figure 52 a map of the forest coverage of the 1859 dataset under this hypothesis is reported

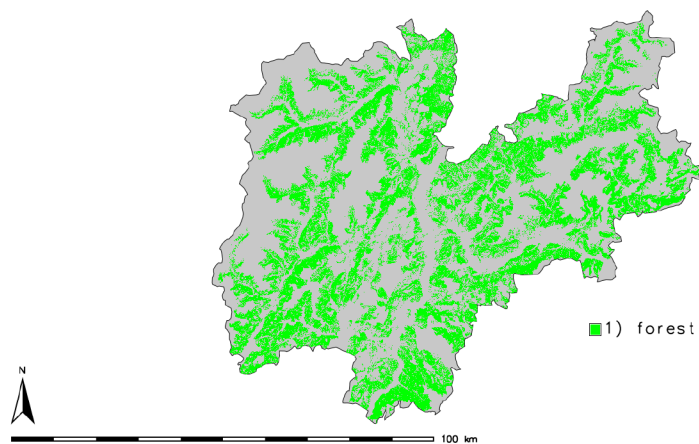


Figure 52: A simulation of the distribution of the trees in 1859.

To analyse the wooded pasture composition, the distribution of this particular landuse was crossed with the map of forest classes described in section 2.2.5. Although the two dataset differ by 150 years, this comparison was set up to understand in which typology of forest wooded pastures evolved.

Forest type	Area [ha]
Holm Oak forests	231
Mixed Oak-Hornbeam forests	614
Hornbeam forests	39
Chestnut and Black Locust forests	287
Maple, Linden, Ash forests	12
Pioneer formations	163
Scots Pine forests	1681
Beech forests	1686
Fir forests	392
Mountain Pine forests	715
Alder tree forests	920
Spruce forests	3854
Larch forests	2757
Larch and stone pine forests	1684
Stone pine forests	50

Table 31: Forest types in wooded pastures

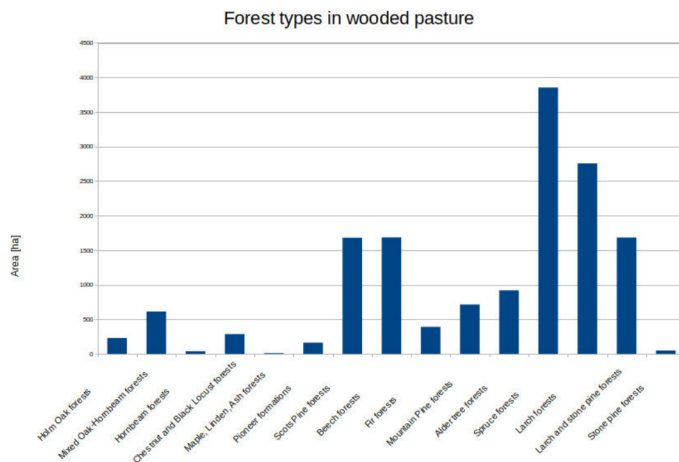


Figure 53: Forest types in wooded pastures

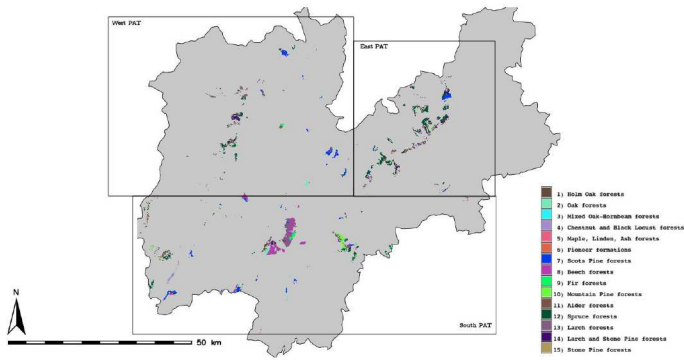


Figure 54: Forest types in wooded pastures

The most present forest type in wooded pasture is the spruce forest, which is also the most recurrent forest class in PAT. Most of the wooded pasture area evolved in Larch, Scots Pine and Beech forest. The Beech forests is mostly present in the southern area of PAT, whereas the other typologies are mostly present in the eastern and western part of PAT. The areas of wooded pasture were located in different altitudes across these three areas as shown in table 32. Beeches, a specie which grows at lower altitude, was mostly present in the wooded pastures areas of south PAT, while the resinous species were preferred in the more elevated pastures in the eastern and western part of PAT.

Zone	Mean Altitude of wooded pasture [mamsl]
Southern PAT	1300
Eastern PAT	1800
Western PAT	1400

Table 32: Mean altitude of the wooded pasture for different areas of PAT

Although the strong assumption that the forest types have not changed since 1859 certainly introduces some errors, it is possible to assert that the maps of forest types is enough well-representative of the situation in those years. This is confirmed by the fact that many of the former wooded pasture areas were occupied by a sparse cultivation of the same species that can be found today as it is confirmed by different historical sources [124, 110].

Finally, all the images for each municipality of PAT were collected from the OpenKat website and analysed one municipality at a time to extract the relevant information for the current case study. During this process were identified misplacement errors, due to the conversion of the maps reference system. The maps of the municipality of Castel di Fiemme, in particular, were not correctly referenced, this issue was reported and corrected by the owner of the maps. Finally, some portions of the maps from the municipality of Daone were missing: this issue was reported as well and corrected by the owner of the maps.

To sum up, the analysis of the 1859 H.C.P.A.T. revealed as a valuable source of information about the distribution of the forest in the XIX century, and provided a basis for the analysis of the evolution of the forests in PAT.

### 3.2.1 *New Tools for the Classification and Filtering of Historical Maps*

The following papers present a new method used to enhance the OBIA classification for some thematic maps, such as historic and cadastral maps. During the current research the maps of 1859 were manually classified because symbols and markings, typical features of cadastral and topographical maps interferes with the classification results. For future needs, a new module in GRASS GIS was developed. This module requires as input the results of a classification and a category of the classification that the users wants to erase. The pixels belonging to this category will be replaced iteratively with pixels belonging to the nearest category. In the case of cadastral maps, the user need to set up classification scheme where a separate class identifies symbols. The module will blend the 'symbol' class with the closest land use category, [125, 46]. The first paper "Object-based image analysis for historic maps classification" was published in the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, volume 42 4W14 in 2019 [125]. The second paper, "New Tools for the Classification and Filtering of Historical Maps" was published in ISPRS International Journal of Geo-Information, volume 8, in 2019 [46].



# OBJECT-BASED IMAGE ANALYSIS FOR HISTORIC MAPS CLASSIFICATION

P. Zatelli<sup>1</sup>, S. Gobbi<sup>1,2,3\*</sup>, C. Tattoni<sup>1</sup>, N. La Porta<sup>2,3</sup>, M. Ciolli<sup>1</sup>

<sup>1</sup> Università degli Studi di Trento, Dipartimento di Ingegneria Civile, Ambientale e Meccanica, Trento, Italy - (paolo.zatelli, stefano.gobbi, clara.tattoni, marco.ciolli)@unitn.it

<sup>2</sup> IASMA Research and Innovation Centre, Fondazione Edmund Mach, San Michele a/Adige, Trento, Italy - nicola.laporta@fmach.it

<sup>3</sup> The EFI Project Centre on Mountain Forests (MOUNTFOR), San Michele a/Adige, Trento, Italy

**KEY WORDS:** Heritage maps, classification, OBIA, image filtering, land use, cadaster

## ABSTRACT:

Heritage maps represent fundamental information for the study of the evolution of a region, especially in terms of landscape and ecologic features. Historical maps present two kinds of hurdle before they can be used in a modern GIS: they must be geometrically corrected to correspond to the datum in use and they must be classified to exploit the information they contain. This study deals the latter problem: the Historical Cadaster Map, created between 1851 and 1861, for the Trentino region in the North of Italy is available as a collection of maps in the ETRS89/UTM 32N datum. The map is a high resolution scan (230 DPI, 24 bit) of the original map and has been used in several ecological studies, since it provides detailed information not only about land property but also about land use. In the past the cadaster map has been manually digitized and for each area a set of attributes has been recorded. Since this approach is time consuming and prone to errors, automatic and semi-automatic procedures have been tested. Traditional image classification techniques, such as maximum likelihood classification, supervised or un-supervised, pixelwise and contextual, do not provide satisfactory results for many reasons: map colors are very variable within the same area, symbols and characters are used to identify cadaster parcels and locations, lines, drawn by hand on the original map, have variable thickness and colors. The availability of FOSS tools for the Object-based Image Analysis (OBIA) has made possible the application of this technique to the cadaster map. This paper describes the use of GRASS GIS and R for the implementation of the OBIA approach for the supervised classification of the historic cadaster map. It describes the determination of the optimal segments, the choice of their attributes and relevant statistics, and their classification. The result has been evaluated with respect to a manually digitized map using Cohen's Kappa and the analysis of the confusion matrix. The result of the OBIA classification has also been compared to the classification of the same map using maximum likelihood classification, un-supervised and supervised, both pixelwise and contextual. The OBIA approach has provided very satisfactory results with the ability to automatically remove the background and symbols and characters, creating a ready to be used classified map. This study highlights the effectiveness of the OBIA processing chain available in the FOSS4G ecosystem, and in particular the added value of the interoperability between GRASS GIS and R.

## 1. INTRODUCTION

The availability of long time series for the description of the evolution the landscape and of the ecologic features of a region is fundamental (Antrop, 2005) (Tattoni et al., 2017). While the analysis of historical maps to extend the time span of available time series has been investigated for some time (Haase et al., 2007), the number of studies exploiting this possibility is still limited (Dittrich et al., 2017).

There are several reasons for this: on one hand suitable historical maps are uncommon, on the other the use of these maps for use in a Geographic Information System (GIS) system requires substantial pre processing. For historical maps to be suitable for landscape and ecological studies, they must homogeneously cover the study area and feature information about land use/land cover (LULC) (Cantiani et al., 2016).

These two conditions are seldom met, as most historical maps have been created for specific purposes and do not carry this type of information. Even when a suitable map is available, the pre processing needed to prepare the map for its use in a GIS can be challenging. In fact, maps must be digitized, reprojected in the reference system and projection in use, and classified into LULC classes of interest.

This paper describes a new approach for the semi-automatic classification of historical maps and the automatic removal of

unwanted artifacts, such as text and symbols on the map. This approach combines R (R Core Team, 2019) and GRASS GIS (GRASS Development Team, 2019) capabilities to digitize and classify an historical map using Object-based Image Analysis (OBIA). The classification is followed by a tailored filtering process to remove text strings and symbols from the map.

## 2. MATERIALS AND METHODS

The main purpose of this research is to develop a semi-automatic procedure for the classification and digitalization of historical maps. The procedure has been set up for the Historical Cadaster Map of the Trentino region and tested on this map.

Historical cadaster maps have already been used to describe the evolution of landscape and ecological features in other parts of Italy (Agnoletti, 2007) and Europe (Forejt et al., 2018), (Skaloš , Engstová, 2010), but the maps have been always digitized by manually creating vector areas in a GIS. Even the same Historical Cadaster Map of the Trentino region has already been used after manual digitalization (Tattoni et al., 2010) (Ciolli et al., 2012), but only on a small region. Given the large area covered by the cadaster map, this approach would be time consuming and prone to errors.

\*Corresponding author

## 2.1 Historic Cadaster Map

The Historical Cadaster Map for the Province of Trento has been created as part of the cadaster map of the Austrian Empire, under the Royal Imperial Patents of the 23rd of December, 1817 by Francis I, the first Emperor of Austria (Buffoni et al., 2016). The reference network has been established in the Trentino region between 1851 and 1855 (Servizio Catasto della Provincia Autonoma di Trento, 2019a), with the complete survey carried out between 1855 and 1861 (Buffoni et al., 2016).

The purpose of this map is not only the geometric description of objects and property boundaries, but also the representation of LULC, with particular attention to different agricultural uses, and natural boundaries. This map is therefore suitable for the analysis of landscape and ecologic features, since LULC patches can be recognized.

The original maps, 52,68 x 65,85 cm wide, covering 288 ha, have been digitized at 230 DPI, 24 bit color depth, and the complete map for the Trentino region is available as a set of JPEG images (Servizio Catasto della Provincia Autonoma di Trento, 2019b). Each image contains one of the 13297 original single map sheets or one of the map mosaics; files have dimension between 5 and 8 Mb each. They are accompanied by a JGW world file with coordinates in the ETRS89/UTM 32 N (EPSG 25832) datum. Additional material, such as surveying maps, cadastral borders descriptions, pictorial and geometric drawings, cadastral monographies and triangulation points are available on a dedicated web site (Servizio Catasto della Provincia Autonoma di Trento, 2019a).

The digital map is available under the Creative Commons Attribution 4.0 license (OPENDATA Trentino, 2019).

All tests have been carried out on the map sheets 22.03, 22.04 and 22.05 of the Comune catastale di Cunevo (Cadastral municipality of Cunevo) because they are representative of all the possible combinations of classes and objects sizes. Map sheet 22.05 is shown in figure 1: all successive images and tables refer to this map.

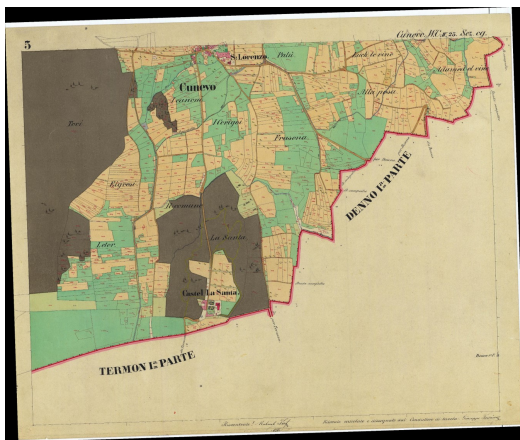


Figure 1. Map sheet 22.05 of the Comune catastale di Cunevo used for the tests.

## 2.2 Geo-referentiation

Digital maps available at (Servizio Catasto della Provincia Autonoma di Trento, 2019b) are already geo-referenced

in the ETRS89/UTM 32 N (EPSG 25832) datum. The geo-referentiation has been carried out by the Servizio Catasto (Cadastral Service) of the Provincia Autonoma di Trento using the four corners of each sheet, whose ETRS89/UTM 32 N coordinates have been previously determined. Coordinates accuracy is of 5-10 m, but can be improved locally by applying a successive transformation using local Ground Control Points, achieving a 1-2 m accuracy (Revolti, R., 2017). The ground resolution of the map is around 0.32 m.

## 2.3 Classification

The classification of historical maps is usually performed by visual interpretation and manual classification, see e.g. (Statuto et al., 2017) (Forejt et al., 2018). This requires the manual digitalization of each parcel boundary and the input of the parcel number and LULC class. This approach is not only time consuming but also prone to errors:

1. boundary location can be subjective and dependent on the scale/zoom level used during the digitalization
2. errors can be made performing a repetitive task
3. wrong LULC classes can be assigned
4. topology errors are frequent, in particular with overlapping polygons or gaps between them

These errors can be partly removed by adding post-processing steps to the procedure, but an automatic or semi-automatic approach would avoid many of these problems altogether.

The classification of the cadaster historical map has been carried out using the image classification methods available in GRASS:

1. Maximum Likelihood Classification (MLC)
  - (a) unsupervised
  - (b) supervised
  - (c) supervised contextual
2. Object-based Image Analysis (OBIA)

The first approach has already been tested for other maps (Ferretti et al., 2018) with unsatisfactory results and is used in this study only as a reference to evaluate the classification step of OBIA. For this reason it will not be described in detail. The output maps of the unsupervised MLC, supervised MLC and supervised contextual MLC are shown in figures 2, 3 and 4 respectively for the LULC classes of table 1; colors for classes have been chosen to mimic the colors on the original map. Note that for unsupervised MLC it has not been possible to find a one to one correspondence between classes on the map and LULC classes of table 1. Classes 6 (Black strings and symbols) and 10 (Red strings and symbols) are used to identify text and symbols on the map and will be filtered in the final map.

An assessment of the performance of MLC is reported in table 3.

## 2.4 Object-based Image Analysis

OBIA consists in the creation on the image of "segments" (homogeneous areas) whose radiometric and geometric features are used to classify them. This approach is available in GRASS (Grippa et al., 2017) though a set of add-on modules. The procedure is accomplished in two steps: segment creation and classification.

Class	Description
1	Forest
2	Crop land
3	Meadow
4	Pasture
5	Buildings
6	Black strings and symbols
7	Roads
8	Boundary lines
9	Gardens and courts
10	Red strings and symbols

Table 1. LULC classes used for MLC

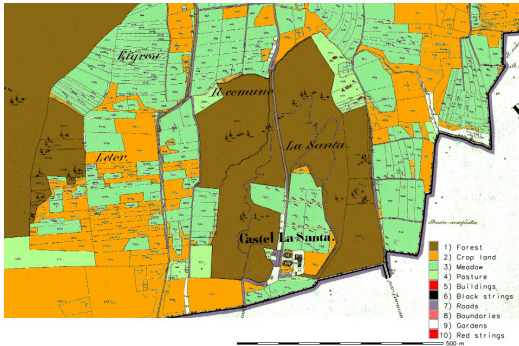


Figure 2. Classified map for unsupervised MLC - particular.

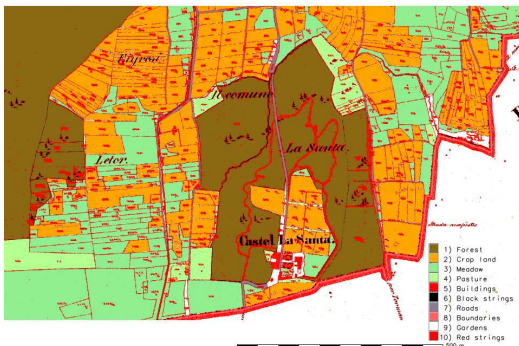


Figure 3. Classified map for supervised MLC - particular.

**2.4.1 Segmentation** Segment creation in GRASS is performed by the *i.segment* module. The region growing image segmentation approach has been used in this study, but mean shift image segmentation is also available using the same module.

The region growing algorithm is controlled by two parameters: a *similarity threshold*, which controls the maximum difference of cells in a segment, and the minimum number of cells in a segment, smaller segments are merged. Both parameters have a fundamental role in determining the outcome of the segments creation and therefore of the map classification. The minimum number of cells in a segment can be set by determining the

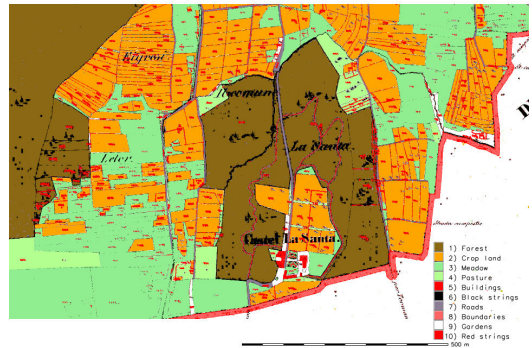


Figure 4. Classified map for contextual supervised MLC - particular.

smallest object that must be classified on the map, while the similarity threshold can be chosen empirically, trying different values and evaluating the result.

GRASS implements another approach in the *i.segment.uspo* add-on module (Grippa et al., 2017), which performs an unsupervised segmentation parameter optimization (USPO), indicating the best segmentation as the one maximizing the *F* function described in (Johnson et al., 2015). Tests for the cadaster map indicate optimum values for the threshold of 0.05 and a minimum number of cells in a segment of 10, corresponding to a value  $F = 1.827$ . However, an analysis of the size of the smallest objects on the map has suggested to increase the minimum number of cells to 25, while the threshold has been raised to 0.1 to avoid the creation of a large number of spurious small segments mostly due to artifacts imputable to JPEG compression. Higher threshold values would hinder the correct classification of small letters and symbols.

A particular of the segmentation of the test map for a threshold value of 0.1 and a minimum number of cells in a segment of 25 is shown in figure 5.



Figure 5. Segments for threshold value of 0.1 and minimum number of cells in a segment of 25 - particular.

**2.4.2 Classification** Segments have been classified into the 12 LULC classes (table 2) on the map using the GRASS

*v.class.mlr* add-on module. With respect to the MLC tests, 2 additional classes ("Black lines" and "Red lines") have been used (table 2), to account for the fact that OBIA relies also on the shape of the objects (segments) for the classification: letters and symbols compactness is very different from compactness of lines.

Class	Description
1	Forest
2	Crop land
3	Meadows
4	Pasture
5	Buildings
6	Black strings and symbols
7	Roads
8	Boundary lines
9	Gardens and courts
10	Red strings and symbols
11	Black lines
12	Red lines

Table 2. LULC classes used for OBIA classification.

This module performs the classification using the direct connection between GRASS and R, exploiting the capabilities of R's Caret library. All the available classifiers, Support Vector Machine (with a radial kernel) (SVMradial), Random Forest (RF), Recursive partitioning (Rpart), and k-Nearest Neighbors (kNN) have been used. All the available voting systems, Simple Majority Vote (SMV), Simple Weighted Vote (SWV), Best Worst Weighted Vote (BWWV) and Quadratic Best Worst Weighted Vote (QBWWV) have also been used.

A training map containing 331 sample points (figure 6) and 12 LULC classes (table 2) has been created for training. LULC classes are assigned to the segments they fall into after the segmentation by joining the tables of the two vector maps, using the segment id as join field.

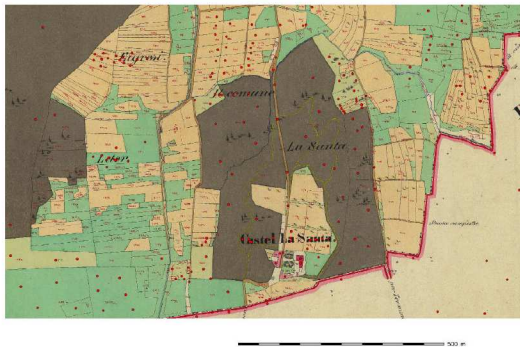


Figure 6. Training points (red dots) - particular.

The output of the classification using the SWV voting system is shown in figure 7.

## 2.5 Text and symbols removal

One of the main obstacles of the application of automatic classification techniques to maps is the presence of text string and symbols, which must be deleted in the classified map, where the user is interested in LULC.

For this reason a new GRASS add-on module, *r.replace*, has

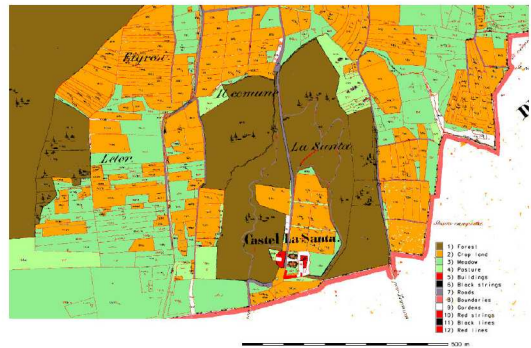


Figure 7. Classified map for OBIA and BWWV - particular.

been written, which substitutes the values of pixels of a given category with values of the surrounding pixels and stores the output in a new raster map. Pixels with other categories are left untouched.

The module iteratively applies a low-pass filter to the pixels with the category to remove, assigning the mode of the category of the neighbor pixels. The user controls the size of the neighborhood: large values can speed the process considerably but can also lead to unwanted effects where pixels with different categories are mixed, small values of the neighborhood size require a larger number of iterations and longer processing times, but provide better results when categories are mixed. The module can also create a MPEG animation of the iterative process.

The module is currently being tested and will be made available as a GRASS add-on module under the GNU General Public License ( $\geq v2$ ) used for GRASS source code.

The application of this module to the classified images using unsupervised MLC, supervised MLC, supervised contextual MLC and OBIA with the BWWV voter is shown in figures 8, 9, 10 and 11.

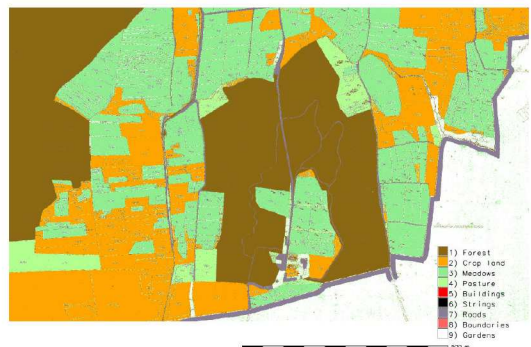


Figure 8. Classified map for unsupervised MLC after text and symbols removal - particular.

The method is effective, as long as segments containing a text string or a symbol are correctly classified. Tests are under way

### 3. RESULTS

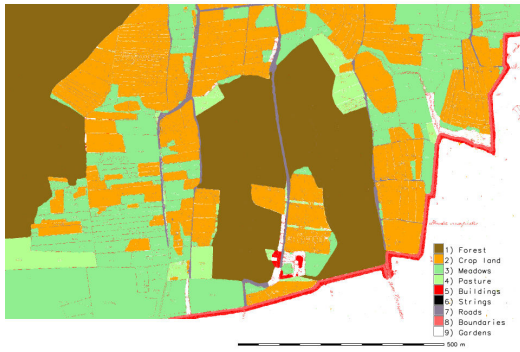


Figure 9. Classified map for supervised MLC after text and symbols removal - particular.

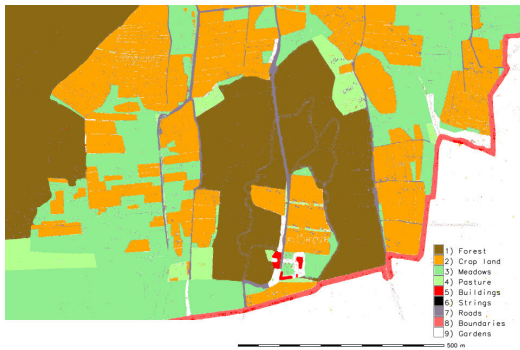


Figure 10. Classified map for contextual supervised MLC after text and symbols removal - particular.



Figure 11. Classified map for OBIA and BWWV after text and symbols removal - particular.

to determine the best value for the neighborhood size for the filter depending on the image.

To check the results of the classification, 96 areas have been manually digitized on the map. Eight areas have been created for each of the 12 categories: 4 areas of the first 10 categories have been used for training the supervised and supervised contextual MLC, the other 4 have been used to evaluate the overall accuracy and Cohen's Kappa parameter. Areas of the last 2 categories have been used to assess the overall accuracy and Cohen's Kappa parameter for the OBIA output maps. Results for the 3 MLC classifiers and 4 voting systems for OBIA are reported in tab 3.

Classifier		OA	Kappa
MLC	Unsupervised	67.9%	0.473
	Supervised	97.6%	0.960
	Supervised Contextual	98.7%	0.978
OBIA	SMV	98.7%	0.978
	SWV	99.0%	0.983
	BWWV	99.0%	0.984
	QBWWV	99.0%	0.984

Table 3. Overall accuracy and Coen's Kappa for different classifiers. OA: Overall accuracy. MLC: Maximum Likelihood Classification. OBIA: Object-based Image Analysis. SMV: Simple Majority Vote. SWV: Simple Weighted Vote. BWWV: Best Worst Weighted Vote. QBWWV: Quadratic Best Worst Weighted Vote.

Table 3 shows how OBIA using BWWV and QBWWV voters achieves the highest values of overall accuracy and Cohen's Kappa. LULC maps are also more uniform, while all maps from MLC contain fragmented areas. Note that the last two voting systems, Best Worst Weighted Vote (BWWV) and Quadratic Best Worst Weighted Vote (QBWWV), assign exactly the same votes, therefore the resulting classified map, the overall accuracy and Kappa are the same.

The overall accuracy and Cohen's Kappa parameter have been evaluated again after strings and lines removal (table 4), having removed classes 6 (Black strings and symbols), 10 (Red strings and symbols), 11 (Black lines) and 12 (Red lines).

Classifier		OA	Kappa
MLC	Unsupervised	68.0%	0.471
	Supervised	98.0%	0.968
	Supervised Contextual	99.5%	0.991
OBIA	SMV	98.7%	0.978
	SWV	99.0%	0.983
	BWWV	99.0%	0.984
	QBWWV	99.0%	0.984

Table 4. Overall accuracy and Coen's Kappa for different classifiers after strings and lines removal. OA: Overall accuracy. MLC: Maximum Likelihood Classification. OBIA: Object-based Image Analysis. SMV: Simple Majority Vote. SWV: Simple Weighted Vote. BWWV: Best Worst Weighted Vote. QBWWV: Quadratic Best Worst Weighted Vote.

Values for maps from OBIA classifiers do not change, because no segment in the remaining 9 classes were (mis)classified as one of the 4 removed classes: 6 (Black strings and symbols), 10 (Red strings and symbols), 11 (Black lines) and 12 (Red lines). As expected, Kappa and OA improve slightly for MLC, mainly because of the removal of pixels of classes 6 (Black strings and symbols) and 10 (Red strings and symbols) in the test areas.

Confusion matrixes for Supervised Contextual MLC and OBIA Best Worst Weighted Vote (BWWV) classifiers are reported in tables 5 and 6.

Tables 5 and 6 show that both Supervised Contextual MLC and OBIA are able to correctly classify “Forst”, “Meadows”, “Pasture” and “Boundaries” classes, with a slightly better result for Supervised Contextual MLC for the “Pasture” class (99.99% vs 98.44% accuracy) and for OBIA for the “Boundaries” class (100% vs 99.76% accuracy). Supervised Contextual MLC outperforms OBIA for the “Crop land” class, 99.98% vs 90.3% accuracy: this is mostly due to one single parcel of the “Gardens” class (figure 12), which has been misclassified because it has been made part of a larger segment correctly classified as “Crop land” (figure 13).

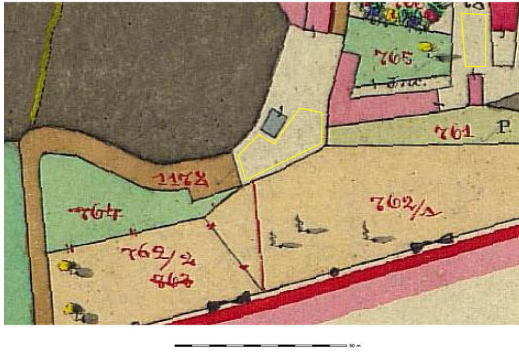


Figure 12. Garden parcel, in the center in light grey, containing a darker grey rectangle, which is misclassified. The extension of one of the check area for the “Gardens” class is shown in yellow.



Figure 13. Classified segments for BWWV voter, the garden parcel of figure 12 is misclassified as “Crop land” because it has been merged in a single segment with crop land parcels. The extension of one of the check area for the “Gardens” class is shown in yellow.

OBIA performance clearly surpass Supervised Contextual MLC for the “Buildings” class, 98.43% vs 55.89% accuracy, where the color on the map, red, is the same as that used for

“Boundaries”: OBIA classifiers are able to discriminate using the geometric features of the segments, whereas Supervised Contextual MLC uses only pixel values and neighborhood classes. A similar situation occurs for the “Road” class, where taking into account geometric features allows OBIA to reach a 98.56% accuracy, while similar colors lead to the misclassification of 6.71% of the “Road” surface as “Forest” for the Supervised Contextual MLC, limiting its accuracy to 92.29%.

OBIA classifiers incorrectly identify 7.16% of the area of the “Boundaries” class as belonging to the “Gardens” class: some segments are misclassified as “Gardens” because the JPEG compression adds a line of lighter colored pixels, whose color is very similar to color for pixels the “Gardens” class, between the two tones of red of the boundary line (figure 14).

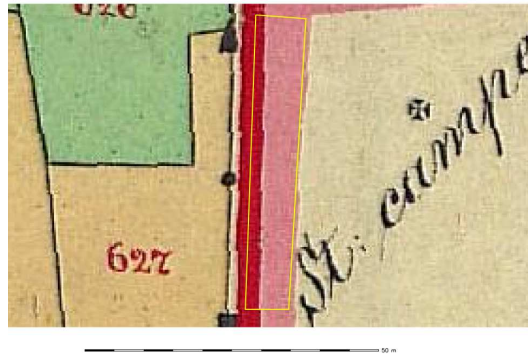


Figure 14. Light tones between two red tones of the boundary line. The extension of one of the check area for the “Boundaries” class is shown in yellow.

Many misclassifications are due to artifacts introduced by the JPEG compression, the access to the original uncompressed images is currently under negotiation. The application of the procedure to the original images should enhance the result, but tests must be carried out to quantify the benefit of using images encoded in a loseless format.

#### 4. CONCLUSIONS

The OBIA approach has proved effective for the semi-automatic digitalization and classification of the Historical Cadaster Map for the Province of Trento.

Special care must be used when choosing the training points, in particular all LULC must be sampled. For the correct classification of text and symbols, it is fundamental to place some training points in the space inside the characters, such as the inner space for the “o” letter, and the symbol drawings, because the classifier must be trained to recognize segments in these spaces as not belonging to the characters and symbols class. At the same time, the similarity threshold and the minimum number of cells in a segment must be carefully chosen to balance the need of properly classifying text with the aim of avoiding the creation of a large number of small segments which can undermine the classification step and lead to fragmented maps.

		Reference							
Classes		Forest	Crop land	Meadows	Pasture	Buildings	Roads	Boundaries	Gardens
Supervised Contextual MLC	Forest	100	0	0	0	0	6.71	0	0.06
	Crop land	0	99.98	0	0	0	0.87	0	0
	Meadows	0	0	100	0	0	0	0	0
	Pasture	0	0	0	99.99	0	0	0	0.10
	Buildings	0	0	0	0	55.89	0	0.24	0
	Roads	0	0.02	0	0	0	92.29	0	0.02
	Boundaries	0	0	0	0	44.11	0	99.76	0
	Gardens	0	0	0	0.01	0	0.12	0	99.83

Table 5. Confusion matrix for Supervised Contextual MLC. Values are percentages of the reference test set (column-based normalization), producer accuracy can be read on diagonal values.

		Reference							
Classes		Forest	Crop land	Meadows	Pasture	Buildings	Roads	Boundaries	Gardens
OBIA BWWV	Forest	99.97	0	0	0	0	1.09	0	0
	Crop land	0	90.63	0	1.56	0	0	0	2.43
	Meadows	0.02	0	100	0	0	0	0	0
	Pasture	0	0	0	98.44	0	0	0	1.38
	Buildings	0	0	0	0	98.43	0.35	0	0
	Roads	0.01	0	0	0	0	98.56	0	0
	Boundaries	0	1.05	0	0	1.57	0	100	7.16
	Gardens	0	8.32	0	0	0	0	0	89.02

Table 6. Confusion matrix for OBIA Best Worst Weighted Vote (BWWV). Values are percentages of the reference test set (column-based normalization), producer accuracy can be read on diagonal values.

The procedure for text strings and symbols removal is operative and very simple to apply, thanks to the creation of a dedicated GRASS add-on module, which will be made available after the current testing phase.

One of the advantages of OBIA is the possibility of creating a training map of points, instead of areas as for the supervised MLC, simplifying the process. Moreover, the training map created for a map sheet or image can be usually applied to other map sheets or images of the same dataset, with a clear advantage, especially on large datasets. In the case of the cadaster historical map for the Province of Trento this is a huge asset, as the dataset contains 13297 map sheets and its classification can be performed in a batch process.

The use of GRASS GIS (GRASS Development Team, 2019) provides a clear advantage for the possibility of rapid development of new procedures and modules (Preatoni et al., 2012), its direct connection with R, and the ease of the dissemination of the results, in particular for educational purposes (Ciolli et al., 2017).

The digitalization of the Historical Cadaster Map of the Trentino region is part of a wider project (Gobbi et al., 2019) which aims to build a dataset containing a large number of digital maps, created from historical maps and historical aerial images (Gobbi et al., 2018), covering a long time span for the Trentino region, as a base for the study of the landscape and ecologic evolution of the area. In particular, the application of this technique to the 1936 Italian Kingdom Forest Map (Ferretti et al., 2018) is planned in the next future.

#### ACKNOWLEDGMENTS

The cadaster historical map for the Province of Trento has been made available by the Servizio Catasto della Provincia Autonoma di Trento, under the Creative Commons Attribution 4.0 license.

#### REFERENCES







- Agnoletti, M., 2007. The degradation of traditional landscape in a mountain area of Tuscany during the 19th and 20th centuries: Implications for biodiversity and sustainable management. *Forest Ecology and Management*, 249, 5 - 17. doi.org/10.1016/j.foreco.2007.05.032.
- Antrop, M., 2005. Why landscapes of the past are important for the future. *Landscape and Urban Planning*, 70, 21-34. doi.org/10.1016/j.landurbplan.2003.10.002.
- Buffoni, D., Girardi, S., Revolti, R., Cortese, G., Mastronunzio, M., 2016. Historicalkat. la documentazione catastale trentina d'impianto è open data. *Proceedings of the XX Conferenza Nazionale ASITA*, 112–117.
- Cantiani, M.G., Geitner, C., Haida, C., Maino, F., Tattoni, C., Vettorato, D., Ciolli, M., 2016. Balancing economic development and environmental conservation for a new governance of Alpine areas. *Sustainability (Switzerland)*, 8. doi.org/10.3390/su8080802.
- Ciolli, M., Federici, B., Ferrando, I., Marzocchi, R., Sguerso, D., Tattoni, C., Vitti, A., Zatelli, P., 2017. FOSS tools and applications for education in geospatial sciences. *ISPRS International Journal of Geo-Information*, 6. doi.org/10.3390/ijgi6070225.
- Ciolli, M., Tattoni, C., Ferretti, F., 2012. Understanding forest changes to support planning: A fine-scale Markov chain approach. *Developments in Environmental Modelling*, 25, 355-373. doi.org/10.1016/B978-0-444-59396-2.00021-3.
- Dittrich, A., von Wehrden, H., Abson, D. J., Bartkowski, B., Cord, A. F., Fust, P., Hoyer, C., Kambach, S., Meyer, M. A., Radzevičiūtė, R., Nieto-Romero, M., Seppelt, R., Beckmann, M., 2017. Mapping and analysing historical indicators of ecosystem services in Germany. *Ecological Indicators*, 75, 101 - 110. doi.org/10.1016/j.ecolind.2016.12.010.

- Ferretti, F., Sboarina, C., Tattoni, C., Vitti, A., Zatelli, P., Geri, F., Pompei, E., Ciolli, M., 2018. The 1936 Italian Kingdom Forest Map reviewed: A dataset for landscape and ecological research. *Annals of Silvicultural Research*, 42, 3-19. doi.org/10.12899/asr-1411.
- Forejt, M., Dolejš, M., Raška, P., 2018. How reliable is my historical land-use reconstruction? Assessing uncertainties in old cadastral maps. *Ecological Indicators*, 94, 237 - 245. doi.org/10.1016/j.ecolind.2018.06.053.
- Gobbi, S., Cantiani, M.G., Rocchini, D., Zatelli, P., Tattoni, C., Ciolli, M., La Porta, N., 2019. Fine spatial scale modelling of Trentino past forest landscape (trentinoland): a case study of foss application. In preparation.
- Gobbi, S., Maimeri, G., Tattoni, C., Cantiani, M.G., Rocchini, D., La Porta, N., Ciolli, M., Zatelli, P., 2018. Orthorectification of a large dataset of historical aerial images: Procedure and precision assessment in an open source environment. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42 4W8, 53–59.
- GRASS Development Team, 2019. Geographic Resources Analysis Support System (GRASS GIS) Web Site. <http://grass.osgeo.org>. Accessed: 2019-04-29.
- Grippa, T., Lennert, M., Beaumont, B., Vanhuysse, S., Stephenne, N., Wolff, E., 2017. An open-source semi-automated processing chain for urban object-based classification. *Remote Sensing*, 9. doi.org/10.3390/rs9040358.
- Haase, D., Walz, U., Neubert, M., Rosenberg, M., 2007. Changes to Central European landscapes-Analysing historical maps to approach current environmental issues, examples from Saxony, Central Germany. *Land Use Policy*, 24, 248-263. doi.org/10.1016/j.landusepol.2005.09.003.
- Johnson, B.A., Bragais, M., Endo, I., Magcale-Macandog, D.B., Macandog, P.B.M., 2015. Image segmentation parameter optimization considering within-and between-segment heterogeneity at multiple scale levels: Test case for mapping residential areas using landsat imagery. *ISPRS International Journal of Geo-Information*, 4, 2292-2305. doi.org/10.3390/ijgi4042292.
- OPENdata Trentino, 2019. Historical Cadaster Map of the Trentino region metadata. <https://dati.trentino.it/dataset/mappe-storiche-d-impianto>. Accessed: 2019-04-29.
- Preatoni, D.G., Tattoni, C., Bisi, F., Masseroni, E., D'Acunto, D., Lunardi, S., Grimod, I., Martinoli, A., Tosi, G., 2012. Open source evaluation of kilometric indexes of abundance. *Ecological Informatics*, 7, 35-40. doi.org/10.1016/j.ecoinf.2011.07.002.
- R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Revolti, R., 2017. Produzione e restauro della cartografia catastale in Trentino. <https://drive.google.com/file/d/0B0abPZ1lrxqdcnUwXzB1T2RvV0E/view>. Accessed: 2019-04-29.
- Servizio Catasto della Provincia Autonoma di Trento, 2019a. HistoricalKat, documentation of the Trentino cadastral archive. <http://historicalkat.provincia.tn.it>. Accessed: 2019-04-29.
- Servizio Catasto della Provincia Autonoma di Trento, 2019b. Mappe storiche di impianto (urmappe). <https://www.catastotn.it/mappeStoriche.html>. Accessed: 2019-04-29.
- Skaloš, J., Engstová, B., 2010. Methodology for mapping non-forest wood elements using historic cadastral maps and aerial photographs as a basis for management. *Journal of Environmental Management*, 91, 831 - 843. doi.org/10.1016/j.jenvman.2009.10.013.
- Statuto, D., Cillis, G., Picuno, P., 2017. Using Historical Maps within a GIS to Analyze Two Centuries of Rural Landscape Changes in Southern Italy. *Land*, 6.
- Tattoni, C., Ciolli, M., Ferretti, F., Cantiani, M.G., 2010. Monitoring spatial and temporal pattern of Paneveggio forest (Northern Italy) from 1859 to 2006. *IForest*, 3, 72-80. doi.org/10.3832/ifor0530-003.
- Tattoni, C., Ianni, E., Geneletti, D., Zatelli, P., Ciolli, M., 2017. Landscape changes, traditional ecological knowledge and future scenarios in the Alps: A holistic ecological approach. *Science of the Total Environment*, 579, 27-36. doi.org/10.1016/j.scitotenv.2016.11.075.



Article

# New Tools for the Classification and Filtering of Historical Maps

Stefano Gobbi <sup>1,2,3</sup> , Marco Ciolli <sup>1,4</sup> , Nicola La Porta <sup>2,3</sup> , Duccio Rocchini <sup>4,5,6,7</sup> ,  
Clara Tattoni <sup>1</sup>  and Paolo Zatelli <sup>1,\*</sup> 

- <sup>1</sup> Department of Civil, Environmental and Mechanical Engineering, University of Trento, Via Mesiano 77, 38123 Trento, Italy; stefano.gobbi@unitn.it (S.G.); marco.ciolli@unitn.it (M.C.); clara.tattoni@unitn.it (C.T.)
  - <sup>2</sup> Fondazione Edmund Mach, Research and Innovation Centre, Centro di Ricerca e Innovazione, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy; nicola.laporta@fmach.it
  - <sup>3</sup> The EFI Project Centre on Mountain Forests (MOUNTFOR), Edmund Mach Foundation, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy
  - <sup>4</sup> C3A, Center Agriculture Food Environment, University of Trento, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy; duccio.rocchini@unitn.it
  - <sup>5</sup> Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, Via Sommarive, 9, 38123 Povo (TN), Italy
  - <sup>6</sup> Fondazione Edmund Mach, Research and Innovation Centre, Department of Biodiversity and Molecular Ecology, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy
  - <sup>7</sup> Department of Applied Geoinformatics and Spatial Planning, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, 16500 Praha - Suchdol, Czech Republic
- \* Correspondence: paolo.zatelli@unitn.it

Received: 7 August 2019; Accepted: 10 October 2019; Published: 14 October 2019



**Abstract:** Historical maps constitute an essential information for investigating the ecological and landscape features of a region over time. The integration of heritage maps in GIS models requires their digitalization and classification. This paper presents a semi-automatic procedure for the digitalization of heritage maps and the successive filtering of undesirable features such as text, symbols and boundary lines. The digitalization step is carried out using Object-based Image Analysis (OBIA) in GRASS GIS and R, combining image segmentation and machine-learning classification. The filtering step is performed by two GRASS GIS modules developed during this study and made available as GRASS GIS add-ons. The first module evaluates the size of the filter window needed for the removal of text, symbols and lines; the second module replaces the values of pixels of the category to be removed with values of the surrounding pixels. The procedure has been tested on three maps with different characteristics, the “Historical Cadaster Map for the Province of Trento” (1859), the “Italian Kingdom Forest Map” (1926) and the “Map of the potential limit of the forest in Trentino” (1992), with an average classification accuracy of 97%. These results improve the performance of classification of heritage maps compared to more classical methods, making the proposed procedure that can be applied to heterogeneous sets of maps, a viable approach.

**Keywords:** heritage maps; image classification; OBIA; map filtering; GRASS GIS

## 1. Introduction

Historical maps are available in many nations and regions of the world, and represent an invaluable source of information in many fields [1,2].

Depending on the purpose for which the maps were conceived, historical maps are typically reliable in terms of precision and spatial accuracy, thus they can usually be compared to modern maps not only to make qualitative assessments but also to quantify the changes that have occurred in the

past [3,4]. For example, in ecological studies, it is widely recognized that past land use affects present biodiversity distribution and status [5–8], ecosystem service provision [9] and traditional ecological knowledge [10]. Thus, the present structure and the ecological dynamics of the landscape in an area can be understood only having a profound knowledge of the land use changes previously occurred in the region [4]. In addition, the interpretation of many of the results of contemporary ecological analyses, including those obtained with field sampling, is more robust if considered in a framework that includes past landscapes, underlining the importance of historical maps [11]. The same concepts can be applied to many other disciplines or study areas like urban studies [12], geomorphology [13], planning [14], social science and economic studies [15], anticipation [16] and history [17].

The use of GIS to spatially reconstruct past landscape changes during long periods by comparing historical geographical data has been sharply increasing in the last few decades and different approaches can be found in the literature dealing with different map scales and time extents [10,18,19]. The vast majority of this research relies on satellite imagery, aerial photographs and historical maps, the latter being often the oldest documents available. Landsat images can date back to the 1970s, aerial photographs at the national scale are sometimes available from the 1940s, although occasionally single aerial photographs taken in earlier periods can be found. However, historical maps are the only material that cover pre “remote sensing era” and can provide information from several past decades or centuries back [3].

Before their use in a GIS, historical maps must be digitized, georeferenced, classified and possibly post processed to filter unwanted features [20]. The digitalization of historical maps is usually carried out manually by one or more operators who basically redraw the boundaries of the map and assign a value to the features: this is time-consuming and prone to error tasks [3,4]. A rather comprehensive list of challenges in the segmentation and classification of printed map can be found in [20,21]. Most of the applications of these approaches have been developed for a specific set of maps or to extract a specific land cover class [22]. This work aims to a more general extraction of land use/land cover (LULC) classes from maps with very different features. Together with thematic maps created from orthoimages [23] will form a comprehensive dataset (Trentinoland) for the study of the forest landscape evolution in the last few centuries [24] in the Italian region of Trentino. Several techniques are available for text extraction from printed documents [25], but they focus on the development of effective document processing methods rather than map categories’ preservation.

This paper presents new map classification and filtering procedures based on Object-based Image Analysis (OBIA) that have been implemented using GRASS GIS [26], developing new techniques and software modules for text and symbol removal. GRASS GIS is the leading GIS Free and Open Source Software (FOSS), used in research and education [27]. Its availability under the GNU GPL license allows for the modification and adaptation of its code to a wide range of GIS applications; see, e.g., [28–30].

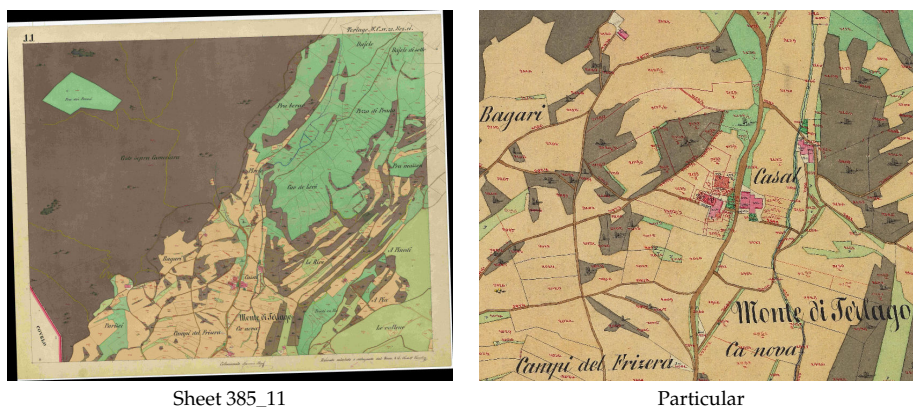
The paper is organized as follows: Section 2 describes the maps used in the tests, Section 3 outlines the processing steps for the map digitalization and the state of the art, Sections 4 and 5 illustrate the segmentation and classification procedures, Section 6 describes the text removal method and its testing, Section 7 reports the results of the tests and the conclusions and future developments are presented in Section 8.

## 2. Test Maps

Three different maps have been used to test the effectiveness of the approach to historical map digitalization proposed in this paper. These maps have been chosen because they represent typical situations encountered when dealing with historical maps: the “Historical Cadaster Map for the Province of Trento” is a 19th century hand-drawn map, the 1936 “Italian Kingdom Forest Map” is a hand-drawn map on a printed base, and the 1992 “Map of the potential limit of the forest in Trentino” is a printed map.

The Historical Cadaster Map for the Province of Trento (Figure 1) is based on the Second Military Survey of the Habsburg Empire [31], continuing the mapping effort started with the Royal Imperial Patents of the 23rd of December of 1817 by Francis I, the first Emperor of Austria [32,33]. The survey for the Trentino (South Tyrol) region has been carried out between 1855 and 1861 [33,34], using a projection system centered at the Pfarrturm in Innsbruck. While its main purpose is mapping property boundaries, the map contains not only the geometric description of the land units but also a representation of natural boundaries and LULC, with specific distinction between agricultural uses. The map consists of 13,297 map sheets approximately  $52.68 \times 65.85$  cm wide, each covering 288 ha. They have been digitized at 230 DPI, 24 bit color depth, with file sizes between 5 and 8 MB, georeferenced in the ETRS89/UTM 32 N (EPSG 25832) datum and made available as a set of JPEG images [35], with JGW word files, under the Creative Commons Attribution 4.0 license [36]. The ground resolution of the digital map is around 0.32 m.

Map geo-referentiation has been performed by the Cadaster Service (Servizo Catasto) of the Provincia Autonoma di Trento using the four corners of each sheet as Ground Control Points (GCPs), whose ETRS89/UTM 32 N coordinates have been previously determined. This process has achieved an accuracy of 5–10 m, but it is possible to improve it locally, up to 1–2 m [37], by applying a successive transformation using local GCPs. Supplementary information about the map and the underlying survey, such as surveying maps, cadastral borders descriptions, pictorial and geometric drawings, cadastral monographs and triangulation points can be found on the HistoricalKat web site [34].



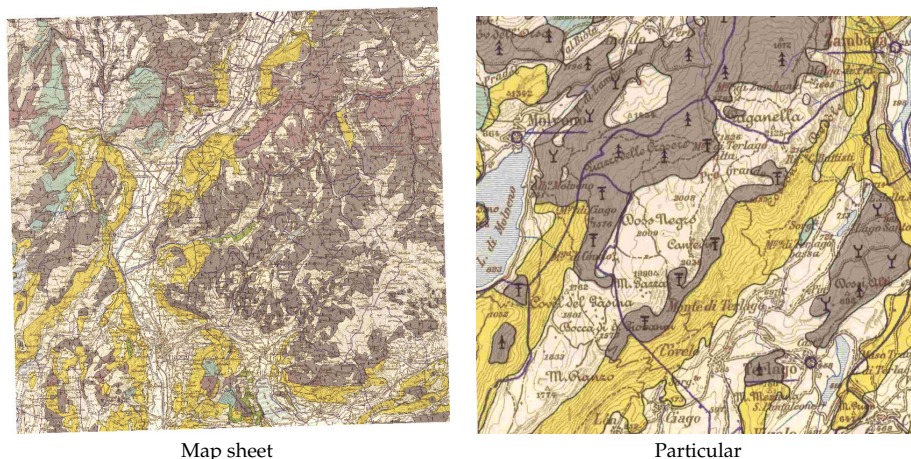
**Figure 1.** Map sheet 385\_11 (“Comune catastale di Terlago”) of the Historical Cadaster Map of the Province of Trento.

The 1936 Italian Kingdom Forest Map (IKFM, “Carta Forestale della Milizia Forestale del Regno d’Italia del 1936”) [38] (Figure 2) represents the forest extent and main species composition in the whole Italian territory and parts of some surrounding countries in 1936 [3]. It is composed of 276 sheets hand-drawn on a printed 1:100,000 scale map. The base map is the Italian official map of the time, developed by the Istituto Geografico Militare Italiano (IGMI, Italian Military Geographic Institute). Ground surveys and field sampling have originally been carried out on 1:25,000 maps and later transferred to the 1:100,000 maps. The original sheets have been scanned with a resolution of 400 DPI, resulting in  $6650 \times 5900$  pixels images saved as TIFF files, approximately 112 MB each. The map has been geo-referenced in the Italian national Rome 40 datum with Gauss-Boaga projection, in the east or west zone (EPSG 3003 and 3004) depending on the location of the map sheet. The accuracy of the coordinates has been estimated to be around 30 m, ground pixels size is around 6.5 m [3].

The IKFM categories describe species composition and silvicultural system with three macro categories (conifers, broadleaves and “degraded forests”) and eight different categories, with some

sub-categories related to forest system features for broadleaves. The main categories are Resinose (Conifers), Faggio (Beech), Rovere e Farnia (Sessile oak and English oak), Cerro (Turkey oak), Sughera (Cork oak), Castagno (Chestnut), Altre specie o misti (other species or mixed wood), and Boschi degradati (Degraded forest). Details about forest classification in the IKFM can be found in [3].

Forest areas have been manually digitized in a previous project, creating approximately 63,000 vector polygons. The map, both in raster and vector format, is available on a dedicated web site [3], with a WebGIS interface, under the Creative commons Attribution—Version 3.0 license.

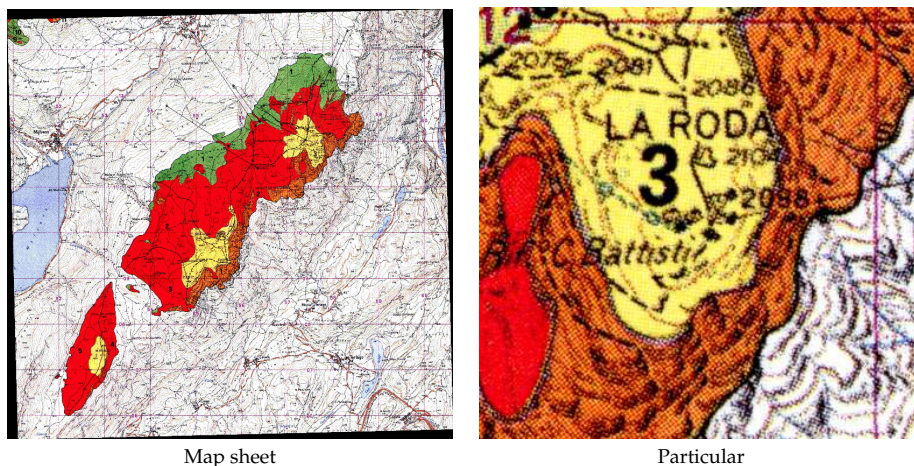


**Figure 2.** Map sheet of the 1936 Italian Kingdom Forest Map.

The “Map of the potential limit of the forest in Trentino” (MPLFT, “Carta del limite potenziale del bosco in Trentino”) [39] (Figure 3) is a map published in 1992 by the *Forest, Hunting and Fishing Service* (Servizio Foreste Caccia e Pesca) of the local government (Provincia Autonoma di Trento) in Trentino. It is based on ground surveys carried out between 1987 and 1990 and represents forest areas as polygons. Colored polygons have been printed with transparency on the 1:50,000 IGMI color map.

The original map is divided into small  $23 \times 23$  cm sheets printed in colors using halftones. The map sheets have been digitized with 600 DPI resolution and saved as TIFF files, approximately 60 MB each. The images have been geo-referenced using the coordinates of the kilometric grid markings as GCPs, in their original ED50/UTM32N datum (EPSG 23032). Ground pixel size is 2.12 m. The geo-referentiation process estimates an accuracy of the coordinates of 0.6 m, higher than the limit due to the map scale. The map contains four classes: forest area (green), real potential [forest] area (red), excluded potential [forest] area (orange) and area beyond the upper [forest] limit (yellow).

Specific classification methods, based on the application of an inverse halftoning scheme and a classifier, have been developed for maps containing halftones [40], but tests described in this paper show that, with a proper choice of segmentation parameters and training points, it is possible to achieve satisfying results using OBIA.



**Figure 3.** Map sheet of the “Map of the potential limit of the forest in Trentino”.

One test sheet has been selected for each map, covering roughly the same area around the village of Terlago, a few kilometers east of Trento, Italy. This area has been chosen because most of the categories are present on the three maps, with a complex background which includes lakes and villages.

### 3. Historical Map Digitalization

Map digitalization aims to create a vector or raster map reproducing geographic features represented on a map image. The transformation from image to map allows for the application of GIS methods to extract, combine and process the information provided by the map.

Unless the map image is already available in digital form, the first step in this process is the scanning of the map, where parameters such as geometric (spatial sampling rate on the paper map) and radiometric (bit depth) resolution, contrast, and brightness, influence the resulting digital image properties. The scanning process can also introduce geometric distortions. The second step consists of the map georeferencing in the datum used for the processing. Different approaches are possible [41,42], and the resulting geometric distortions must be evaluated to assess the spatial accuracy of successive processing [43]. Maps used in this work are already available as georeferenced images, therefore these processing steps are not discussed here. The maps have been digitized using different parameters (datums, scan resolution, etc.) because the process has been carried out by different organizations for different purposes. This fact allows for testing the classification and filtering procedure on maps with different features.

Finally, geographic features must be identified and classified. Different techniques have been developed: histogram thresholding, color space clustering, edge detection, region-based approaches, artificial neural networks and semantic segmentation [20,44]. Aiming to extract LULC classes from maps with very different features, the OBIA technique has been used for the classification, with the segmentation of the maps using a region-based approach and successive segment classification using machine learning. This approach differs from histogram thresholding and color space clustering because, in these methods, spatial contiguity is not explicitly taken into account, from edge detection techniques for the fact that they require an additional classification step and are usually performed for gray-scale images, while neural network-based and semantic segmentation approaches require extended training time and must be specialized for each specific type of map.

The flowchart of the processing chain of the proposed procedure for the classification and filtering of historical maps is comprised of the following steps (Figure 4): optimal segmentation parameters determination, segmentation, training set creation, classification, filtering and performance evaluation.

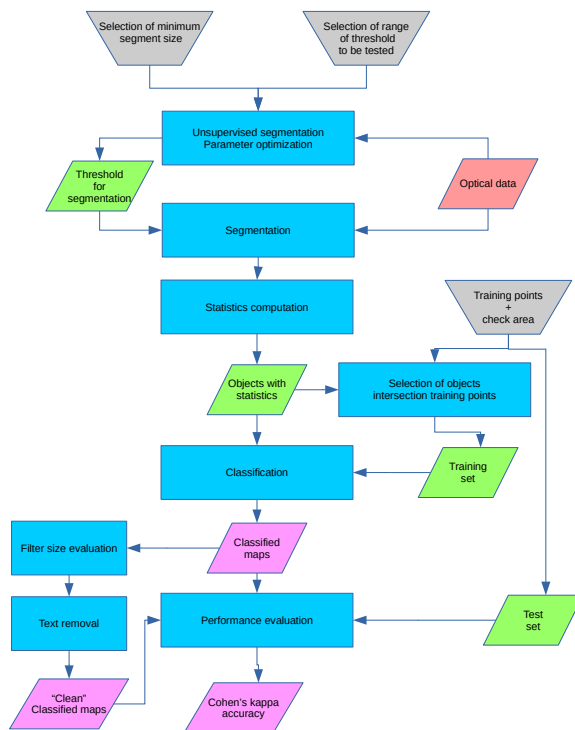


Figure 4. Flowchart of the classification and filtering procedure.

#### 4. Segmentation

Map digitalization can be achieved with different approaches to perform an LULC classification of a digital map:

- Pixelwise, an algorithm classifies every single pixel of the image in an LULC category, using its spectral signature;
- Object oriented, an algorithm creates “objects”, or cluster of adjacent pixels with a similar spectral response, and then classifies the objects using their geometric and spectral features.

Both of these approaches have been extensively discussed in the literature, proving their effectiveness and viability [45]. The choice of the family of the algorithm should be driven by the resolution of the image or, in other words, what can be identified as an object on the image. In the case of high resolution aerial images (pixel sizes of centimeters or meters), a building, a road or a single tree are “objects” defined by a set of pixels part of the same structure. As the resolution gets coarser (e.g., Landsat imagery with 30 m resolution), single objects become composed by less and less pixels, until the dimension of the pixel itself becomes greater than the single object [46]. In the case of high resolution images the pixelwise and object oriented approaches are both viable, but, for lower resolution, a pixelwise approach is preferable [47].

The combinations of scan resolutions and original map scales of the historical maps studied in the current work result in objects such as buildings and roads, but also text and symbols, to be represented by a number of pixels large enough to use an object oriented approach. This choice is supported by the comparison between object oriented classification and maximum likelihood classification for a map sheet of the cadaster map reported in [48].

The first step in this approach is the creation of the objects themselves. Areas on the image corresponding to objects are called *segments* and their creation is called *segmentation*. This is achieved by using GRASS GIS *i.segment* module, which can perform the segmentation of an image using two different approaches: a region-growing algorithm, used in this study, or a mean-shift algorithm [49].

Segmentation is the first step in the processing sequence that composes OBIA classification, and its results affect all the further analysis; therefore, it must be carefully calibrated. Ideally, the result of the segmentation creates a map where each segment corresponds to a geographical object on the map. The region-growing algorithm used in this study requires two input parameters: a threshold value and a minimum size value.

The *threshold* parameter, ranging between 0 and 1, defines the value of the (normalized) similarity below which neighbor pixels are considered part of the same segment. Lowering the threshold value makes the segmentation algorithm more sensitive to radiometric changes; for a threshold value of 0, the segmented maps are identical to the original maps since each pixel is considered a distinct object [50]. A threshold value of 1 creates a single segment for the whole image, since all pixels are considered similar. The *minsize* value is the minimum area in a pixel of a single segment [50]; smaller segments are merged. It must be chosen small enough to allow the recognition of small objects but not too small to avoid the division of larger objects into small segments that can cause problems during the classification step.

Usually, the goodness of a segmentation is evaluated by the user with a visual assessment [51]. The problems with this approach are its subjectivity, which makes it difficult to quantify the result quality, especially considering that a single image can contain more than a billion segments [51], and the impossibility of its use in an automatic procedure. Different objective approaches have been developed to avoid these problems, the solution described in [51] and implemented in GRASS' *i.segment.USPO* (Unsupervised Segment Parameter Optimization) module [52] has been used here. The module provides an indication of the best combination of the threshold and minimum size parameters among all their combinations chosen by the user indicating a search interval and step for each parameter. The performance of the segmentation is evaluated using two indexes, intra-object variance weighted by object size (VW) spatial autocorrelation (SA) [53]: VW measures the the homogeneity of the segment and SA measures the difference between neighbor objects in terms of spectral response. The ideal segmentation process should lead to a result where both of these indexes are at a maximum.

The algorithm used by *i.segment.USPO* normalizes VW and SA as follows [51]:

$$F = \frac{X_{max} - X}{X_{max} - X_{min}}, \quad (1)$$

where  $X$  represents either VW or SA, and  $F(X)$  is the normalized value of the analyzed index. A high value (close to 1) for normalized VW indicates undersegmentation, i.e., segments are larger than objects on the map, while a high value of normalized SA indicates the opposite effect, or oversegmentation, i.e., segments are smaller than objects on the map. The module calculates a synthetic "overall goodness" index either as a simple sum of the normalized VW and SA parameters [53] or as a function of the twos [54]:

$$F = (1 + \alpha^2) \frac{AS_{norm} \times VW_{norm}}{\alpha^2 \times AS_{norm} + VW_{norm}}. \quad (2)$$

$F$  is a normalized function varying from 0 to 1, with 0 indicating a poor quality and 1 a good quality segmentation, while  $\alpha$  is a parameter set by the user to control the relative weight of SA and VW [51]. While this index is effective in giving an unsupervised indication of the goodness of a segmentation, the authors of the *i.segment.USPO* module recommend a visual check of the segmentation results, especially in those areas where the objects represented are small, e.g., single trees in a wide open area, to check the consistency of the final segmented map [51]. The use of *i.segment.USPO* can be computationally intensive, since the module creates all the segment maps with all the possible

parameter combinations indicated by the user: for this reason, it is possible to perform the test on sub regions, selected as representative of the whole image.

For the test maps described in Section 2, the minimum segment size has been chosen to be smaller than the smallest symbol (Table 1), since the correct separation and classification of text and symbols is a fundamental requirement for their successive elimination described in Section 6. This approach inevitably leads to oversegmentation of some area, but this effect is preferable to undersegmentation [55,56] because the first situation can be corrected during the classification step, contrary to the latter. The optimal threshold values have been determined applying the *i.segment.USPO* with fixed minimum segment size to a region of each image where all the combinations of colors and shapes are present.

**Table 1.** Values of the optimized segmentation parameters for the test maps. Minsize in pixels. The optimization criteria is the sum of the normalized intra-object Variance Weighted by object size (VW) and Spatial Autocorrelation (SA) parameters described in Section 4.

Map	Threshold	Minsize	Optimization Criteria
Cadaster	0.06	20	1.283
IKFM	0.03	20	1.077
MPLFT	0.03	60	1.340

## 5. Classification

The second step of the OBIA is the classification of the segments. The supervised approach requires the creation of a training map containing a set of training areas representative of the LULC classes to be recognized. While it is possible to use any vector map containing areas, a procedure has been set up so that training points instead of areas are used: the information about the LULC class is transferred from each training point to the segment it falls into. The advantage of this approach is that manually digitizing points instead of areas is remarkably faster and the process uses the geometry of the segments created automatically. Furthermore, the process is independent from the segments configuration, which can change if different parameters' values are used in the segmentation step, for example while testing the procedure.

The creation of vector segments and the evaluation of their radiometric and geometric features is performed in GRASS by the *i.segment.stats* module [57]. Radiometric statistics include, for each image band: min, max, range, mean, mean of absolute values, standard deviation, variance, coefficient of variation, sum, sum of absolute values, first quartile, median, third quartile and the 90th percentile. For the classification of the test maps, the mean, median, variance, first quartile and third quartile have been used. Geometric attributes, calculated via the *r.object.geometry* module [58], include area, perimeter, fractal dimension (FD), compact square (CS), and compact circle (CC), are described as:

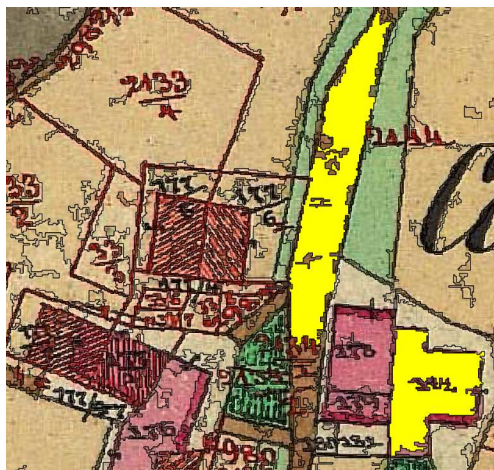
$$FD = 2 \times \frac{\log(perimeter)}{\log(area + 0.0001)} \quad (3)$$

$$CS = 4 \times \frac{\sqrt{area}}{perimeter} \quad (4)$$

$$CC = \frac{perimeter}{2 \times \sqrt{\pi \times area}} \quad (5)$$

Only the fractal dimension, square ratio and circle ratio have been used for the classification of the test maps because shape, rather than size, discriminates between segments belonging to different LULC classes in these maps (Figure 5).





**Figure 5.** Details of the segments map for the cadaster map: the two segments in yellow, corresponding to a building and a road section, have similar area and perimeter but different shape; therefore, fractal dimension, square ratio and circle ratio facilitate their correct classification.

The training map is created extracting the segments containing training points and assigning them their LULC class.

The final step for the classification is performed by the GRASS module *v.class.mlR*, which combines GRASS and the R Caret library [59]. Five different classification techniques have been used: k-nearest-neighbor classifier (k-NN), naive Bayesian classifier (NB), support vector machine (SVM), multi-layer perceptron (MLP) and decision tree (DT). Each classifier assigns (“votes”) each segment to an LULC class according to the descriptive parameters of the LULC classes provided by the training areas [60]. Discordance is possible between the classifiers; therefore, a voting method is applied [60]:

- Simple Majority Vote (SMV): all the classifiers have the same authority and the decision on the LULC class of an object is determined by a majority vote;
- Simple Weighted Vote (SWV): the weight of the vote of each classifier in the decision process is proportional to its accuracy in the identification of the class of the training areas;
- Best-Worst Weighted Vote (BWWV): as in the SWV process, all the classifiers are ranked according to their accuracy; the worst voter is excluded from the voting process;
- Quadratic Best-Worst Weighted Vote (QBWWV): as in BWWV, but the weights of the best voters are squared in order to give them more authority.

All the voting methods have been used for the classification of the test maps and the best one has been selected by ranking the classification generated by the different voters according to their overall accuracy and Cohen’ Kappa values, evaluated against a reference map using the *r.kappa* GRASS module. The classes corresponding to text, symbols and boundary lines have been removed in the final digitized maps.

## 6. Text and Symbol Removal

Text and symbol recognition techniques on images, technical drawings and maps have been researched for a long time [61], but new approaches have been proposed in recent years [21,25,62]. The general problem of text/non-text separation is relevant for a wide range of documents, such as manuscripts, images and video, in every situation in which a semantic analysis is used to index and rank them. Text/non-text separation is usually performed by the classification of features on the image

in text and non-text classes. This classification can be done at the pixel level, using only the radiometric information, or at feature level, using radiometric and morphological, shape-based, information [25]. The latter approach allows the use of additional information, such as texture, which can be derived from pixels' radiometric information.

For images and video text/non-text, separation is performed with a step-wise approach, which separates text detection and recognition phases, or an integrated approach [63]. The first method uses segmentation to extract features belonging to strings, while integrated techniques use characters recognition procedures.

Text/non-text separation for technical drawings and for maps presents some common characteristics, but for drawings the focus is on the detection of symbols, their localization and the identification of their connections [64] after image binarization, with the creation of specific tools [65].

A survey of techniques for text/non-text separation on maps can be found in [25]: methods described there are mostly based on the combination of pixelwise classification and thresholding and the application of filter rules based on a priori knowledge of the morphological features of the characters, symbols and lines to recognize. Other approaches include binarization and template matching [66,67], "text areas" detection and the application of standard optical character recognition (OCR) software [68], or the use of a tailored connected component analysis and segmentation [69]. An assessment of the possible results of the application of text detection techniques and OCR can be found in [70]. A different approach has been used for the digitalization of historical map with a focus on the separation of land use classes in [22], with the removal of text, symbols, and lines by a pre-process step with dilation, median and low-pass filtering before an unsupervised pixelwise classification using a k-means algorithm.

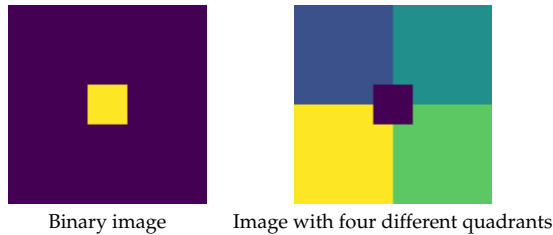
In this research, a new technique has been developed for text, symbols and lines removal from classified maps, substituting the category of their pixels with the category of the surrounding pixels. The input map must contain one or more categories corresponding to the text, symbols and lines to remove; therefore, images representing maps must be classified beforehand. If text, symbols, and lines correspond to different categories—for example, because the map contains text strings with different colors or different sets of symbols—two approaches are possible: either all the categories corresponding to text, symbols and lines to be removed are reclassified in a single category or the procedure is applied interactively to the sequence of categories to be removed. The results can be different for the two approaches because the application of the procedure to one category can influence the application to the successive ones.

The procedure applies a low pass filter [71] to pixels of the category to remove, substituting their values with those of the surrounding pixels. A moving window assigning the mode of the values to the central pixel is used to remap its value to one of the existing discrete categories. A binary map containing only the category to be removed, with all other pixels being set to NULL, is used for the identification of the pixels to which to apply the procedure. The process is iterated until no pixel of the category to remove is left.

This procedure has been implemented in a new GRASS GIS [26] module, *r.fill.category* [72]. The user controls the process by setting the neighborhood size in pixels and the maximum number of iterations. The neighborhood size sets the size of the moving window used to evaluate the mode of the values to assign to pixels of the category to be replaced. It is possible to keep the maps created at each iteration; by default, maps generated at each iteration are removed as soon as they are used to save space. The module is also able to combine the maps for each iteration as frames of an MPEG animation. The module is available as a GRASS add-on in the standard repository under the GNU General Public License (>=v2).

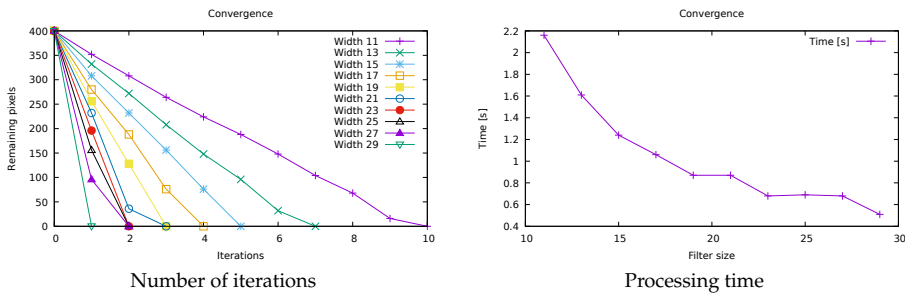
The size of the moving window used by the *r.fill.category* module must be larger than half the minimum number of pixels separating two sides of the text/symbol segment to replace; otherwise, the mode of the values in the moving window inside the area is equal to the category to remove; therefore, pixels values are left unchanged. Larger filter sizes lead to a smaller number of iterations and faster

processing time, but can lead to the substitution of the pixels values with values of pixels far from the segment. Tests have been carried out to evaluate how the filter size influences the number of iterations and the processing time needed to remove all the pixels of a given class. The *r.fill.category* module has been applied to the two images in Figure 6: both images are  $100 \times 100$  pixels with an inner  $20 \times 20$  square, but the first one has a uniform background, while the second image is divided into four different quadrants with different values.

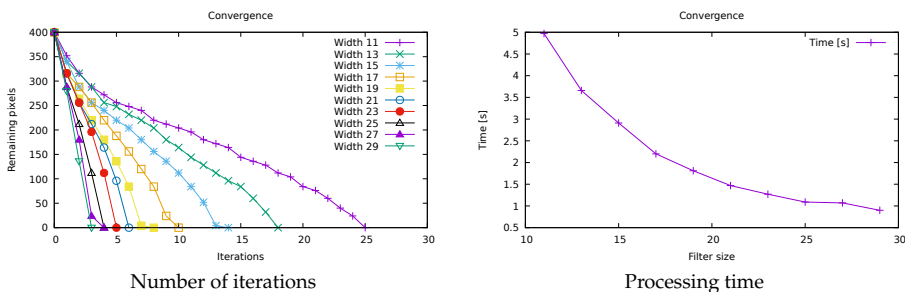


**Figure 6.** Synthetic  $100 \times 100$  images used for the tests. The image on the left has value 1 on the inner  $20 \times 20$  square and 0 outside, the image on the right has values from 1 to 4 clockwise starting from the top left quadrant and 0 in the inner square.

The number of iterations and the time needed to remove the inner square of  $20 \times 20$  pixels depending on the filter size is shown in Figure 7 for the first image and Figure 8 for the second one.



**Figure 7.** Number of iterations and processing time needed to remove a  $20 \times 20$  pixels square on a binary  $100 \times 100$  image depending on the filter size.

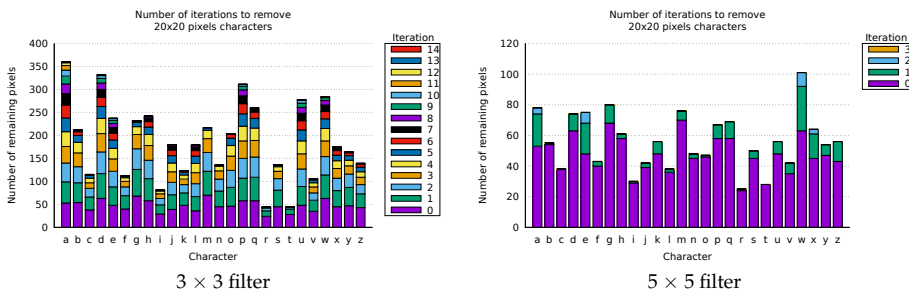


**Figure 8.** Number of iterations and processing time needed to remove a  $20 \times 20$  pixels square depending on the filter size on a  $100 \times 100$  image with background with different values for the four quadrants.

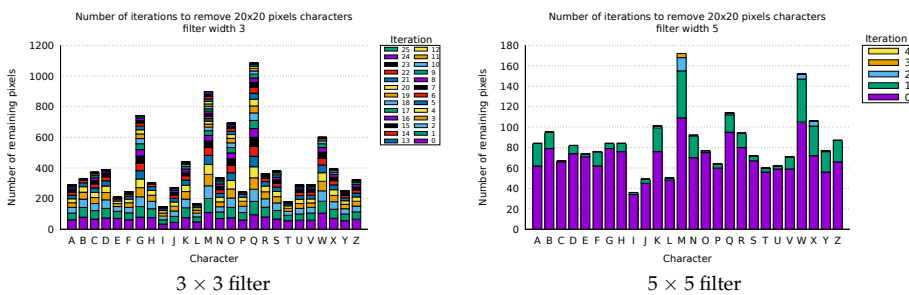
As expected, the size of the filter is fundamental in determining the number of iterations and processing time: for the binary image, the number of iterations increases ten-fold while processing time quadruples passing from a  $29 \times 29$  to a  $11 \times 11$  filter size, while, for the image divided in quadrants, both the number of iterations and the processing time increase five-fold. The complexity of the background

also plays a major role, with more than double the amount of iterations and processing time needed to remove the central square in the case of background with different categories. This is obviously due to the slower convergence in the evaluation of the mode value in the case of mixed background.

Tests have been carried out to determine how the shape of a single character influences the number of iterations needed to remove the text. Images containing single characters in the Garamond font with 10 pt size, occupying approximately  $20 \times 20$  pixels, both lower and upper-case have been created. The results of the application of  $3 \times 3$  and  $5 \times 5$  filters to lower and upper-case characters are shown in Figures 9 and 10. The variability of the number of iterations required to remove different characters of the same font and the same size is evident. For lowercase characters, the number of iterations required for their complete removal ranges from 3 for the *r* letter to 13 for the letter *a*, using a  $3 \times 3$  filter (Figure 9, left). With a wider  $5 \times 5$  filter (Figure 9, right), not only the number of required iterations drops, as expected, but the character requiring the most iterations changes as well. The *r* character remains one of the faster to remove, and the characters *a*, *e*, *w*, and *x* are still the slowest, but, for the  $5 \times 5$  filter characters, *d* and *p* are removed in two iterations like most of the other letters, while they are in the slowest group for the  $3 \times 3$  filter. For the uppercase characters, a  $3 \times 3$  filter size (Figure 10, left) requires 23 iterations for the removal of the *Q* character, whereas the *I* character needs only 7. Increasing the filter size from  $3 \times 3$  to  $5 \times 5$  not only decreases the number of iterations from 23 to 4, but it also changes the character needing the higher number of iterations: the maximum corresponds to the *M* character (4 iterations), while the minimum still corresponds to the *I* character (2 iterations). As expected, the complexity of the character shape plays an essential role in the number of iterations needed for its removal.

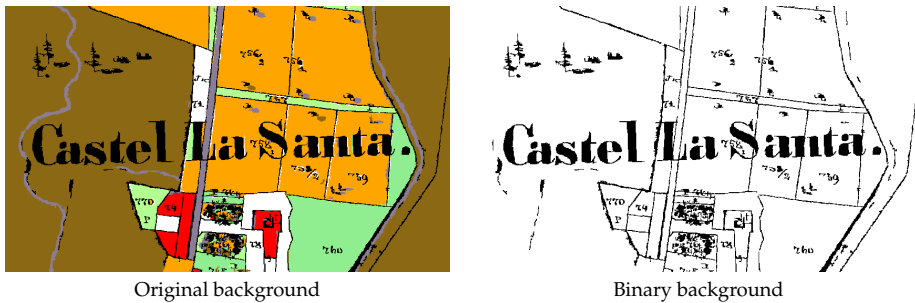


**Figure 9.** Number of iterations needed to remove 10 pt, occupying approximately  $20 \times 20$  pixels, lowercase characters with a  $3 \times 3$  (left) and a  $5 \times 5$  (right) filter. The single bar segment indicates the number of pixels removed from the character in each iteration. Iteration 0 indicates the number of initial pixels for each character.



**Figure 10.** Number of iterations needed to remove 10 pt, occupying approximately  $20 \times 20$  pixels, uppercase characters with a  $3 \times 3$  (left) and a  $5 \times 5$  (right) filter. The single bar segment indicates the number of pixels removed from the character in each iteration. Iteration 0 indicates the number of initial pixels for each character.

Finally, tests have been carried out on strings and symbols extracted from a historical map. The map used is part of the output of the classification of the Historical Cadaster Map for the Province of Trento, for a different map sheet than that digitized for this paper: the map and the methods used for text, symbols and lines classification are described in [48]. The removal technique has been applied to the original map (Figure 11, left) and to a binary map created by extracting the category corresponding to text, symbols and lines (Figure 11, right). The small trees and their shadows in the top left corner are symbols used to represent forest coverage and must be removed.



**Figure 11.** Test map containing text, symbols and lines from the classification of a real map. The left map has the original categories, the right one has been binarized with value 1 for text, symbols and lines and 0 for all other categories.

The removal process for the original, non binary, map is shown in Figure 12: the procedure correctly replaces the category of pixels belonging to text, symbols and lines with the category of the surrounding pixels.



**Figure 12.** Text, symbol and lines removal with  $23 \times 23$  filter: original map, map after two and nine iterations and the final map, with all text, symbols and lines removed, after 24 iterations.

Figure 13 describes the number of iterations needed to remove text, symbols and lines for the original map and the binary one. The procedure applied to the original map does not remove all the text strings for filter sizes below 21, while with the binary background a filter size of 11 is already effective.

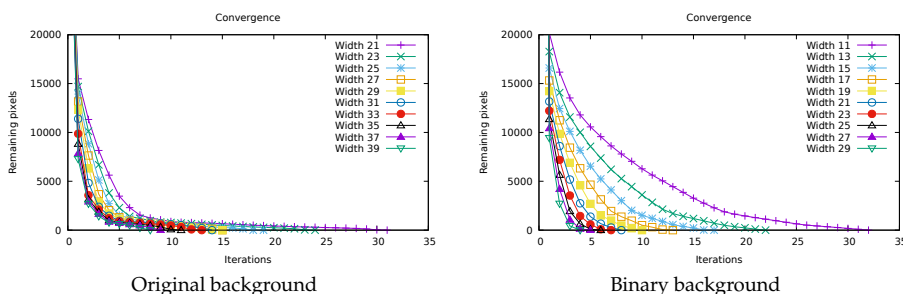


Figure 13. Number of iterations needed to remove the text for and historical map (Figure 11), with the original background (left) and after binarization (right), using different filter sizes.

Different filter sizes generate slightly different maps because the value of the mode can change. This is significant where text, symbols and lines are surrounded by pixels of mixed categories. This effect is visible in Figure 14 where the output map for the smallest and largest filter sizes used in the tests are compared.

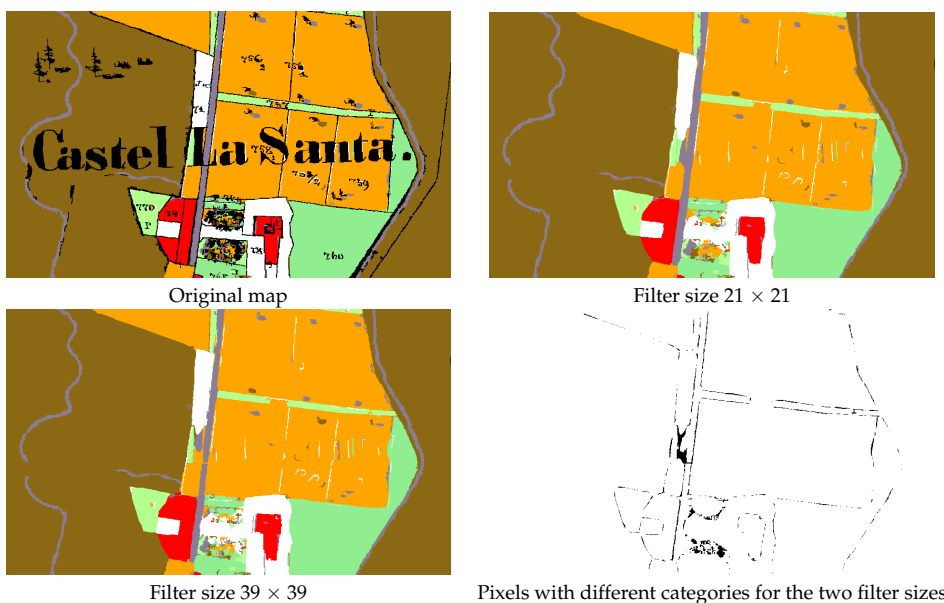


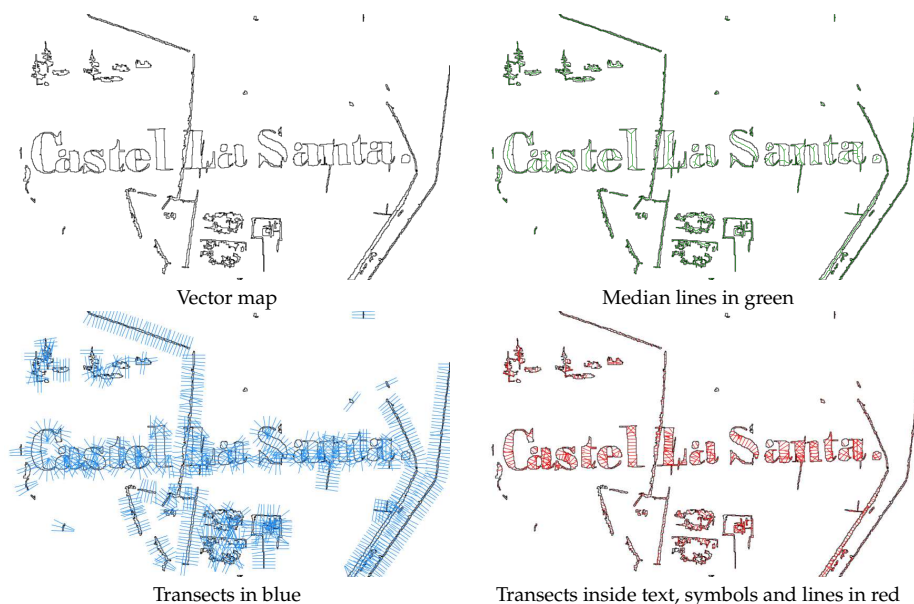
Figure 14. Output difference between 21 × 21 and 39 × 39 filter sizes: original map, output for the 21 × 21 and 39 × 39 filters and map of pixels with category difference between the two.

As a general indication, if the aim is to replace the category of text, symbols and lines with the categories of the surrounding pixels, smaller filter sizes provide better outcomes.

A new GRASS GIS module, *r.object.thickness* [73], has been created to estimate the smallest filter size needed to entirely remove text, symbols and lines on a map. The module finds the median line for each area to be removed, creates transects perpendicular to the median lines, and intersects them with the areas (Figure 15). The length of each transect is evaluated, and the minimum, maximum and mean

values are reported, both in map units and number of pixels. The number of pixels can be used to set the minimum filter width by using its half value plus one. The module is available as a GRASS add-on in the standard repository under the GNU General Public License ( $\geq v2$ ).

For the test map of Figure 11, the module estimates a maximum transect length of 13.09 m in map units, which, combined with the map resolution of around 0.32 m, corresponds to a maximum length of 41 pixels: the minimum filter size to remove all text, symbols and lines is 21 pixels. This value coincides with the minimum empirically found during the tests for the original (non-binary) map. The application of the module to the binary image containing a square (Figure 7 left) correctly returns as maximum value the length of the sides, 20 pixels.



**Figure 15.** Procedure for the evaluation of the minimum filter size to remove all the text, symbols and lines: original vector map, median lines, transects and transects inside text, symbols and lines' areas.

In this application, text and symbols are treated as noise that must be removed, but the use is planned of these features as information for the automatic creation of map labels as points in a vector map. The word spotting technique [74] will be tested with the *query-by-string* (QBS) approach [62,75], using the available list of toponyms as input. This approach has already been tested for maps, but text portions on the maps are obtained through image binarization [76,77], usually applying Otsu's global thresholding [78], rather than segmentation, and the focus is on words recognition rather than the separation of areas and labels on the map. Recent developments in character grouping in topographic maps [79] can be applied to accomplish the automatic creation of a label vector layer.

## 7. Results

The segmentation, classification and filtering methods outlined in Sections 4–6 have been applied to the maps described in Section 2.

The application of the segmentation parameters' optimization technique described in Section 4 to the three test maps, using the simple sum of the normalized VW and SA as optimization criteria, has given the results in Table 1. The *minsize* parameter values have been chosen to match the different sizes of text and symbols that have to be removed after the classification on each map. The results of the application of the Unsupervised Segment Parameter Optimization (USPO) for the optimal *threshold* determination reflect the features of the maps: the value is higher for the cadaster map because the

map, completely hand-drawn, presents more variable colors for the same LULC even inside a single segment. The IKFM and MPLFT maps have more uniform colors; therefore, a lower value of the threshold is already effective in separating the segments.

The classification of the test maps has been carried out using the LULC classes of Table 2.

**Table 2.** Land use/land cover (LULC) classes for the test maps: 1859 cadaster map (a), 1936 Italian Kingdom Forest Map (IKFM) (b) and 1992 Map of the Potential Limit of the Forest in Trentino (MPLFT) (c).

(a) Cadaster	
Class	Description
1	Forest
2	Crop land
3	Meadows
4	Pasture
5	Buildings
6	Roads
7	Gardens
8	Tracks
9	Water bodies
10	Map frame
11	Boundaries
12	Black text
13	Red text
14	Black lines
15	Red lines
(b) IKFM	
Class	Description
1	Conifers
8	Beech tree
20	Chestnut
23	Other species
24	Degraded forest
30	Background
31	City blocks
32	Brown text
33	Black symbols
34	Brown lines
35	Contour lines
36	Blue lines
37	Purple lines
(c) MPLFT	
Class	Description
1	Forest
2	Real potential [forest]
3	Excluded potential [forest]
4	Beyond [forest] limit
5	Water
6	Buildings
7	Background
8	Roads
9	Black text
10	Blue text
11	Black lines
12	Blue lines
13	Contour lines
14	Purple lines



For the cadaster map, LULC classes 1–11 have been set according to the classes of the original map, and the last four classes represent text and lines that are removed in the filtering step. Classes 1–24 of the IKFM correspond to the classes of the original map, and classes above 30 are filtered. For the MPLFT map, classes 1–4 match the classes of the original map, and classes 6–14 are eliminated after the classification.

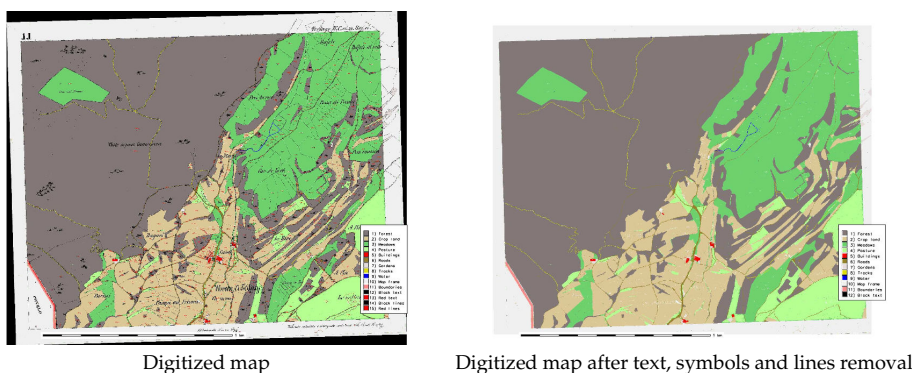
The method described in Section 6 has been applied to the classified maps using the *r.object.thickness* module, obtaining the minimum, mean and maximum thickness of the objects to be removed (Table 3). The filter size, used for the application of the *r.fill.category* module, is given by the smallest odd number larger than half the thickness in pixels. For the cadaster map, a filter size of 53 pixels instead of 51 has been used to account for possible approximations in the evaluation of the maximum thickness, very close to 102 pixels.

**Table 3.** Evaluation of the thickness of the objects to be removed and the corresponding filter size.

Map		Min	Mean	Max	Filter Size
Cadaster	[m]	0.004318	1.625650	32.705302	53
	[pixel]	0.013467	5.070013	101.999999	
IKFM	[m]	0.004409	59.438563	1249.561948	97
	[pixel]	0.000678	9.135900	192.061727	
MPLFT	[m]	0.004493	10.868955	276.521482	67
	[pixel]	0.002118	5.122713	130.329011	

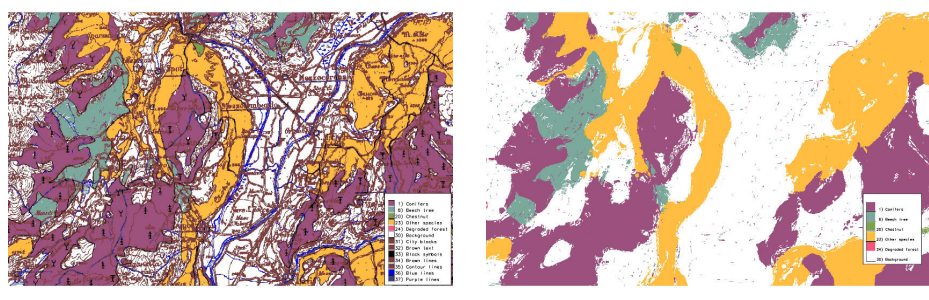
Maps have been filtered with *r.fill.category* using the filter sizes in Table 3, with four iterations for the cadaster map, 15 for the IKFM and four for the MPLFT map.

The results of the classification and filtering procedures are shown in Figures 16–18, while Table 4 reports the assessment of their accuracy. The colors of the digitized maps are chosen to mimic the colors of the corresponding classes on the original maps.



**Figure 16.** Digitized maps for the sheet 385\_11 sheet of the Historical Cadaster Map of the Province of Trento, “Comune catastale di Terlago”, Simple Weighted Vote (SWV) voting system.

The classification of the cadaster map provides good results, despite the fact that the map is hand-drawn because objects, text, lines and symbols are well separated. This facilitates the filtering phase. Some artifacts are due to the JPEG compression. It is also possible to create a map with distinct cadaster parcels with LULC classes as attributes combining the classified map with the map of the segments.

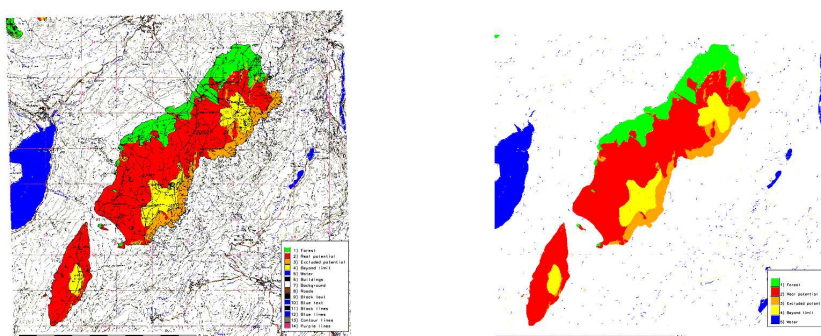


Digitized map

Digitized map after text, symbols and lines removal

**Figure 17.** Digitized maps for the Italian Kingdom Forest Map (IKFM) (particular), SWV voting system.

The classification of the IKFM gives positive results, taking into account the complexity of the background and the fact that forest areas are hand-drawn, with frequent overlapping features or gaps [3]. This map is the more difficult to classify and the filtering step takes considerably longer with respect to the other two maps.



Digitized map

Digitized map after text, symbols and lines removal

**Figure 18.** Digitized maps for the Map of the potential limit of the forest in Trentino (MPLFT) (particular), Best-Worst Weighted Vote (BWWV) voting system.

The classification of the MPLFT produces satisfactory results, even more so considering it uses halftones; therefore, colors are very mixed. Filtering text, symbols, and lines is fast, but some noise is left from the misclassification of some segments belonging to these classes.

The digitized maps have been checked against control maps created by manually digitizing the LULC classes areas. Only LULC classes representing forest have been digitized for the IKFM and MPLFT maps. Cohen's *Kappa* and the overall accuracy is shown in Table 4 before and after the post-process removal of text, symbols and lines described in Section 6. Since manually digitized areas include text, symbols and lines without distinction of LULC class, the values after post-processing must be used for the procedure assessment. The difference of *Kappa* and accuracy between pre- and post-processed maps reflects the complexity of their background: it is minimum for the cadaster map, which contains only a few text string and parcel boundary lines, and it is maximum for the IKFM, which has a very mixed background of text, contour lines and symbols. Table 4 shows a good agreement between the results of the procedure and the control maps, with values for post-processed maps of *Kappa* ranging from 0.96 to 0.97 and overall accuracy between 97% and 99%.

**Table 4.** Values of Kappa and accuracy (% observed correct) for the classification of the cadaster map, the Italian Kingdom Forest Map (IKFM) and the Map of the Potential Limit of the Forest in Trentino (MPLFT) with the four voting systems, Simple Majority Vote (SMV), Simple Weighted Vote (SWV), Best-Worst Weighted Vote (BWWV), Quadratic Best-Worst Weighted Vote (QBWWV), pre and post the application of the *r.fill.category* module.

Map	SMV		SWV		BWWV		QBWWV	
	Kappa	Accuracy	Kappa	Accuracy	Kappa	Accuracy	Kappa	Accuracy
Cadaster	0.945367	96.706618	0.952895	97.156840	0.949426	96.944917	0.949426	96.944917
Cadaster post	0.961574	97.727262	0.969815	98.212840	0.967224	98.058683	0.967224	98.058683
IKFM	0.552870	67.439228	0.551012	67.237156	0.553261	67.442162	0.553261	67.442162
IKFM post	0.887051	93.064664	0.892390	93.399099	0.892722	93.422756	0.892722	93.422756
MPLFT	0.727045	82.038126	0.745224	83.390072	0.745680	83.420707	0.745680	83.420707
MPLFT post	0.935200	96.101119	0.936114	96.152864	0.937423	96.264055	0.936315	96.165954

The lower values for the IKFM are explained by the difficulty in both the classification and filtering phases due to the very complex background with many overlapped features representing different LULC classes. The MPLFT map presents the additional complication of being printed using halftones, with high color variability inside a single segment.

These results mark a definitive improvement over the use of Maximum Likelihood Classification (unsupervised, supervised or supervised contextual) on similar maps [48], making the application of (semi) automatic classification a viable approach to the digitalization of heritage maps. The values for *Kappa* after post-processing are significantly higher than those reported for the forest class in [22].

## 8. Conclusions

This study provides an assessment of the application of standard OBIA procedures for the semi-automatic digitalization of heritage maps. Notably, this approach can be applied to heterogeneous sets of map using the same workflow, with the obvious care of evaluating the proper segmentation and classification parameters for each set. The text and symbol removal technique developed during this research has proved effective and enhances considerably the quality of the resulting maps, as shown in Table 4.

The next step in the development will be the study of the feasibility of the application of this techniques to large datasets, where a map is comprised of a hundred or a thousand map sheets. The main challenge is the reliability of the use of a single training set, based on one map sheet, for the classification of the whole dataset. Tests carried out on aerial images are encouraging, but more specific trials must be performed.

A new procedure is under development for identifying and counting symbols in each area, adding the type of symbol and the number of times it is found inside each area as attributes to the output vector map. A further future development involves the automatic creation of labels by identifying words on the map using word spotting techniques [74], with the creation of an additional vector layer containing toponyms and labels.

All the software used in this research, including modules developed ad hoc, is available as Free and Open Source Software in public repositories. Two of the maps analyzed, the “Historical Cadaster Map for the Province of Trento” and the 1936 “Italian Kingdom Forest Map”, are available under a Creative Commons license on the Web [3,36]. The 1992 “Map of the potential limit of the forest in Trentino” has been digitized and georeferenced for the first time during this study: consultations are under way with the copyright holder with the aim to make it available under a Creative Commons license. The availability of both software and input maps with open licenses makes it possible to check the results and easily extend the methods presented in this paper to other datasets.

**Author Contributions:** Research conceptualization, Paolo Zatelli, Stefano Gobbi and Marco Ciolli; methodology and software Paolo Zatelli and Stefano Gobbi; validation, Paolo Zatelli, Stefano Gobbi, Clara Tattoni and Marco Ciolli; writing—original draft preparation, Paolo Zatelli, Stefano Gobbi and Marco Ciolli; writing—review and editing, Stefano Gobbi, Marco Ciolli, Nicola La Porta, Duccio Rocchini, Clara Tattoni and Paolo Zatelli; funding acquisition, Marco Ciolli, Nicola La Porta and Paolo Zatelli

**Funding:** Stefano Gobbi is funded by Doctoral Programme in Civil, Environmental and Mechanical Engineering established at the Department of Civil, Environmental and Mechanical Engineering at the University of Trento and Fondazione Edmund Mach.

**Acknowledgments:** The cadaster historical map for the Province of Trento has been made available by the Servizio Catasto of the Provincia Autonoma di Trento, under the Creative Commons Attribution 4.0 license.

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

## Abbreviations

The following abbreviations are used in this manuscript:

BWWV	Best-Worst Weighted Vote
CC	Compact Circle
CS	Compact Square
DPI	Digit Per Inch
DT	Decision Tree
EPSG	European Petroleum Survey Group (EPSG Geodetic Parameter Set)
FD	Fractal Dimension
GCP	Ground Control Point
GIS	Geographic Information System
GPL	Gnu Public License
FOSS	Free and Open Source Software
IGMI	Istituto Geografico Militare Italiano (Italian Military Geographic Institute)
IKFM	Italian Kingdom Forest Map
k-NN	k-Nearest-Neighbor classifier
LULC	Land Use/Land Cover
MLP	Multi-Layer Perceptron
MPLFT	Map of the Potential Limit of the Forest in Trentino
NB	Naive Bayesian classifier
OBIA	Object-Based Image Analysis
OCR	Optical Character Recognition
QBS	Query-By-String
QBWWV	Quadratic Best-Worst Weighted Vote
SA	Spatial Autocorrelation
SMV	Simple Majority Vote
SVM	Support Vector Machine
SWV	Simple Weighted Vote
USPO	Unsupervised Segment Parameter Optimization
VW	intra-object Variance Weighted by object size

## References

1. Loran, C.; Haegi, S.; Ginzler, C. Comparing historical and contemporary maps - a methodological framework for a cartographic map comparison applied to Swiss maps. *Int. J. Geogr. Inf. Sci.* **2018**, *32*, 2123–2139. [[CrossRef](#)]
2. Ciolli, M.; Bezzi, M.; Comunello, G.; Laitempergher, G.; Gobbi, S.; Tattoni, C.; Cantiani, M. Integrating dendrochronology and geomatics to monitor natural hazards and landscape changes. *Appl. Geomat.* **2019**, *11*, 39–52. [[CrossRef](#)]
3. Ferretti, F.; Sboarina, C.; Tattoni, C.; Vitti, A.; Zatelli, P.; Geri, F.; Pompei, E.; Ciolli, M. The 1936 Italian Kingdom Forest Map reviewed: A dataset for landscape and ecological research. *Ann. Silv. Res.* **2018**, *42*, 3–19. [[CrossRef](#)]

4. Tattoni, C.; Ciolli, M.; Ferretti, F. The fate of priority areas for conservation in protected areas: A fine-scale markov chain approach. *Environ. Manag.* **2011**, *47*, 263–278. [[CrossRef](#)] [[PubMed](#)]
5. Sitzia, T.; Semenzato, P.; Trentanovi, G. Natural reforestation is changing spatial patterns of rural mountain and hill landscapes: A global overview. *For. Ecol. Manag.* **2010**, *259*, 1354–1362. [[CrossRef](#)]
6. Senapathi, D.; Carvalheiro, L.; Biesmeijer, J.; Dodson, C.A.; Evans, R.; McKerchar, M.; Morton, D.; Moss, E.; Roberts, S.; Kunin, W.; et al. The impact of over 80 years of land cover changes on bee and wasp pollinator communities in England. *Proc. R. Soc. B Biol. Sci.* **2015**, *282*. [[CrossRef](#)]
7. Cousins, S.; Auffret, A.; Lindgren, J.; Tränk, L. Regional-scale land-cover change during the 20th century and its consequences for biodiversity. *Ambio* **2015**, *44*, 17–27. [[CrossRef](#)] [[PubMed](#)]
8. Feurdean, A.; Munteanu, C.; Kummerle, T.; Nielsen, A.; Hutchinson, S.; Ruprecht, E.; Parr, C.; Perşoiu, A.; Hickler, T. Long-term land-cover/use change in a traditional farming landscape in Romania inferred from pollen data, historical maps and satellite images. *Reg. Environ. Chang.* **2017**, *17*, 2193–2207. [[CrossRef](#)]
9. Auffret, A.; Kimberley, A.; Plue, J.; Skånes, H.; Jakobsson, S.; Waldén, E.; Wennbom, M.; Wood, H.; Bullock, J.; Cousins, S.; et al. HistMapR: Rapid digitization of historical land-use maps in R. *Methods Ecol. Evol.* **2017**, *8*, 1453–1457. [[CrossRef](#)]
10. Tattoni, C.; Ianni, E.; Geneletti, D.; Zatelli, P.; Ciolli, M. Landscape changes, traditional ecological knowledge and future scenarios in the Alps: A holistic ecological approach. *Sci. Total Environ.* **2017**, *579*, 27–36. [[CrossRef](#)]
11. Geri, F.; La Porta, N.; Zottele, F.; Ciolli, M. Mapping historical data: Recovering a forgotten floristic and vegetation database for biodiversity monitoring. *ISPRS Int. J. Geo-Inf.* **2016**, *5*. [[CrossRef](#)]
12. Spooner, P.; Shoard, J. Using historic maps and citizen science to investigate the abundance and condition of survey reference ‘blaze’ trees. *Aust. J. Bot.* **2016**, *64*, 377–388. [[CrossRef](#)]
13. Alberico, I.; Cavuoto, G.; Di Fiore, V.; Punzo, M.; Tarallo, D.; Pelosi, N.; Ferraro, L.; Marsella, E. Historical maps and satellite images as tools for shoreline variations and territorial changes assessment: The case study of Volturno Coastal Plain (Southern Italy). *J. Coast. Conserv.* **2018**, *22*, 919–937. [[CrossRef](#)]
14. Tortora, A.; Statuto, D.; Picuno, P. Rural landscape planning through spatial modelling and image processing of historical maps. *Land Use Policy* **2015**, *42*, 71–82. [[CrossRef](#)]
15. Gabriel, N. Mapping urban space: The production, division and reconfiguration of natures and economies. *City* **2013**, *17*, 325–342. [[CrossRef](#)]
16. Gretter, A.; Ciolli, M.; Scolozzi, R. Governing mountain landscapes collectively: Local responses to emerging challenges within a systems thinking perspective. *Landsc. Res.* **2018**, *43*, 1117–1130. [[CrossRef](#)]
17. Svenningsen, S.; Levin, G.; Perner, M. Military land use and the impact on landscape: A study of land use history on Danish Defence sites. *Land Use Policy* **2019**, *84*, 114–126. [[CrossRef](#)]
18. Ciolli, M.; Tattoni, C.; Ferretti, F. Understanding forest changes to support planning: A fine-scale Markov chain approach. *Dev. Environ. Model.* **2012**, *25*, 355–373. [[CrossRef](#)]
19. Willcock, S.; Phillips, O.; Platts, P.; Swetnam, R.; Balmford, A.; Burgess, N.; Ahrends, A.; Bayliss, J.; Doggart, N.; Doody, K.; et al. Land cover change and carbon emissions over 100 years in an African biodiversity hotspot. *Glob. Chang. Biol.* **2016**, *22*, 2787–2800. [[CrossRef](#)]
20. Chiang, Y.Y.; Leyk, S.; Knoblock, C.A. A Survey of Digital Map Processing Techniques. *ACM Comput. Surv.* **2014**, *47*, 1–44. [[CrossRef](#)]
21. Liu, T.; Xu, P.; Zhang, S. A review of recent advances in scanned topographic map processing. *Neurocomputing* **2019**, *328*, 75–87. [[CrossRef](#)]
22. Herrault, P.A.; Sheeren, D.; Fauvel, M.; Paegelow, M. Automatic extraction of forests from historical maps based on unsupervised classification in the CIELab color space. In *Lecture Notes in Geoinformation and Cartography*; Springer: Cham, Switzerland, 2013; pp. 95–112. [[CrossRef](#)]
23. Gobbi, S.; Maimeri, G.; Tattoni, C.; Cantiani, M.; Rocchini, D.; La Porta, N.; Ciolli, M.; Zatelli, P. Orthorectification of a large dataset of historical aerial images: Procedure and precision assessment in an open source environment. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences—ISPRS Archives, Dar es Salaam, Tanzania, 29–31 August 2018; pp. 53–59. [[CrossRef](#)]
24. Gobbi, S.; Cantiani, M.; Rocchini, D.; Zatelli, P.; Tattoni, C.; Ciolli, M.; La Porta, N. Fine spatial scale modelling of Trentino past forest landscape (Trentinoland): A case study of FOSS application. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences—ISPRS Archives, Bucharest, Romania, 26–30 August 2019; pp. 71–78. [[CrossRef](#)]

25. Bhowmik, S.; Sarkar, R.; Nasipuri, M.; Doermann, D. Text and non-text separation in offline document images: A survey. *Int. J. Doc. Anal. Recognit.* **2018**, *21*. [[CrossRef](#)]
26. GRASS Development Team. Geographic Resources Analysis Support System (GRASS) Software. Open Source Geospatial Foundation. Available online: [grass.osgeo.org](http://grass.osgeo.org) (accessed on 20 September 2017).
27. Ciolli, M.; Federici, B.; Ferrando, I.; Marzocchi, R.; Sguerso, D.; Tattoni, C.; Vitti, A.; Zatelli, P. FOSS tools and applications for education in geospatial sciences. *ISPRS Int. J. Geo-Inf.* **2017**, *6*. [[CrossRef](#)]
28. Simeoni, L.; Zatelli, P.; Floretta, C. Field measurements in river embankments: Validation and management with spatial database and webGIS. *Nat. Hazards* **2014**, *71*, 1453–1473. [[CrossRef](#)]
29. Preatoni, D.; Tattoni, C.; Bisi, F.; Masseroni, E.; D’Acunto, D.; Lunardi, S.; Grimod, I.; Martinoli, A.; Tosi, G. Open source evaluation of kilometric indexes of abundance. *Ecol. Inform.* **2012**, *7*, 35–40. [[CrossRef](#)]
30. Federici, B.; Giacomelli, D.; Sguerso, D.; Vitti, A.; Zatelli, P. A web processing service for GNSS realistic planning. *Appl. Geomat.* **2013**, *5*, 45–57. [[CrossRef](#)]
31. Timár, G. System of the 1:28 800 scale sheets of the Second Military Survey in Tyrol and Salzburg. *Acta Geod. Et Geophys. Hung.* **2009**, *44*, 95–104. [[CrossRef](#)]
32. von Nischer-Falkenhof, E. The Survey by the Austrian General Staff under the Empress Maria Theresa and the Emperor Joseph II., and the Subsequent Initial Surveys of Neighbouring Territories during the Years 1749–1854. *Imago Mundi* **1937**, *2*, 83–88. [[CrossRef](#)]
33. Buffoni, D.; Girardi, S.; Revolti, R.; Cortese, G.; Mastronunzio, M. HISTORICALKat. La documentazione catastale trentina d’impianto è Open Data. In Proceedings of the XX Conferenza Nazionale ASITA, Trieste, Italy, 12–14 November 2019; pp. 112–117.
34. Servizio Catasto della Provincia Autonoma di Trento. HistoricalKat, Documentation of the Trentino Cadastral Archive. Available online: <http://historicalkat.provincia.tn.it> (accessed on 29 April 2019).
35. Servizio Catasto della Provincia Autonoma di Trento. Mapped storiche di impianto (Urmapped). Available online: <https://www.catastotn.it/mappedstoriche.html> (accessed on 29 April 2019).
36. OPENData Trentino. Historical Cadaster Map of the Trentino Region Metadata. Available online: <https://dati.trentino.it/dataset/mappedstoriche-d-impianto> (accessed on 29 April 2019).
37. Revolti, R. Produzione e restauro della cartografia catastale in Trentino. Available online: <https://drive.google.com/file/d/0B0abPZ1IrxqdcnUwXzBIT2RvV0E/view> (accessed on 29 April 2019).
38. Brengola, A. La carta Forestale d’Italia. [The Italian Forest Map]. In *Rivista Forestale Italiana. Milizia Nazionale Forestale*; Comando Centrale: Rome, Italy, 1939; pp. 1–10.
39. Piussi, P. *Carta del limite potenziale del bosco in Trentino*; Technical Report; Servizio Foreste Caccia e Pesca: Provincia Autonoma di Trento, Italy, 1992.
40. Liu, Y.; Guo, J.; Lee, J. Halftone Image Classification Using LMS Algorithm and Naive Bayes. *IEEE Trans. Image Process.* **2011**, *20*, 2837–2847. [[CrossRef](#)]
41. Hackeloeer, A.; Klasing, K.; Krisp, J.; Meng, L. Georeferencing: A review of methods and applications. *Ann. GIS* **2014**, *20*, 61–69. [[CrossRef](#)]
42. Zitová, B.; Flusser, J. Image registration methods: A survey. *Image Vis. Comput.* **2003**, *21*, 977–1000. [[CrossRef](#)]
43. Jenny, B.; Hurni, L. Studying cartographic heritage: Analysis and visualization of geometric distortions. *Comput. Graph.* **2011**, *35*, 402–411. [[CrossRef](#)]
44. Yu, H.; Yang, Z.; Tan, L.; Wang, Y.; Sun, W.; Sun, M.; Tang, Y. Methods and datasets on semantic segmentation: A review. *Neurocomputing* **2018**, *304*, 82–103. [[CrossRef](#)]
45. Vaithyanathan, V.; Rajappa, U. A review on clustering techniques in image segmentation. *Int. J. Appl. Eng. Res.* **2013**, *8*, 2685–2688.
46. Belgiu, M.; Drăgău, L. Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. *Int. J. Remote Sens.* **2014**, *96*, 67–75. [[CrossRef](#)] [[PubMed](#)]
47. Gao, Y.; Mas, J. A comparison of the performance of pixel based and object based classifications over images with various spatial resolutions. *Online J. Earth Sci.* **2008**, *2*, 27–35.
48. Zatelli, P.; Gobbi, S.; Tattoni, C.; La Porta, N.; Ciolli, M. Object-based image analysis for historic maps classification. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences—ISPRS Archives, Bucharest, Romania, 26–30 August 2019; pp. 247–254. [[CrossRef](#)]

49. Momsen, E.; Metz, M. GRASS Development Team i.segment. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6. 2016. Available online: <https://grass.osgeo.org/grass76/manuals/i.segment.html> (accessed on 25 July 2019).
50. Peng, Q. An image segmentation algorithm research based on region growth. *J. Softw. Eng.* **2015**, *9*, 673–679.
51. Grippa, T.; Lennert, M.; Beaumont, B.; Vanhuyse, S.; Stephenne, N.; Wolff, E. An open-source semi-automated processing chain for urban object-based classification. *Remote Sens.* **2017**, *9*. [[CrossRef](#)]
52. Lennert, M. GRASS Development Team Addon i.segment.uspo. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6. 2016. Available online: <https://grass.osgeo.org/grass72/manuals/addons/i.segment.uspo.html> (accessed on 25 July 2019).
53. Espindola, G.; Camara, G.; Reis, I.; Bins, L.; Monteiro, A. Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. *Int. J. Remote Sens.* **2006**, *27*, 3035–3040. [[CrossRef](#)]
54. Johnson, B.; Bragais, M.; Endo, I.; Magcale-Macandog, D.; Macandog, P. Image segmentation parameter optimization considering within-and between-segment heterogeneity at multiple scale levels: Test case for mapping residential areas using landsat imagery. *ISPRS Int. J. Geo-Inf.* **2015**, *4*, 2292–2305. [[CrossRef](#)]
55. Schiewe, J. Segmentation of high-resolution remotely sensed data-concepts, applications and problems. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences—ISPRS Archives, Volume XXXIV, Part 4, Ottawa, QC, Canada, 9–12 July 2002; pp. 380–385.
56. Carleer, A.; Debeir, O.; Wolff, E. Assessment of very high spatial resolution satellite image segmentations. *Photogramm. Eng. Remote Sens.* **2005**, 1285–1294. [[CrossRef](#)]
57. Metz, M.; Lennert, M. GRASS Development Team Addon i.segment.stats. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6. 2016. Available online: <https://grass.osgeo.org/grass76/manuals/addons/i.segment.stats.html> (accessed on 25 July 2019).
58. Metz, M.; Lennert, M. GRASS Development Team Addon r.object.geometry. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6. 2016. Available online: <https://grass.osgeo.org/grass72/manuals/addons/r.object.geometry.html> (accessed on 25 July 2019).
59. Kuhn, M. The R Caret Library. 2019. Available online: <https://topepo.github.io/caret/> (accessed on 25 July 2019).
60. Moreno-seco, F. Structural, Syntactic, and Statistical Pattern Recognition. *Struct. Syntactic Stat. Pattern Recognit.* **2006**. [[CrossRef](#)]
61. Fletcher, L.A.; Kasturi, R. A Robust Algorithm for Text String Separation from Mixed Text/Graphics Images. *IEEE Trans. Pattern Anal. Mach. Intell.* **1988**, *10*, 910–918. [[CrossRef](#)]
62. Mhiri, M.; Desrosiers, C.; Cheriet, M. Word spotting and recognition via a joint deep embedding of image and text. *Pattern Recognit.* **2019**, *88*, 312–320. [[CrossRef](#)]
63. Ye, Q.; Doermann, D. Text Detection and Recognition in Imagery: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1480–1500. [[CrossRef](#)] [[PubMed](#)]
64. Moreno-García, C.F.; Elyan, E.; Jayne, C. New trends on digitisation of complex engineering drawings. *Neural Comput. Appl.* **2019**, *31*, 1695–1712. [[CrossRef](#)]
65. Kang, S.O.; Lee, E.B.; Baek, H.K. A Digitization and Conversion Tool for Imaged Drawings to Intelligent Piping and Instrumentation Diagrams (P&ID). *Energies* **2019**, *12*. [[CrossRef](#)]
66. Chiang, Y.Y.; Knoblock, C.A. An approach for recognizing text labels in raster maps. In Proceedings of the International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 3199–3202. [[CrossRef](#)]
67. Rokon, M.O.F.; Masud, M.R.; Islam, M.M. An iterative approach to detect region boundary eliminating texts from scanned land map images. In Proceedings of the 3rd International Conference on Electrical Information and Communication Technology, EICT 2017, Khulna, Bangladesh, 7–9 December 2017; pp. 1–4. [[CrossRef](#)]
68. Poudroux, J.; Gonzato, J.; Pereira, A.; Guitton, P. Toponym Recognition in Scanned Color Topographic Maps. In Proceedings of the Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), Washington, DC, USA, 23–26 September 2007; Volume 1, pp. 531–535. [[CrossRef](#)]
69. Roy, P.P.; Lladós, J.; Pal, U. Text/Graphics Separation in Color Maps. In Proceedings of the 2007 International Conference on Computing: Theory and Applications (ICCTA'07), Kolkata, India, 5–7 March 2007; pp. 545–551. [[CrossRef](#)]
70. Chiang, Y.Y.; Leyk, S.; Honarvar Nazari, N.; Moghaddam, S.; Tan, T.X. Assessing the impact of graphical quality on automatic text recognition in digital maps. *Comput. Geosci.* **2016**, *93*, 21–35. [[CrossRef](#)]

71. Naidu, P.S.; Mathew, M. (Eds.) Chapter 4 Digital filtering of maps I. In *Analysis of Geophysical Potential Fields*; Elsevier: Amsterdam, The Netherlands, 1998; Volume 5, *Advances in Exploration Geophysics*; pp. 145–221. [[CrossRef](#)]
72. Gobbi, S.; Zatelli, P. GRASS Development Team Addon r.fill.category. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6. 2019. Available online: <https://grass.osgeo.org/grass76/manuals/addons/r.fill.category.html> (accessed on 25 July 2019).
73. Zatelli, P. GRASS Development Team Addon r.object.thickness. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6. 2019. Available online: <https://grass.osgeo.org/grass76/manuals/addons/r.object.thickness.html> (accessed on 25 July 2019).
74. Giotis, A.; Sfikas, G.; Gatos, B.; Nikou, C. A survey of document image word spotting techniques. *Pattern Recognit.* **2017**, *68*, 310–332. [[CrossRef](#)]
75. Krishnan, P.; Dutta, K.; Jawahar, C. Word spotting and recognition using deep embedding. In Proceedings of the 13th IAPR International Workshop on Document Analysis Systems, Vienna, Austria, 24–27 April 2018; pp. 1–6. [[CrossRef](#)]
76. Roy, P.; Pal, U.; Lladós, J. Query driven word retrieval in graphical documents. In Proceedings of the ACM International Conference Proceeding Series, Boston, MA, USA, 9–11 June 2010; pp. 191–198. [[CrossRef](#)]
77. Tarafdar, A.; Pal, U.; Roy, P.; Ragot, N.; Ramel, J.Y. A two-stage approach for word spotting in graphical documents. In Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, Washington, DC, USA, 25–28 August 2013; pp. 319–323. [[CrossRef](#)]
78. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
79. Xu, P.; Miao, Q.; Liu, T.; Chen, X.; Nie, W. Graphic-based character grouping in topographic maps. *Neurocomputing* **2016**, *189*, 160–170. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



### 3.2.2 Assessment of the results of the classification

The stratified sampling approach was applied to the maps of forest, extracted from the land use classification for the 1954, 1994, 2006 and 2015 datasets to assess the accuracy of the OBIA procedure in recognizing the forest areas.

This step was crucial for discussing the results because it reveals how much the forest cover and metric analysis are reliable. The accuracy was calculated by placing a total of 750 points randomly scattered over the PAT territory, as explained in section 2.5.6. The accuracy was then assessed by calculating the percentage of points for each land use that were classified according to ground truth.

After changing training area and checking manually the overall result of the classification procedure, the best classification was tested for its accuracy. The results of this procedure are reported in table 33.

Year	Accuracy calculated [%]
1954	90
1994	91
2006	95
2015	94

Table 33: Accuracy level of the OBIA procedure for forest class

The results show an overall accuracy greater than 90% for all the datasets, with a considerable improvement of accuracy level for the multiband imagery datasets. Although this result may seem satisfactory, the analysis of landscape metrics revealed a series of issues raised in the 1954 and 1994 dataset.

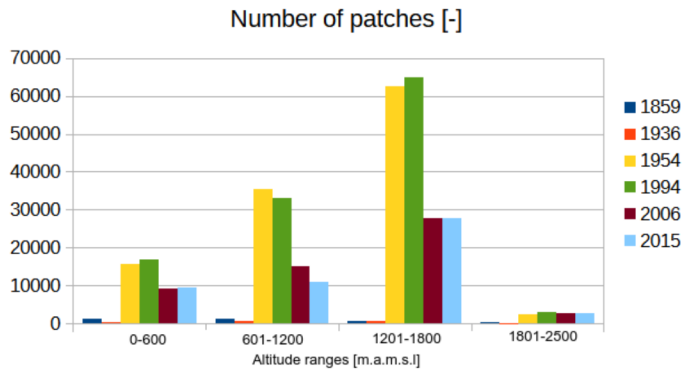


Figure 55: The calculus of the number of patches metric for the forest for the different years, divided by altitude ranges.

As the graphs 55 shows, the analysis of the number of patches revealed a situation where the difference in NP between 1994 and 2006 is considerably high and, in counter tendency from what revealed from other studies [72, 108, 16], it seemed that the number of patches is increasing from 1954 to 1994, for all the altitude ranges.

By a further and more detailed visual analysis it seemed that the problem in the dataset of 1954 arose in the segmentation phase. The main issue was that objects including areas of

forests were segmented in open areas and vice-versa. To avoid repeating the entire classification process it was decided to automatically correct the results of the classification for the 1954 dataset.

The procedure used to perform this correction is based on classifying pixels in the forest or meadows land use category according to their pixel value. The concept applied was to evaluate a threshold value which could detect those false positive pixels for the forest area that were misidentified during the segmentation phase and those false negative pixels which were misidentified as open areas during the segmentation phase. The areas of forest extracted with the OBIA procedure were analysed with the *r.stats.quantile* module in GRASS GIS to find out the quantile distribution of the pixel value, and the same was made for the pasture and meadows class. The pixels misclassified in the forest class were the highest in pixel value value, therefore a threshold value on the pixel value needed to be set up to exclude all those false positive pixels in the forest area. In the same way the false negative pixels in the meadows and pasture class were detected by imposing a threshold value for each map: darker pixels, in other words with a value of pixel value minor than the threshold were detected as false negatives and then aggregated to the forest class. By figuring these two threshold the aim was to eliminate as much of the false positives and detect as much as the false negatives disturbing the least possible the real positives pixels for the forest areas

The procedure was tested on two images of the 1954 dataset 18\_3821 and 21\_4575, one image per type of histogram. To test the effect of the two threshold values in detecting false positives and negatives some meadows and forest areas were manually digitized on the two images which were incorrectly identified. The percentage of pixels remaining in the incorrect class after applying the correction process were evaluated.

Quantile	Threshold	% remaining false positives
60	112	0
65	118	0
70	124	0
75	131	0
80	139	0
85	148	0
90	159	0.013
95	174	0.081
100	196	1.47

Table 34: Evaluation of the threshold value of the false positives for the forest class in the 18\_3821 image.

Quantile	Threshold	% remaining false positives
60	111	0
65	118	0
70	126	0
75	134	0.0070
80	142	0.0070
85	152	0.091
90	164	1.23
95	178	7.05
100	198	38.66

Table 35: Evaluation of the threshold value of the false positives for the forest class in the 21\_4575 image.

Quantile	Threshold	% remaining false negatives
10	70	14.86
15	84	9.88
20	98	5.18
25	110	1.99
30	121	0.73
35	129	0.33
40	136	0.19
45	143	0.085
50	149	0.047
55	154	0.023
60	160	0.019
65	166	0.019
70	173	0.0047
75	180	0.0047
80	187	0.0047

Table 36: Evaluation of the threshold value of the false negatives for the forest class in the 18\_3281 image

Quantile	Threshold	% remaining fake negatives
10	68	32.21
15	95	5.52
20	113	2.52
25	124	1.76
30	133	1.28
35	140	0.79
40	147	0.57
45	153	0.46
50	159	0.29
55	164	0.18
60	170	0.13
65	175	0.09
70	181	0.03
75	187	0.02
80	193	0.01

Table 37: Evaluation of the threshold value of the fake negatives for the forest class in the 21\_4575 image

It was then chosen to use a threshold value of the 80% of the reflectance distribution in the forest areas of each image of the 1954 dataset for detecting the false positives of the forest class. This value, in fact, was the lowest threshold value in both the analysed images which eliminated more than the 99% of false positives without compromising the real positives. Similarly, it was decided to use the 30% of the reflectance distribution in the meadows and pasture class of each image of the 1954 dataset for detecting the false negatives in the forest class. This value was the highest threshold in both the analysed images which eliminated more than the 99% of fake negatives without compromising the real positives.

This process was performed on each image of the 1954 dataset, with the identification of the two different threshold values and generated for each image a map of the forest false positives and forest false negative.

False negatives were then added to the map of the 1954 forest, and the false positives were subtracted.

The issue in the 1994 dataset have arisen in some valleys with a East-West orientation (Val di Sole, Valsugana) where some areas of shaded forest were incorrectly classified as shadows. By running a new OBIA process, adding training areas for the forest class selected in the darkest parts of the forests it was possible to perform a new classification which allowed to better recognize those darkest parts of the forest.

The various classification were then combined using *r.mapcalc* to generate a new map of forest for the 1994 dataset pixels classified as forest in one or both maps are labelled as forest in the output map.

The stratified random approach was then re-evaluated for the newly generated maps of the forests in the 1954 dataset and in the 1994 dataset.

Year	Accuracy [%]
1954	93
1994	93
2006	95
2015	94

Table 38: Accuracy level of the OBIA procedure for forest class, after the correction of the 1954 and 1994 dataset.

The improvements were detected by the procedure, but the stratified approach proved its limits in analysing the accuracy of LULC for regional scale datasets. This issue highlight how it is recommendable to always perform a visual assessment of the LULC result before running a systematic procedure for accuracy assessment, especially if the analysis is performed in wide, diverse landscapes.

### 3.3 THE RESULTS OF THE CLASSIFICATION

The aim of the classification was to obtain, for each dataset, a map of the forest coverage as accurate as possible. From the classified maps were extracted the classes 1 and 2, according to 5 the areas of forest and shaded forests. The map of forest as difference between DSM and DTM, described in 2.2.5 was used in this phase as a mask to re-classify all the areas classified as 'shadows' inside it as forest across the different datasets. This was performed assuming that this map represent the maximal extension of forest in the PAT, therefore shadows outside this map in each dataset likely does not contain forests, and areas classified as shadows inside this area are likely forest in all the dataset used in the analysis.

The figures 56 reports the extracted areas of forests for the years 1954, 1994, 2006, 2015. To have an homogeneous and comparable series of results, before comparing the different datasets, the forest maps were all resampled to a resolution of 10x10m pixels.

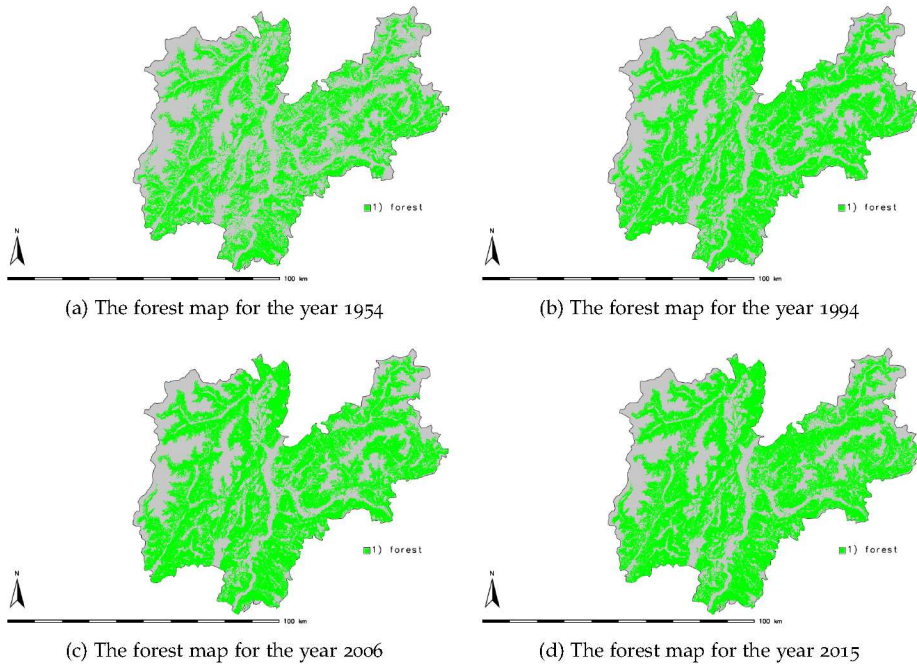


Figure 56: The forest area extracted from the Land use classification process.

Under the hypothesis explained in section 3.2 the maps of an hypothetical forest distribution for the years 1859 and 1936 were produced.

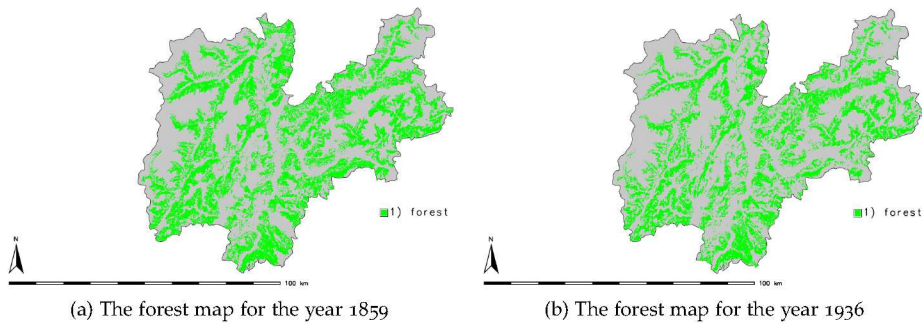


Figure 57: The forest area extracted from the 1859 and 1936 dataset adapted to the forest of 1954.

The study of the forest area through the years reveals a tendency of forest expansion, as reported in table 39.

Year	Coverage [ha]	% coverage on the whole PAT area
1859	212477	34.22
1936	153818	24.77
1954	277423	44.68
1994	347439	55.96
2006	357016	57.50
2015	359308	57.87

Table 39: Total forest coverage for the PAT over the years

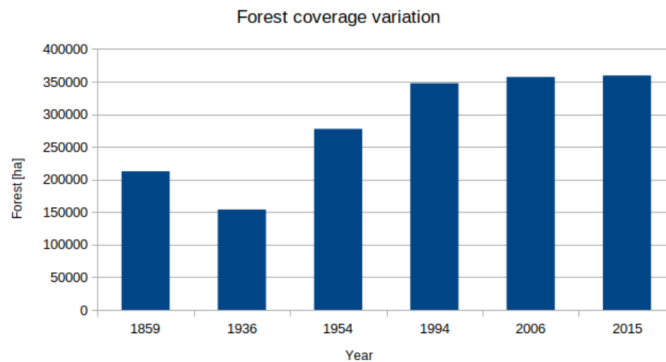


Figure 58: The evolution of forest coverage through the years.

From a 44% of the landscape in PAT composed by forests in 1954, in 40 years the forests in 1994 covered the 55% of PAT territory, a phenomenon that continued steadily but slowly from the year 1994 to the year 2006.

In fact, assuming that the expansion kept a constant rate over a single year, from the year 1954 to the year 1994 the net expansion was about 1750 ha/year, while from 1994 to 2006 the net expansion rate was about 800 ha/year; furthermore, in the period 2006-2015 the net expansion rate was about 250 ha/year. The slowdown of this process of expansion could be due to the fact that along a progressive abandonment of rural areas, which drove the forest expansion, the areas where trees could easily grow were becoming less and less available. Therefore it is possible to assume that in absence of a strong human intervention or external factors (contrary to what happened in 2018, as explained in section 3.6) the process of afforestation had an asymptotic nature toward the limit of maximum possible area that could be occupied by trees.

As the data of 1859 and 1936 was concerned it was possible to notice a tendency in deforestation from 1859 to 1936. The reasons behind this process was that forest in PAT suffered a strong exploitation in the years of the World War first from the both the armies involved (Austro-Hungarian and Italian) and in the following years, from the local population, who, as always after a war, was indigent and heavily appealed to every single resource that its environment offered [59, 52]. Finally, it is worth reminding that the original data of the 1936 I.K.F.M. had a much coarser scale respect to the other mapsets in the analysis and the result may be affected by this. Nevertheless the data indicate a situation where forest covered a fraction of the area present in the year 1954. In the period between 1936 and 1954 the Italian government also supported the artificial reforestation of some areas, contributing to a general recovery of forest coverage in Trentino [7].

The data about forest cover was, for each year, divided in altitude ranges, derived from the ISTAT classification of altitude ranges for the Italian municipalities [56]:

- 0-600 m a.m.s.l. valley bottom range;
- 601-1200 m a.m.s.l. low mountain range;
- 1201-2200 m a.m.s.l. mid mountain range;
- 2201-3000 m a.m.s.l. high mountain range;
- 3001-4000 m a.m.s.l. glacier range.

The data about forest coverage in these ranges is reported in table 40.

Altitude [mamsl]	1859	1936	1954	1994	2006	2015
0-600	13069.03	10444.56	18480.76	26703.9	29444.93	28179.02
601-1200	77603.58	67302.09	94171.29	120105.42	124865.17	126216.94
1201-2200	121690.23	101900.51	158669.73	199718.36	201727.83	203755.43
2201-3000	109.56	28.14	900.1	900.59	963.04	1142.91

Table 40: Forest coverage (in ha) divided for altitude ranges

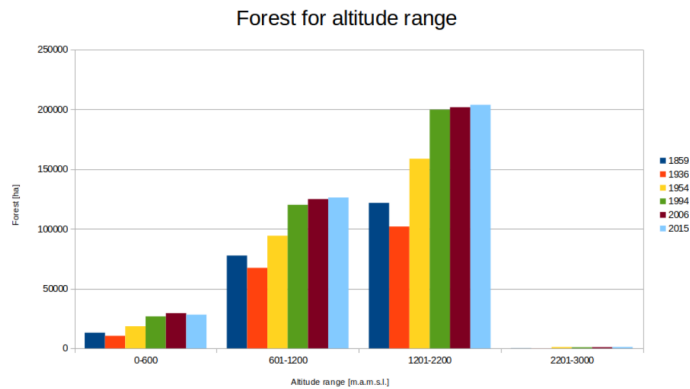


Figure 59: The evolution of forest coverage trough the years divided per altitude ranges.

This unbundling of the data underlined that the process of afforestation had the same dynamics in every altitude range, with the exception of the high-mountain range, which saw a dramatic increase from the 1994 to the 2006. This could be an effect of the global warming, which have retreated the snow limit toward higher altitude and have warmed the climate at higher altitude, making it possible for trees to colonize those areas where the climatic conditions were previously too harsh to thrive.

The data about forest coverage in different altitude ranges were analysed furthermore by unbundling the forest area in the altitude range 2201-2500 mamsl, about the altitude of the treeline in the PAT latitude range [108]. This was made to highlight how the rate of the afforestation at higher altitudes proceeded during the years of the datasets in the analysis.

From 1994 to 2006 and 2015 the range of 2201-2500 mamsl was increasingly afforested, while the range of 2301-2400 mamsl which steadily had the same amount of forest since the year 2006 in the last years saw a fast increase in forest presence, as indicated in table 41.



Altitude range [maslm]	1859	1936	1954	1994	2006	2015
2201-2300	108	26	807	805	863	1011
2301-2400	6	0	80	87	85	114
2401-2500	0	0	1	2	3	3

Table 41: Forest coverage (in ha) for the altitude belts over the treeline limit

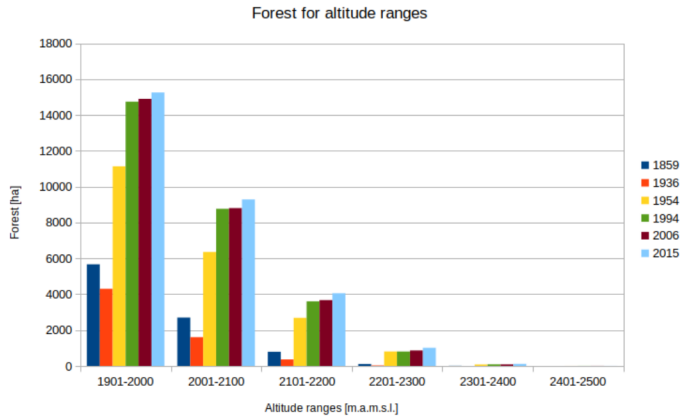


Figure 60: The evolution of forest coverage trough the years around the treeline limit

The forest coverage was then divided following the indication of Piussi [87] to analyse how the forest expanded in the areas described as Real potential, Excluded potential and Beyond limit.

	1859	1936	1954	1994	2006	2015
Forest	22359	19890	29571	37665	37530	37013
Real potential	4061	2688	9738	12364	12672	13213
Excluded potential	5969	3494	12388	15326	15926	16826
Beyond limit	3593	3121	5983	6820	7047	7153

Table 42: Forest coverage (in ha) for the potential or excluded potential areas

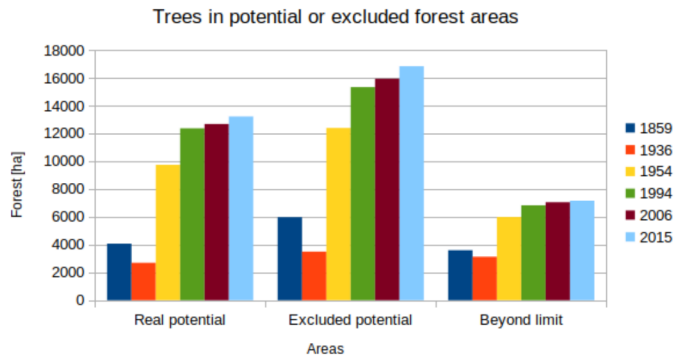


Figure 61: The evolution of forest coverage through the years in the potential and excluded areas

The data shows a similar behaviour to what explained for the general forest coverage and the data unbundled in altitude ranges. Forest are present in those areas beyond the limit and in areas excluded from the potential expansion and even this classification of forest areas underlines an general afforestation in PAT.

The evolution of the forest types areas in PAT is reported in table 43. This analysis was carried out by overlaying the maps of forest extracted for the different years to the maps of forest types distributed by the PAT territory (figure 14). A strong hypothesis was made in this analysis: forest typologies had not changed in the last 60 years. From this analysis were excluded the maps of 1936 and 1859. Those maps were already simplified by the hypothesis of distributing trees accordingly to the situation of 1954 and the time difference between those maps and the forest types maps is over 80 years, therefore the information coming from these maps were not considered reliable enough to perform an analysis of the changes in forest class area.

Forest class maps are based on a map of the 2006 forested areas, described in 2.2.5. The forest class maps considers the public forests, and some parcels may be excluded from the analysis, because they are private property.

Forest type	1954	1994	2006	2015
Holm Oak forests	514.51	872.32	1071.73	953.54
Oak forests	15805.75	22542.19	24164.16	24109.33
Mixed Oak-Hornbeam forests	547.12	930.81	1015.78	972.61
Chestnut and Black Locust forests	4280.99	6379.68	6515.53	6853.29
Maple, Linden, Ash forests	769.12	1011.95	1050.11	1094.18
Pioneer formations	1318.37	1864.35	2128.53	2086.19
Scots Pine forests	21385.08	29892.91	31409.26	30822.75
Beech forests	36429.12	46371.23	47566.89	49137.42
Fir forests	38687.61	42402.74	42904.92	42525.35
Mountain Pine forests	7334.41	8819.65	9613.57	9927.04
Alder forests	4884.51	6318.89	6757.43	7267.63
Spruce forests	89171.95	113228.98	113023.17	112009.64
Larch forests	33490.76	43913.4	44794.49	45309.98
Larch and Stone Pine forests	5553.57	8370.22	8028.72	7935.38
Stone Pine forests	451.31	851.63	818.65	758.14

Table 43: Forest coverage for forest classes

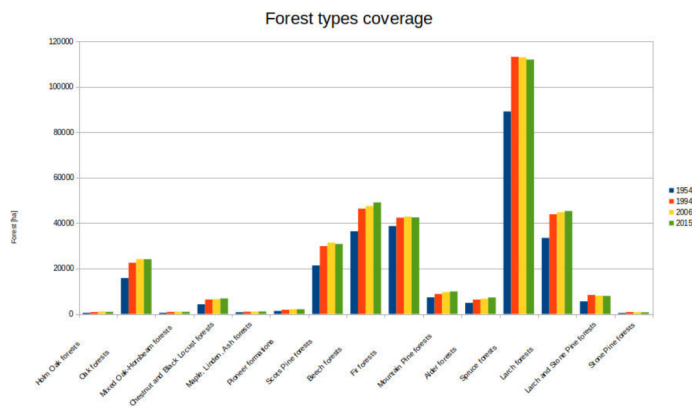


Figure 62: The evolution of forest class coverage

The tendency was of expansion for the majority of the forest classes. From 2006 to 2015 some classes seem to have slightly decreased in their extension. This marginal effect can be consequence of the lower accuracy on the classification of forest area in 2015, as reported by table 38. On the other hand a notable increase in extension could be seen in Beech forest. This was the result of a policy promoted by the PAT administration in order to preserve the presence of Beeches in the PAT forests. While some forest classes had decreased their extension, the forest as a whole increased in extension, because there was an increase as well in the areas of private and not publicly owned forest which were excluded from the map of forest classes.

The results here reported suggested that the afforestation of the PAT territory over the past 160 year was a process which has involved homogeneously all the forest types and all the altitude ranges.

### 3.4 QUANTITATIVE ASSESSMENT OF THE FOREST LANDSCAPE: A LANDSCAPE METRICS ANALYSIS

Landscape metrics can help to qualify the process of afforestation explained in section 3.3. It is worth reminding that for all the data landscape metrics calculation was performed by homogeneously setting the resolution of all the forest maps to 10x10m pixels. This choice allowed to perform a comparison of the calculated metrics between all the datasets, avoiding the effect on metrics that different resolution (as shown in table 2) can set. It is possible to accept this resampling by the hypothesis that a single tree remotely sensed easily covers an area of 10x10m, and that in every dataset the minimum dimension of the object was greater than 10x10m, as resumed in table 4. The metric calculation was performed using a 8 cell neighbour rule [126].

It is worth reminding that metrics were not supposed to give an exact result on the composition of a landscape, because a number of different factors concur in the creation of the land use maps and the evaluation of metrics such, as map resolution and neighbour size and shape. Instead under the hypothesis that land use maps of the same area were calculated with an homogeneous procedure and that those maps share the same resolution then metrics can be used to set a tendency in the way landscape evolved. Table 44 and graphs 63-68 reports on the metrics calculated for the forest in its whole area for PAT territory. In the evaluation of the metrics, with the caveats described earlier in the text regarding the datasets of 1859 and 1936, were considered those datasets as well.

Metric	1859	1936	1954	1994	2006	2015
Aggregation index [-]	88.99	88.85	90.14	94.55	94.77	94.55
Mean patch area [ha]	1.87	1.76	1.77	4.95	7.77	7.81
Largest Patch index [-]	5.04	8.09	8.79	47.44	49.67	50.52
Number of Patches [-]	107262	95990	153936	70137	45967	46020
Patch density [patches/ha]	53.60	56.97	56.51	20.19	12.88	12.81
Te [km]	75507.48	64481.40	107591.17	75910.77	74947.41	78588.31

Table 44: Evolution of the landscape metrics calculated for the whole PAT forest

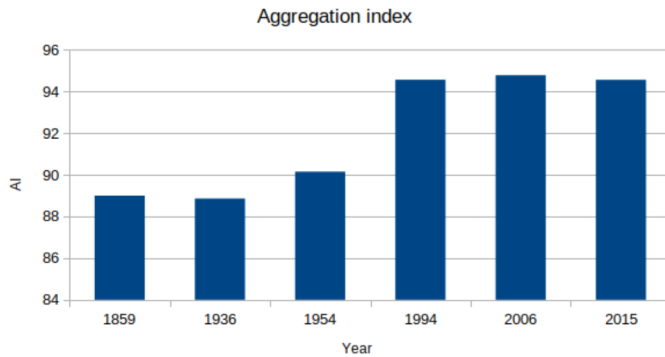


Figure 63: Evolution of the aggregation index for the total forest in PAT

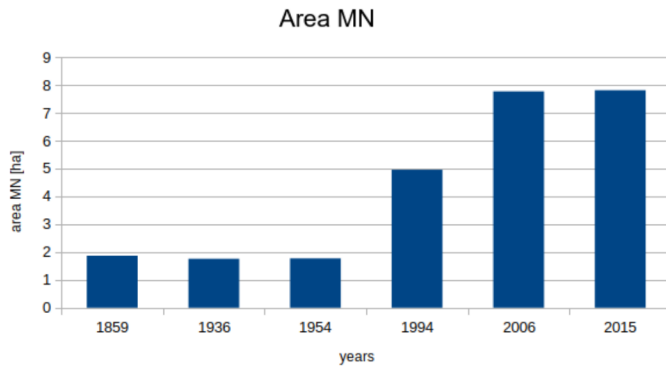


Figure 64: Evolution of the mean patch area for the total forest in PAT

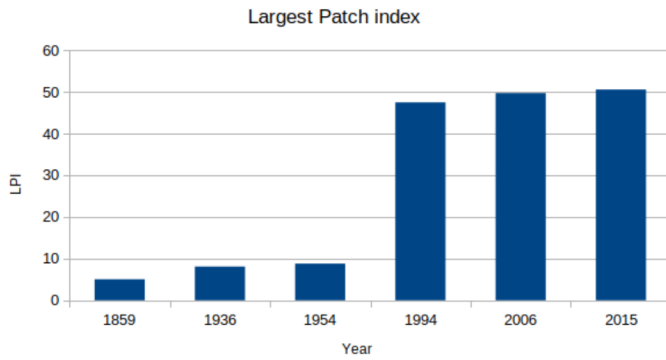


Figure 65: Evolution of the largest patch index for the total forest in PAT

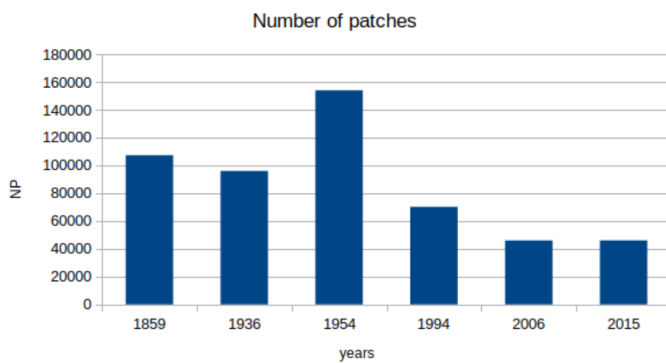


Figure 66: Evolution of the number of patches for the total forest in PAT

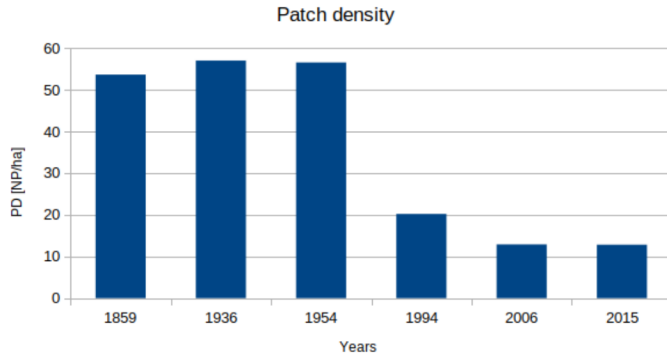


Figure 67: Evolution of the patch density for the total forest in PAT

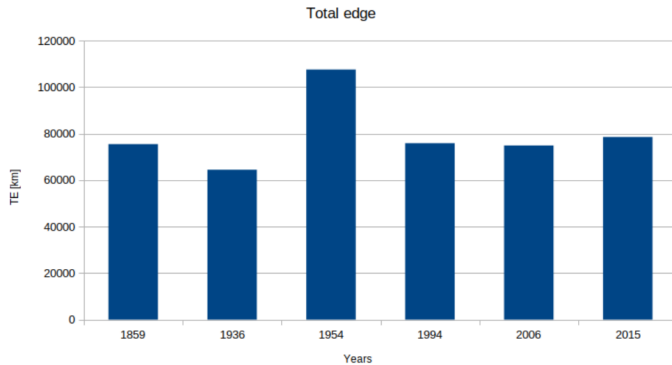


Figure 68: Evolution of the total edge for the total forest in PAT

From the year 1859 to the year 1954 there was an increase in NP, while Area MN, LPI, PD and AI roughly remained the same. This is an effect of the hypothesis made on the distribution of trees of the datasets of 1859 and 1936, but it could indicate that the forest expanded sparsely from 1859 to 1954, while from this year onward the forest distribution changed. NP decreased drastically, along PD, indicating a different evolution of the forests. Area MN and LPI increased in their value, indicating that trees gradually started to form a more compact forest, closing the open areas. The increase in AI confirms this compaction of the total forest area. This effect is visible even in the TE which decreased from 1954 onwards. This indicates a decrease of ecotonal zones. The afforestation in PAT meant a more compact forest, which has a potentially negative effect on the biodiversity in the mountain landscape of PAT.

The following table and graphs report on the calculus of landscape metrics of forests divided per altitude ranges. This reclassification of forest areas was made to underline the effect of the afforestation process in the different parts of the altitude gradient.

Alt range [mamsl]	Metric	1859	1936	1954	1994	2006	2015
0-600	Aggregation index [-]	88.02	88.74	86.43	90.44	91.72	90.59
601-1200		91.42	91.45	90.56	94.46	95.06	94.98
1201-2200		90.57	90.29	89.78	94.53	94.39	94.18
2201-3000		74.10	60.22	69.61	76.95	77.27	78.62
0-600	Mean patch area [ha]	1.50	1.65	1.16	1.83	3.37	2.96
601-1200		2.69	2.49	1.99	5.64	8.77	11.65
1201-2200		2.27	2.09	1.72	5.41	7.78	7.33
2201-3000		0.18	0.08	0.14	0.31	0.38	0.41
0-600	Largest Patch index [-]	0.49	0.35	0.60	1.07	1.33	0.76
601-1200		1.93	2.05	1.52	3.72	6.04	4.65
1201-2200		1.85	2.13	5.04	7.66	6.88	7.91
2201-3000		0.00	0.00	0.01	0.01	0.01	0.01
0-600	Number of Patches [-]	8704	6349	15949	14588	8740	9511
601-1200		28872	27012	47431	21311	14232	10837
1201-2200		53698	48812	92138	36932	25934	27786
2201-3000		600	341	7848	2916	2559	2759
0-600	Patch density [patches/ha]	4.10	3.53	5.85	4.20	2.45	2.65
601-1200		13.59	15.03	17.41	6.13	3.99	3.02
1201-2200		25.27	27.17	33.82	10.63	7.26	7.73
2201-3000		0.28	0.19	2.88	0.84	0.72	0.77
-	Shannon index	0.86	0.86	0.89	0.90	0.91	0.91
-	Simpson Index	0.53	0.53	0.54	0.54	0.55	0.55

Table 45: Calculated landscape metrics for the forest divided by altitude belts.

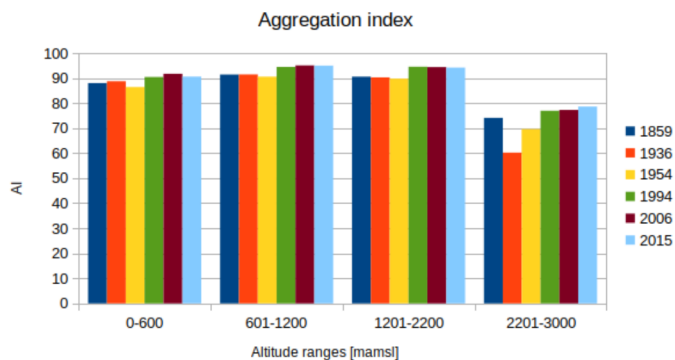


Figure 69: Evolution of the aggregation index divided by altitude ranges

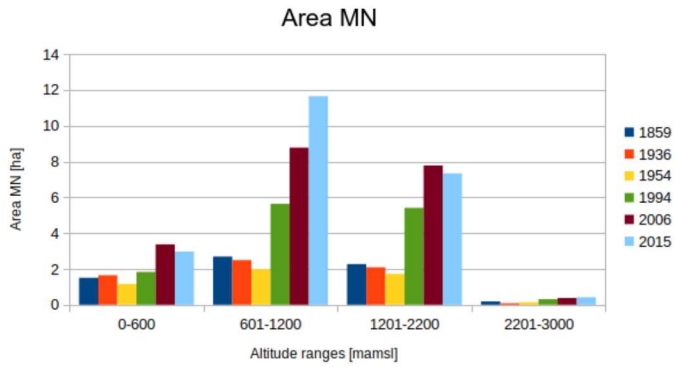


Figure 70: Evolution of the mean patch area divided by altitude ranges

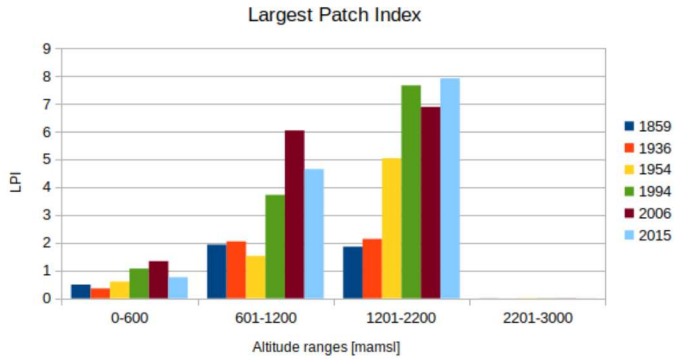


Figure 71: Evolution of the largest patch index divided by altitude ranges

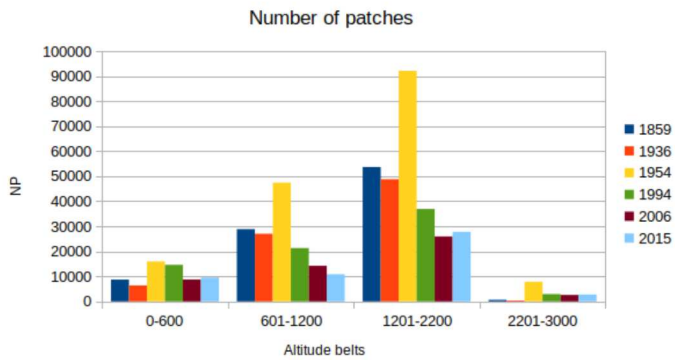


Figure 72: Evolution of the number of patches divided by altitude ranges



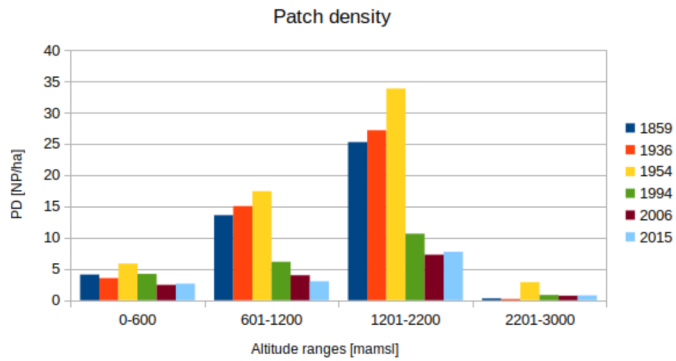


Figure 73: Evolution of the patch density divided by altitude ranges

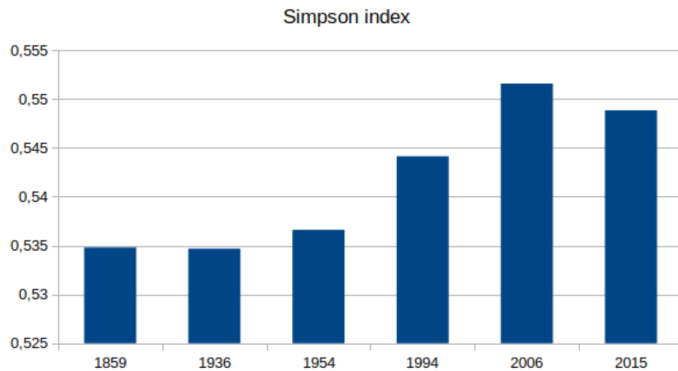


Figure 74: Evolution of the Simpson Index for forest divided by altitude ranges

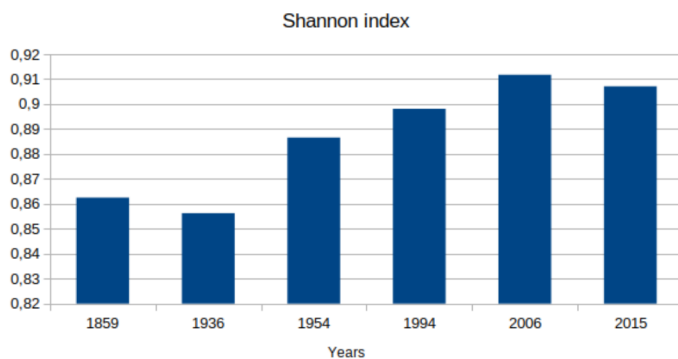


Figure 75: Evolution of the Shannon Index for forest divided by altitude ranges

Even in this case for all the altitude ranges the evolution of the forest followed the trend of the forest in its whole. Compaction of the forests homogeneously involved all the altitude ranges. The TE metrics was not calculated for this case, because the effect of the forest re-classification in altitude ranges is non negligible for this metric and could lead easily to a

misinterpretation. The SIMPSON and SHANNON indexes are indexes of the biodiversity, in this case intended as the diffusion of the forest at higher altitude. The two indexes are concordantly increasing from 1954 toward 2015, indicating a widespread diffusion of the forests in PAT in all the altitude ranges.

These set of metrics represent a significant results because they indicate that in the process of afforestation there was not a preferred altitude range where forest invaded the open areas.

The following tables and graphs show the evaluation of the landscape metrics for the forest reclassified in classes in the different years. Similarly to what explained in section 3.3, the reclassification of the forest area in classes was not performed for the dataset of 1859 and 1936.

Forest class	Metric	1954	1994	2006	2015
Holm Oak forests	Aggregation Index [-]	77.76	81.76	84.01	85.39
Oak forests		85.00	88.47	89.69	91.82
Mixed Oak-Hornbeam forests		77.08	80.73	81.08	82.99
Chestnut and Black Locust forests		81.71	88.45	88.07	91.31
Maple, Linden, Ash forests		78.45	82.22	81.57	84.94
Pioneer formations		73.07	79.43	79.21	82.59
Scots Pine forests		84.70	89.64	90.85	92.61
Beech forests		89.35	92.76	93.18	95.16
Fir forests		92.45	94.64	95.19	96.04
Mountain Pine forests		83.73	85.76	87.39	89.35
Alder forests		79.57	82.27	83.63	87.10
Spruce forests		86.37	91.61	91.37	93.35
Larch forests		85.01	90.08	90.56	92.63
Larch and Stone Pine forests		80.15	86.99	86.17	88.67
Stone Pine forests		75.59	84.08	83.64	85.13
Holm Oak forests		Mean patch area [ha]	0.38	0.51	0.78
Oak forests	0.94		1.59	2.31	4.12
Mixed Oak-Hornbeam forests	0.29		0.42	0.46	0.60
Chestnut and Black Locust forests	0.60		1.49	1.61	3.07
Maple, Linden, Ash forests	0.34		0.49	0.49	0.71
Pioneer formations	0.23		0.38	0.43	0.63
Scots Pine forests	0.93		2.48	3.25	5.19
Beech forests	1.83		3.91	4.78	9.76
Fir forests	4.42		12.04	13.76	21.70
Mountain Pine forests	0.69		0.79	1.21	1.83
Alder forests	0.41		0.60	0.77	1.29
Spruce forests	1.31		3.96	3.98	7.06
Larch forests	0.83		2.09	2.33	3.81
Larch and Stone Pine forests	0.49		1.20	1.03	1.71
Stone Pine forests	0.30		0.75	0.73	1.10

Table 46: Calculated landscape metrics for the forest divided in forest classes.

Forest class	Metric	1954	1994	2006	2015	
Holm Oak forests	Largest Patch index [-]	0.04	0.06	0.07	0.06	
Oak forests		0.25	0.21	0.24	0.24	
Mixed Oak-Hornbeam forests		0.01	0.01	0.01	0.01	
Chestnut and Black Locust forests		0.18	0.22	0.24	0.24	
Maple, Linden, Ash forests		0.03	0.02	0.02	0.03	
Pioneer formations		0.02	0.02	0.02	0.02	
Scots Pine forests		0.21	0.26	0.28	0.40	
Beech forests		2.17	1.83	1.80	1.73	
Fir forests		0.52	0.43	0.42	0.40	
Mountain Pine forests		0.10	0.11	0.13	0.13	
Alder forests		0.04	0.04	0.05	0.07	
Spruce forests		0.86	1.45	1.38	2.33	
Larch forests		0.25	0.44	0.45	0.45	
Larch and Stone Pine forests		0.06	0.07	0.07	0.07	
Stone Pine forests		0.01	0.02	0.02	0.02	
Holm Oak forests		Number of Patches [-]	1164	1405	1115	917
Oak forests			15344	12854	9510	5852
Mixed Oak-Hornbeam forests			1618	1917	1875	1616
Chestnut and Black Locust forests			6617	3966	3778	2230
Maple, Linden, Ash forests	2072		1895	1934	1535	
Pioneer formations	4887		4181	4201	3321	
Scots Pine forests	21242		11108	8930	5935	
Beech forests	19351		11486	9623	5037	
Fir forests	8482		3412	3021	1960	
Mountain Pine forests	9593		10188	7340	5436	
Alder forests	10596		9645	8035	5641	
Spruce forests	63191		26610	26505	15856	
Larch forests	37855		19764	18170	11888	
Larch and Stone Pine forests	10177		6137	6988	4647	
Stone Pine forests	1277		946	958	688	

Table 47: Calculated landscape metrics for the forest divided in forest classes.

Forest class	Metric	1954	1994	2006	2015	
Holm Oak forests	Patch density [patches/ha]	0.48	0.45	0.35	0.27	
Oak forests		6.28	4.11	2.97	1.71	
Mixed Oak-Hornbeam forests		0.66	0.61	0.59	0.47	
Chestnut and Black Locust forests		2.71	1.27	1.18	0.65	
Maple, Linden, Ash forests		0.85	0.61	0.60	0.45	
Pioneer formations		2.00	1.34	1.31	0.97	
Scots Pine forests		8.69	3.56	2.79	1.74	
Beech forests		7.92	3.68	3.01	1.47	
Fir forests		3.47	1.09	0.94	0.57	
Mountain Pine forests		3.92	3.26	2.30	1.59	
Alder forests		4.33	3.09	2.51	1.65	
Spruce forests		25.85	8.52	8.29	4.64	
Larch forests		15.48	6.33	5.68	3.48	
Larch and Stone Pine forests		4.16	1.96	2.19	1.36	
Stone Pine forests		0.52	0.30	0.30	0.20	
-		Shannon index	1.96	1.99	2.01	2.02
-		Simpson index	0.81	0.82	0.82	0.82

Table 48: Calculated landscape metrics for the forest divided by forest classes.

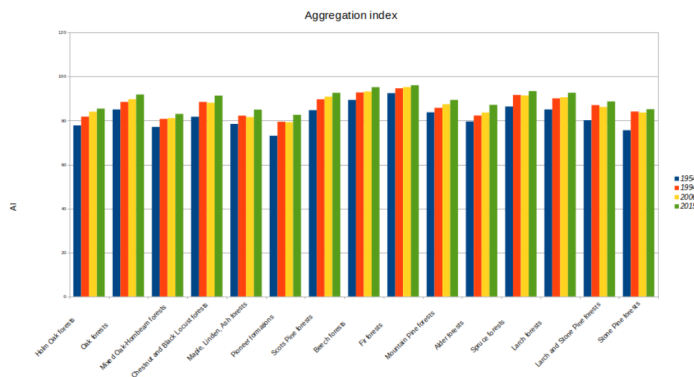


Figure 76: Evolution of the aggregation index divided by forest classes

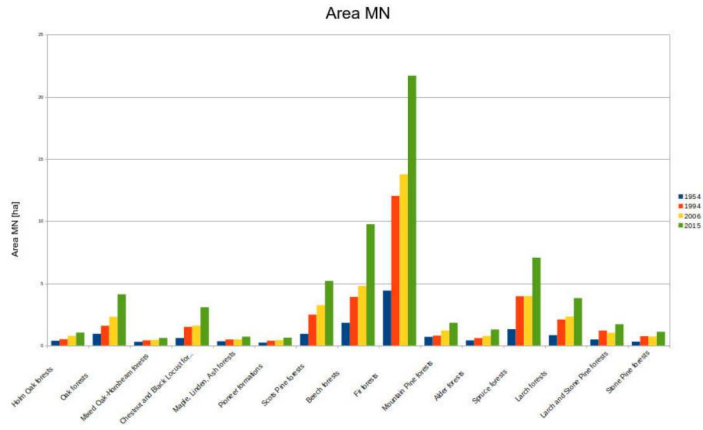


Figure 77: Evolution of the mean patch area divided by forest classes

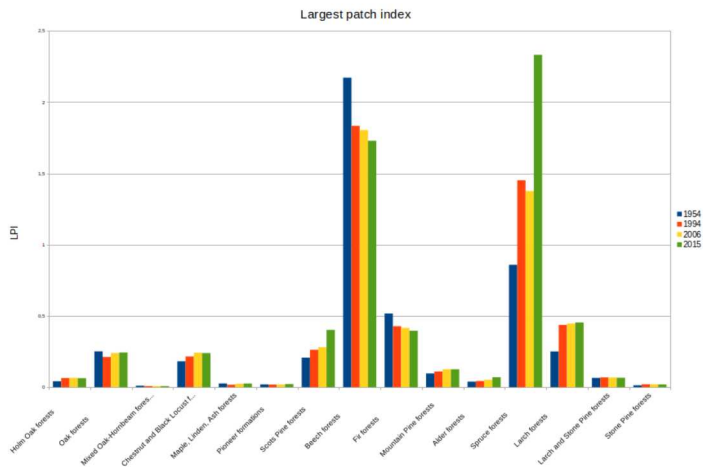


Figure 78: Evolution of the largest patch index divided by forest classes

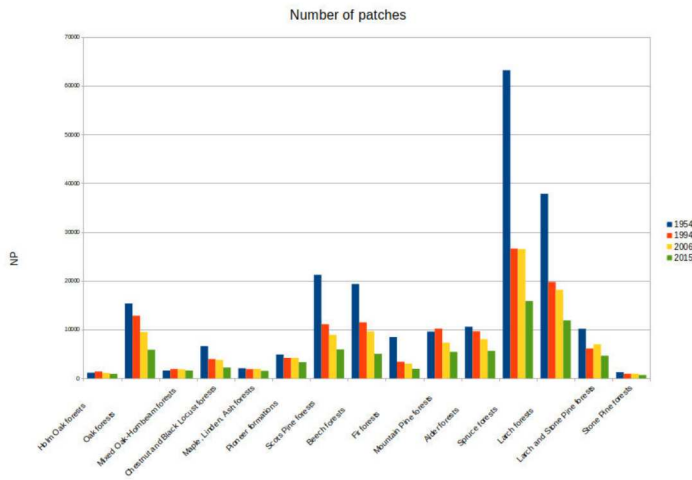


Figure 79: Evolution of the number of patches divided by forest classes

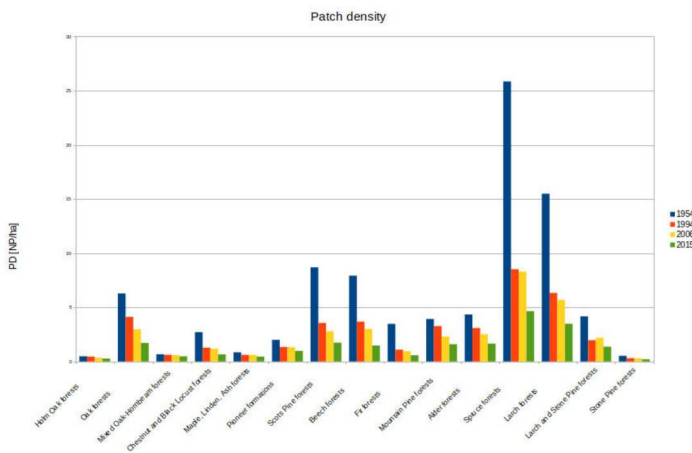


Figure 80: Evolution of the patch density divided by forest classes

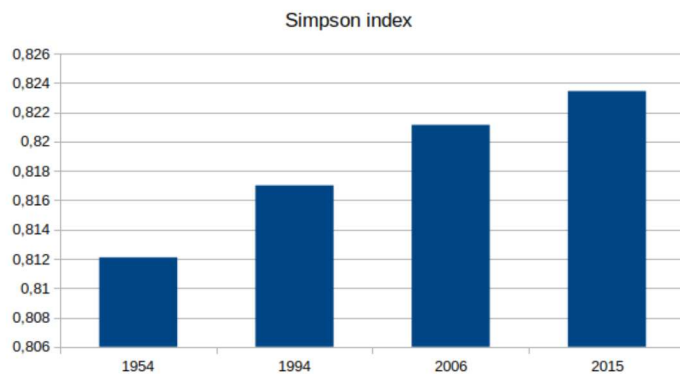


Figure 81: Evolution of the Simpson Index for forest divided by forest classes

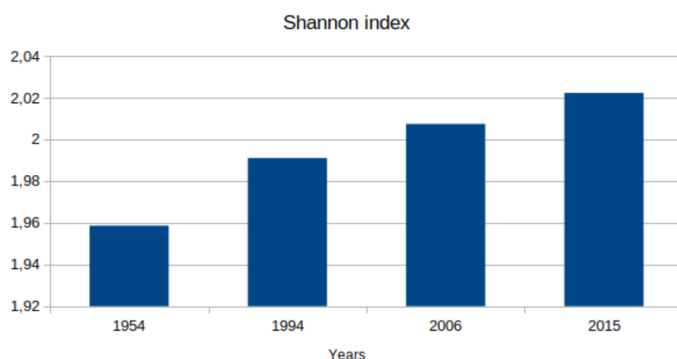


Figure 82: Evolution of the Shannon Index for forest divided by forest classes

Even in this case the compaction and expansion of forest was a detectable process, which involved all the forest types. The dynamics that emerged from the graphs and tables are still similar to the whole forest expansion, a diminishing in NP along with a increase of mean area and LPI. Beech forests and Fir forests, whose LPI decreases from 1954 toward 2015 presented this as outlier data. This unique behaviour is still unclear, because it is an isolated tendency not supported by any other metric in the analysis. This effect could be bound to a lack of accuracy in the forest map, which in some areas interested by these forest types could have led to a over-detection of forest areas, which contributed to the identification of bigger areas of beech and fir forest than they actually were. As the accuracy and the quality of the classification improved in the datasets, the size of the largest patch became more similar to the expected value, thus this effect of over-detection disappeared, with the effect of a decrease in LPI.

Even in this case SIMPSON and SHANNON increased from 1954 toward 2015, an index of that the afforestation did not interest only a particular forest class, but each single forest class. In fact, if a forest type would prevail over the other types these two indexes would show a slight decrease, or would remain constant.

Aggregating the forest data in forest classes shows again a process of afforestation which involved all the forest types. Aggregating and disaggregating the data about forest coverage show no significant differences in the dynamics of afforestation which appears in each forest class and each altitude range. The calculation of metrics, along with the data about forest extension, underlined a process which involved the closure of open areas, instead of a distribution of forests in 'hotspots' where the trees expanded, and other areas which were excluded from the expansion of forests.

### 3.5 *relevance of the cell neighborhood size in landscape metrics evaluation and free or open source software implementations*

The following paper discusses on the relevance of a basic input parameter used for calculating landscape metrics, the shape of the cell neighbourhood (CN from now on). CN can be of two different type:

- Von Neumann, or the neighbourhood is considered as the 4 orthogonal cells of a central one;
- Moore, or the cell neighbourhood is considered as the 8 closest cells of a central one.

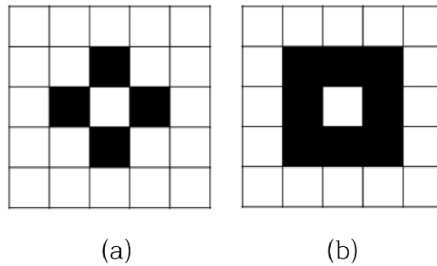


Figure 83: Cell neighbourhood shape, (a) Von Neumann, (b) Moore.

The metrics calculated in section 3.4 were evaluated using a 8 CN rule, but during the current study, by applying landscape metrics calculation to a landscape, different results were yielded depending on the software used for the calculation of the metrics. This discrepancies depended on how each single program considers the shape of the cell neighbourhood and the possibility of choosing this parameter by the users. In the current literature this parameter is seldom declared, this makes the comparison between different works difficult to explain. In the paper a literature review was carried out to find, on a sample of other researches how many openly stated the CN was used. Finally, the paper compares the differences between a same metric calculated on the same landscape using different CN with the Aggregation Index, as a measure of the aggregation of a landscape independent from the CN parameter. The paper was published in ISPRS International Journal of Geo-Information, volume 8, in 2019 [126].



Article

# Relevance of the Cell Neighborhood Size in Landscape Metrics Evaluation and Free or Open Source Software Implementations

Paolo Zatelli <sup>1,\*</sup>, Stefano Gobbi <sup>1,2,3</sup>, Clara Tattoni <sup>1,4</sup>, Maria Giulia Cantiani <sup>1</sup>,  
Nicola La Porta <sup>2,3</sup>, Duccio Rocchini <sup>5,6,7,8</sup>, Nicola Zorzi <sup>1</sup> and Marco Ciolli <sup>1,5</sup>

<sup>1</sup> Department of Civil, Environmental and Mechanical Engineering, University of Trento, Via Mesiano 77, 38123 Trento, Italy; stefano.gobbi@unitn.it (S.G.); clara.tattoni@unitn.it (C.T.); maria.cantiani@unitn.it (M.G.C.); nicomede83@gmail.com (N.Z.); marco.ciolli@unitn.it (M.C.)

<sup>2</sup> Fondazione Edmund Mach, Research and Innovation Centre, Centro di Ricerca e Innovazione, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy; nicola.laporta@fmach.it

<sup>3</sup> The EFI Project Centre on Mountain Forests (MOUNTFOR), Edmund Mach Foundation, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy

<sup>4</sup> Dipartimento di Scienze e Tecnologie Agrarie, Alimentari, Ambientali e Forestali, University of Florence, Piazzale delle Cascine, 18, 50144 Firenze, Italy

<sup>5</sup> C3A, Center Agriculture Food Environment, University of Trento, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy; duccio.rocchini@unitn.it

<sup>6</sup> Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, Via Sommarive, 9, 38123 Povo (TN), Italy

<sup>7</sup> Fondazione Edmund Mach, Research and Innovation Centre, Department of Biodiversity and Molecular Ecology, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy

<sup>8</sup> Czech University of Life Sciences Prague, Faculty of Environmental Sciences, Department of Applied Geoinformatics and Spatial Planning, Kamýcká 129, 16500 Praha-Suchdol, Czech Republic

\* Correspondence: paolo.zatelli@unitn.it

Received: 24 October 2019; Accepted: 9 December 2019; Published: 13 December 2019



**Abstract:** Landscape metrics constitute one of the main tools for the study of the changes of the landscape and of the ecological structure of a region. The most popular software for landscape metrics evaluation is FRAGSTATS, which is free to use but does not have free or open source software (FOSS). Therefore, FOSS implementations, such as QGIS's LecoS plugin and GRASS' r.li modules suite, were developed. While metrics are defined in the same way, the "cell neighborhood" parameter, specifying the configuration of the moving window used for the analysis, is managed differently: FRAGSTATS can use values of 4 or 8 (8 is default), LecoS uses 8 and r.li 4. Tests were performed to evaluate the landscape metrics variability depending on the "cell neighborhood" values: some metrics, such as "edge density" and "landscape shape index", do not change, other, for example "patch number", "patch density", and "mean patch area", vary up to 100% for real maps and 500% for maps built to highlight this variation. A review of the scientific literature was carried out to check how often the value of the "cell neighborhood" parameter is explicitly declared. A method based on the "aggregation index" is proposed to estimate the effect of the uncertainty on the "cell neighborhood" parameter on landscape metrics for different maps.

**Keywords:** landscape metric; landscape ecology; forest landscape; free and open source geospatial; map filtering

## 1. Introduction

In recent decades, landscape metrics were one of the main tools for the study, quantification, and possibly the parametrization of the changes of the landscape and of the ecological structure of a territory [1,2]. Landscapes consist of natural or anthropogenically modified mosaics of patches, and the spatial arrangement of these patches is called landscape pattern [3]. The quantification of spatial heterogeneity over multiple spatial and temporal scales is necessary to clarify the relationships between ecological processes and spatial patterns [4,5]. Landscape metrics became standard and gained the status of an indispensable instrument for those who investigate landscape change dynamics not only in ecology [6,7] but also in many other disciplines such as soil protection and water management [8–10]. Many are metrics that can be calculated, but some of them are considered particularly robust and significant and are used more frequently in scientific literature [2,6]. It is a common practice to compare the different landscapes according to the results of some of these indices like, for example, patch number (NP) or mean patch size (MPS) [11]. However, we argue that the results of some of the most used and reliable indices can be severely affected by the settings used during the processing [11] and, if these settings are not accurately described in the scientific works, this can create a cascade effect that affects the comparability of the results. In particular, we realized that the cell neighborhood size is one of the most sensible parameters but a literature analysis highlighted that not many works have investigated its influence. Additionally, each software provides information about the setting of cell neighborhood size with different emphasis.

The most popular software for landscape metrics evaluation is FRAGSTATS [12], which is free to use but not free or open source software (FOSS). Therefore, FOSS implementations, such as QGIS's LecoS plugin and GRASS' *r.li* modules suite, were developed. Landscape metrics are defined in the same way in all these software. However, the cell neighborhood parameter, specifying the configuration of the moving window used for the analysis, is managed differently: FRAGSTATS can use values of 4 or 8 (8 is default), LecoS uses 8 and *r.li* 4.

Are the software users aware of which cell neighborhood size they are using and what the performances and the effects it can produce are? How reliable are the comparisons among different studies that do not mention the cell neighborhood size settings? Is it possible to safely compare the results of works carried out with different software?

In order to answer to these questions the aims of our work are: (i) to investigate how the cell neighborhood size affects the results of landscape metrics calculation using both tests map and real maps; (ii) to examine the transparency of the settings of the cell neighborhood size in different software; (iii) to understand if the cell neighborhood size settings are usually reported in the scientific papers and, therefore, how reliable are the comparisons of different works developed with the same or different software.

The paper is organized as follows: Section 2 illustrates the literature review and its results; Section 3 describes the materials and methods used for the analysis, the software, the metrics, and the maps used for the tests; Section 4 illustrates the results; Section 5 describes a new method to evaluate the sensitivity of landscape metrics to the cell neighborhood size and finally the conclusions and future developments are presented in Section 6.

## 2. Literature Review

An analysis of the scientific literature for papers regarding the use of FRAGSTATS, the *LecoS* add-on, or the *r.li/r.le* modules in GRASS GIS was carried out to determine if the cell neighborhood parameter was mentioned or explained in scientific papers describing applications using landscape metrics.

The Scopus database was used, searching for the term "cell neighborhood" coupled with "Fragstats", "r.li", "r.le", "LecoS", "landscape patch", and "landscape metrics" as keywords. The joint use of these terms yielded no results, so the keywords have been used separately.

The papers for the current review were chosen following these conditions: (i) the articles must be research papers, peer reviewed, although with no minimum limit on the number of citations; (ii) at

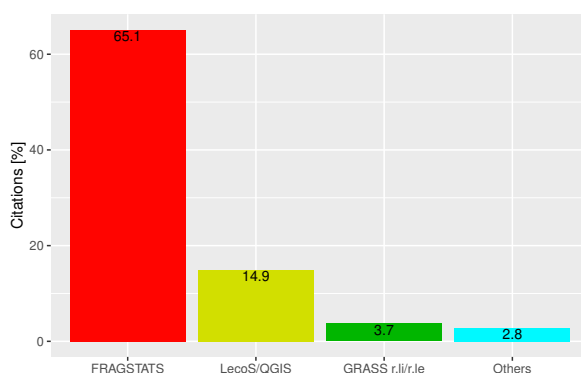
least one of the FRAGSTATS, LecoS, or *r.li/r.le* software must have been used in the study; (iii) more than one landscape metric must have been used; (iv) no limit has been set for the year of publication.

In the Scopus archive the term “landscape metrics” produces an excess of 37,000 results. Refining the search by adding the keyword “Fragstat” narrows the result to 800 papers. Adding the terms “lecos” and “landscape metrics” refines the result to six papers; finally, adding “landscape metrics” with “r.li” refines the result to four papers.

For further analysis the pool of results was narrowed to 93 papers:

- 70 (65.1%) papers cited FRAGSTATS as software for the evaluation of landscape metrics;
- 16 (14.9%) papers cited LecoS as QGIS add-on used for the calculation of the landscape metrics;
- Four (3.7%) papers cited *r.li/r.le* as GRASS modules for the assessment of the metrics;
- Three (2.8%) papers cited other software for the assessment of the metrics;
- Two papers developed new software for the evaluation of landscape metrics, using as reference for the calculations FRAGSTATS (these papers have been counted in the 70 articles citing FRAGSTATS).

Figure 1 shows the distribution of the software used for landscape analysis in the examined papers.



**Figure 1.** Comparison of the times each landscape analysis software is cited in the analyzed papers.

These papers can be classified according to their purpose:

- papers that deal with landscape fragmentation as an issue in itself;
- papers that deal with landscape fragmentation as an issue for habitat connectivity;
- papers that investigate which landscape metric could explain other metrics;
- papers that investigate the size of a moving window algorithm and how it affects landscape analysis;
- papers that present new software for landscape metrics calculation, basing their structure and methods on FRAGSTATS.

Table 1 shows the most used landscape metrics, with the numbers of times the selected metric was used in the reviewed papers as absolute values and percentages.

**Table 1.** Number and percentage of papers in the analyzed set citing landscape metrics.

Metric	Number of Citing Papers	Percentage of Citing Papers
Patch Number	47	51.1%
Patch density	31	33.7%
Mean patch size	51	55.4%
Edge density	34	37.0%
Landscape shape index	32	34.8%
Aggregation index	21	22.8%

Additionally, the number of times nine other metrics are indicated in the set of papers was counted: Largest patch index (34, 37%), Shannon Index (35, 38%), Simpson Index (18, 19.6%), Percentage of landscape (28, 30.4%), Perimeter/Area ratio (20, 21.7%), Euclidean distance (14, 15%), Total edge (27, 29.3%), Shape index (32, 34.8%), Fractal index (22, 23.9%), and Mean nearest neighbor (25, 27.2%).

The value of the cell neighborhood parameter was specified in eight different papers: seven chose a Moore definition [13–20] and one a *von Neumann* definition [21]. The 16 papers which exclusively employed LecoS [22–37] used an 8-cells neighborhood rule, as it is the only setting available in the add-on; a similar hypothesis is applicable to the four works which used *r.li* in GRASS [38–41], which uses a 4-cells neighborhood definition. The papers which did not specify the definition of the neighborhood in FRAGSTATS are assumed to have used a 8-cells definition, since this is the default setting [42], although it is not possible to prove this assumption. Results in Section 4 demonstrate that some of the most commonly used metrics are affected by the choice of this parameter. Therefore, the fact that the cell neighborhood configuration is not explicitly or implicitly specified in most of the research papers analyzed (by the indication of the software used) could make the results difficult to repeat and interpret.

The full list of the examined papers, a table indicating the papers citing each software, and a breakdown of the landscape metrics cited by each paper is available as an additional material of this paper (see Supplementary Materials).

### 3. Materials and Methods

To analyze the effect of the cell neighborhood size on the values of landscape metrics calculation, three different geographic information system (GIS) software were selected (Section 3.1) to evaluate a common set of metrics (Section 3.2) on artificial (Section 3.3) and real (Section 3.4) maps.

#### 3.1. Test Software

Tests on artificial (Section 3.3) and real (Section 3.4) maps were carried out using three different software systems: FRAGSTATS [12], GRASS GIS [43], and QGIS [44]. The rationale behind this choice is that using software in the public domain (FRAGSTATS) or available as free and open software (FOSS) (GRASS GIS and QGIS) allows the replication of the experiments described in this paper. Additionally, FOSS source code is available for analysis. Therefore, possible discrepancies between metrics' values can be resolved examining the code.

According to its manual [45], "FRAGSTATS is a spatial pattern analysis program for quantifying the structure (i.e., composition and configuration) of landscapes.". FRAGSTATS [12] was the reference software for the evaluation of landscape metrics since its inception in 1995. It is available as executable in the public domain for the MS Windows operative system only. Written in Microsoft Visual C++, the use of a 32-bit address space limits its capacity to process large maps. Its source code is not accessible but its features are well documented in the manual available on-line and its widespread use in landscape and ecology research speaks to its accuracy and reliability in metrics evaluation. Units of input maps are assumed to be meters. Input map can only contain signed integer values and have square cells with sides larger than  $10^{-3}$  m. FRAGSTATS can evaluate 36 *area and edge metrics*, 79 *shape metrics*, 46 *core area metrics*, 17 *contrast metrics*, and 64 *aggregation metrics*, for a total of 242 different metrics and 392 parameters describing them [45]. Additionally, FRAGSTATS can evaluate 9 *Diversity Metrics*.

GRASS GIS [43] is a multi-purpose FOSS GIS for geospatial data production, analysis, and mapping [46] available under the GNU Public License (GPL), used in research and education [47]. It is part of the Open Source Geospatial Foundation (OSGeo). GRASS is mainly written in ANSI C, with some parts in the C++ and, more recently, in the Python programming languages; it is highly modular and easily scriptable. Besides the source code, it is available as ready to install packages for the MS Windows, Apple Mac OSX, and Linux operating systems. Limitations on data size are mostly related to the capacity of the operative system to deal with large files [48], with the possibility of enabling large file support (LFS). Landscape analysis in GRASS is carried out using the *r.li* modules

suite, which can evaluate 10 *patch indices* and 7 *diversity indices* and provides a graphical user interface for the configuration of the analysis. A previous implementation of landscape metrics evaluation, called *r.le* is used in old GRASS versions.

QGIS is a user friendly open source geographic information system (GIS) licensed under the GNU General Public License [44], part of the Open Source Geospatial Foundation (OSGeo). It is written in the C++ language using the Qt toolkit for its interface. Plugins can be written either in C++ or Python. It is available as source code and packages for the Linux, Unix, Mac OSX, MS Windows, and Android operative systems. The maximum size for maps depends on the combination of operative system and Geospatial Data Abstraction Library (GDAL)/OGR Simple Features Library (OGR) limitations for the data type in use. The *Landscape ecology statistics (LecoS)* plugin performs landscape analysis in QGIS using the Python numpy library [49]. LecoS, like version 3.0.0, can evaluate 20 *landscape metrics*, 8 *zonal statistics* and 3 *diversity indexes*.

Other modules to calculate landscape metrics are integrated in the most popular GIS desktop software: V-Late and PA4 in ArcGIS (set of GIS programs created by the Environmental Systems Research Institute, ESRI), Pattern and Texture modules in IDRISI, and the API called Land-metrics DIY developed by Zaragozí et al. [50] and released under the GPL license. Recently, a new R package dedicated to calculation of landscape metrics was released; with *landscapemetrics* it is now possible to perform landscape analysis on rasters inside the R environment [51].

FRAGSTATS 4.2, GRASS 7.6, and QGIS 3.4 (LecoS 3.0.0) versions were used for the tests. A comprehensive list of the software features with respect to landscape metrics is reported in Appendix A.

### 3.2. Landscape Metrics

Landscape metrics are indexes used to describe and quantify the spatial characteristics of the landscape, described as a set of categorical data, typically representing land use [6]. These indexes are useful to study the evolution of the landscape features over time [7,41,52,53] or to compare different areas [6].

The great variety of metrics can be classified according to the spatial level they are dealing with: patches, classes of patches, or the whole landscape [54]. Commonly, in landscape ecology, a patch is defined as a discrete area of homogeneous environmental conditions at a specific scale [45]. These metrics fall into two categories: those that quantify the composition of the map without reference to spatial attributes (for example, the number of patches) and those that quantify the spatial configuration of the map, such as spatial arrangement, position, or orientation of the patches [54].

Landscape metrics calculated at the patch level describe simple statistics about each land use patch (area, perimeter, number of patches, ...) and serve primarily as the computational basis for other landscape metrics [45]. Landscape metrics calculated at the class level report information about all the patches of a given type. These metrics are useful to quantify the amount and spatial configuration of each patch type and thus are useful to estimate the fragmentation of each patch type in the landscape [45]. Finally, landscape metrics can provide information about the pattern, like class metrics: these may be integrated by a simple or weighted averaging or may reflect aggregate properties of the patch mosaic [45].

For the purpose of this work, a limited number of landscape metrics was selected among the many available according to the following criteria: (i) the availability for calculation across software; (ii) the use of the cell neighborhood parameter for their calculation; (iii) the relative simplicity of their formulation, in order to understand the effect of the cell neighbor parameter on the result and to be able to manually evaluate their values for simple maps.

**Patch Number (NP)** is the number of patches of each type; it is an adimensional metric. It is defined as:

$$NP = N_k, \quad (1)$$

where:

- $N$  is the number of patches in the  $k$ -th category.

**Patch density (PD)** equals the number of patches of the corresponding patch type divided by total landscape area [m], multiplied by 10,000 and 100 (to convert to 100 hectares). The unit of measure is number of patches per 100 hectares. It is expressed as:

$$PD = \frac{N}{A}(10^4)(100), \quad (2)$$

where:

- $N$  is the number of patches in the landscape;
- $A$  is the total landscape area in [m<sup>2</sup>];
- $10^4$  and 100 are constant and used to express the index in (100) ha.

**Mean patch size (MPS)** is the average area of all the patches of a given type. It is measured in square meters. When used in combination with NP, MPS gives information about how the patches of a given land use class are growing or merging over time [53]. It is defined as:

$$MPS = \frac{A}{NP}, \quad (3)$$

where:

- $A$  is the area of the patches in [m<sup>2</sup>];
- $NP$  is the number of patches.

**Edge density (ED)** equals the sum of the lengths (m) of all edge segments involving the corresponding patch type, divided by the total landscape area (m), multiplied by 10,000 (to convert to hectares). The unit is meters per hectare.  $ED > 0$ . This index is useful in ecological studies dealing with ecotone species. It is expressed as:

$$ED = \frac{\sum_{k=1}^m \sum_{i=1}^n e_{ik}}{A}(10^4), \quad (4)$$

where:

- $k$  is the category of the patches;
- $m$  is the total number of different categories of patches;
- $n$  is the number of boundary edges for the patch;
- $e_{ik}$  is the total length of boundary edges for the  $k$ -th category of patches;
- $A$  is the total landscape area;
- $10^4$  is a constant to convert the index in [m/ha].

**Landscape shape index (LSI)** measures the perimeter-to-area ratio for the landscape as a whole. All edge segments (m) within the landscape boundary involving the corresponding patch type are divided by the square root of the total landscape area (m).  $LSI > 1$ , adimensional. This index is a measure of the overall geometric complexity of the landscape and is defined as:

$$LSI = 0.25 \frac{E}{\sqrt{A}}, \quad (5)$$

where:

- $E$  is the sum of the lengths of all the boundary edges of the patches;
- $A$  is the sum of all the areas of the patches;
- 0.25 is an adjustment coefficient.

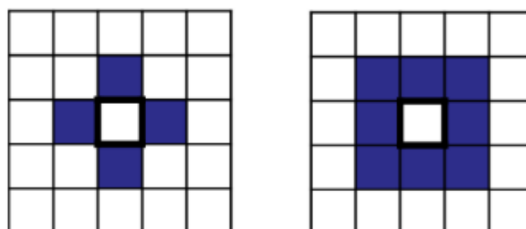
**Aggregation index (AI)** equals the number of adjacent patches involving the corresponding class, divided by the maximum possible number of like adjacencies, which is achieved when the class is maximally clumped into a single, compact patch, multiplied the proportion of the landscape comprised of the corresponding class, summed over all classes and multiplied by 100 (to convert to a percentage). Unit: Percent, range  $0 < AI < 100$  [45]. It is useful to quantify spatial patterns and fragmentation of the landscape. It is defined as:

$$AI = \left[ \frac{g_{ii}}{\max(g_{ii})} \right] (100), \quad (6)$$

where:

- $g_{ii}$  is the number of like adjacencies between pixels of patch class  $i$  based on the *single-count* method;
- $\max(g_{ii})$  is the maximum number of like adjacencies between pixels of patch class  $i$  based on the *single-count* method;
- 100 is a constant used to express the index as percentage.

Cell neighborhood is defined as the nearest cells to a specific one. Only cells in the neighborhood are used to evaluate the metrics above. There are two possible concepts of “cell neighborhood” (CN): the *von Neumann* definition, which considers the four adjacent cells in the cardinal directions, and the *Moore* one, which considers both the corners and edges of adjacent cells (Figure 2). While *von Neumann* neighborhood considers 4 cells, *Moore* neighborhood uses 8 cells for the calculations.



(a) *Von Neumann* neighborhood      (b) *Moore* neighborhood

**Figure 2.** Patterns of the *von Neumann* and *Moore* neighborhood definitions.

The different software systems for landscape analysis use both of these definitions:

- the *r.li/r.le* modules in GRASS GIS uses the *von Neumann* definition [55];
- the *LecoS* plugin for QGIS uses the *Moore* definition [49];
- FRAGSTATS allows the user to choose between the two definitions [42].

The evaluation of some landscape metrics is influenced by the choice of this parameter. Corner edge cells of the same land use class will be considered part of the same patch in a 8-cells neighborhood, while they are not considered such using a 4-cells neighborhood algorithm.

For this reason, using the same map in different software will lead to different results for some landscape metrics. While *LecoS*, *r.le*, and *r.li* allow the use of one single type of neighbors rule, in FRAGSTATS the default option is the 8-cells *Moore* definition but the user can select the 4-cells *von Neumann* one.

### 3.3. Artificial Maps

Four artificial maps were created to test the sensitivity of the different metrics to the variation of the *cell neighborhood* value and to understand which patches' configurations are more susceptible to different values of the *cell neighborhood* (CN) value. The four maps are binary, with value 1 indicating

the patches and value 0 the background. The maps are square, with 100 cell sides, for a total of 10,000 cells and each pixel measures  $10 \times 10$  map units.

The first map is configured as a chessboard (Figure 3), alternating cells with value 1 and 0. Since each pixel of value 1 represents a separate patch, the total number of patches (NP) is 5000, the patch density (PD), which is maximum for this configuration, is equal to 10,000 (number of patches/100 ha), the (uniform) mean patch size (MPS) is 0.01 (ha), the edge density (ED) is 4000 (m/ha), and the landscape shape index (LSI) is 70.71. This configuration is the one that maximizes the difference of outcomes for the different choices of CN values.

1	0	1	0	1	0	1	0	1	0
0	1	0	1	0	1	0	1	0	1
1	0	1	0	1	0	1	0	1	0
0	1	0	1	0	1	0	1	0	1
1	0	1	0	1	0	1	0	1	0
0	1	0	1	0	1	0	1	0	1
1	0	1	0	1	0	1	0	1	0
0	1	0	1	0	1	0	1	0	1
1	0	1	0	1	0	1	0	1	0
0	1	0	1	0	1	0	1	0	1
1	0	1	0	1	0	1	0	1	0
0	1	0	1	0	1	0	1	0	1

Figure 3. Scheme of the configuration of the first artificial test map, displayed on a chessboard.

The second map (Figure 4) contains two large patches covering the whole map, separated by a single column of 0 values. For this map the total number of patches (NP) is 2, the patch density (PD) is 2.02 (number of patches/100 ha), the mean patch size (MPS) is 49.5 [ha], the edge density (ED) is 60.4 [m/ha], and the landscape shape index (LSI) is 1.5. The presence of a single edge should result in the same landscape metrics values for both choices of CN.

1	1	1	1	1	0	1	1	1	1
1	1	1	1	1	0	1	1	1	1
1	1	1	1	1	0	1	1	1	1
1	1	1	1	1	0	1	1	1	1
1	1	1	1	1	0	1	1	1	1
1	1	1	1	1	0	1	1	1	1
1	1	1	1	1	0	1	1	1	1
1	1	1	1	1	0	1	1	1	1
1	1	1	1	1	0	1	1	1	1
1	1	1	1	1	0	1	1	1	1

Figure 4. Scheme of the configuration of the second artificial test map, consisting of two large patches separated by a single column of 0 values.

The third map (Figure 5) consists in a single, compact, large patch covering half of the map. For this map the total number of patches (NP) is 1, the patch density (PD) is 2 (number of patches/100 ha), the mean patch size (MPS) is 50 (ha), the edge density (ED) is 60 (m/ha) and the landscape shape index (LSI) is 1.06. This configuration is expected to provide the same results for any CN choice, since there is only one edge and the patch is larger than the area corresponding to the largest possible value of CN, 8.



1	1	1	1	1	0	0	0	0	0
1	1	1	1	1	0	0	0	0	0
1	1	1	1	1	0	0	0	0	0
1	1	1	1	1	0	0	0	0	0
1	1	1	1	1	0	0	0	0	0
1	1	1	1	1	0	0	0	0	0
1	1	1	1	1	0	0	0	0	0
1	1	1	1	1	0	0	0	0	0
1	1	1	1	1	0	0	0	0	0
1	1	1	1	1	0	0	0	0	0

**Figure 5.** Scheme of the configuration of the third artificial test map, consisting of a single large patch occupying the left half of the map.

The fourth test map (Figure 6) contains a single patch consisting of just one pixel at the center of the map. For this map the total number of patches (NP) is 1, the patch density (PD) is 10,000 (number of patches/100 ha), the mean patch size (MPS) is 0.01 (ha), the edge density (ED) is 4000 (m/ha) and the landscape shape index (LSI) is 1.

The only cell of value 1 is surrounded by 8 cells of value 0. Therefore, the two choices of CN should yield the same results for the landscape metrics.

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

**Figure 6.** Scheme of the configuration of the fourth artificial test map, consisting of a single patch of one pixel at the center of the map.

### 3.4. Real Maps

A map of the forest coverage in the Val di Fassa, Italy, in 2006 was used for the tests on real maps. Val di Fassa is a valley in the Alps, in the north east of the Trentino region in Northern Italy (Figure 7). This area was affected by a marked expansion of the forested area in the last century [6,41,51,52] and is currently being investigated in the Trentinoland project, which is analyzing the forest cover of the whole Trentino region in Northern Italy [56].

Val di Fassa has an area of 567.28 km<sup>2</sup> (56,728 ha), of which 132.36 km<sup>2</sup> (13,236 ha, 23.33%) were covered by forest in 2006. The region boundaries are 5157094 N, 5132029 S, 721983 E, 699351 W, in the ETRS89/UTM 32 N (EPSG 25832) datum. The raster map has a 10 m resolution, with 2509 rows and 2265 columns, for a total of 5,682,885 cells.

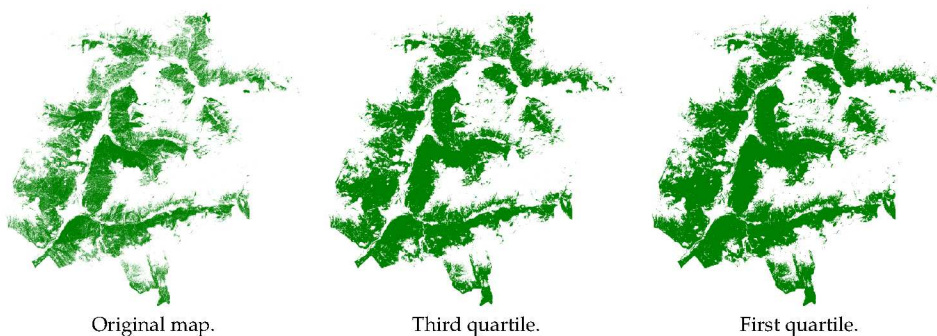
Two maps were derived from this map, by applying a low pass filter with a 3 × 3 pixels window, assigning the third quartile for the first map (Figure 8, center) and the first quartile for the second one (Figure 8, right), with the aim to test how the fragmentation of the maps influences the variation of the

landscape metrics depending on the CN value. The application of a low-pass filter obviously leads to less fragmented maps. These maps will be used to develop an index predicting how different values of CN influence landscape metrics values when CN is unknown in Section 5.

These two maps cover the same area and have the same resolution of the original one.

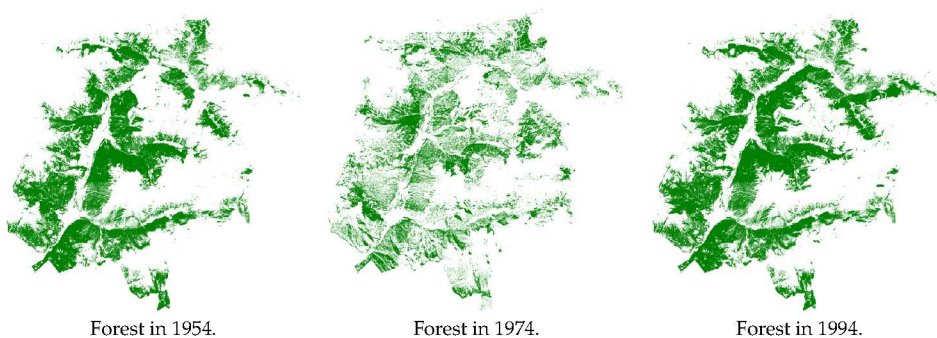


**Figure 7.** Location of Trentino region, in orange (in Italy), the Val di Fassa, in red (in the Trentino region), and the forest area in green (in Val di Fassa).



**Figure 8.** Maps of the forest coverage of Val di Fassa in 2006, original map (left), filtered with the third quartile (center) and first quartile (right).

Additionally, three maps representing the forest coverage in 1954, 1974, and 1994 (Figure 9), which present different degrees of fragmentation, have been used to compare the number of patches evaluated using different values of CN.



**Figure 9.** Forest coverage in Val di Fassa in 1954 (left), 1974 (center), and 1994 (right).

All these maps were created by automatic classification of aerial images using the same classification scheme. An area is considered forest if 20% of the ground area is covered by trees and it is larger than 2000 m<sup>2</sup> with a minimum width of 20 m.

#### 4. Results

For each test map described in Sections 3.3 and 3.4, the landscape metrics listed in Section 3.2 were evaluated using GRASS GIS, FRAGSTATS, and QGIS. Values for the artificial maps described in Section 3.3 are reported in Tables 2–5.

For the first artificial test map (Figure 3), configured as a chessboard, Table 2 shows how a 4-cells CN (GRASS, FRAGSTATS using 4 cells) is able to recognize the 5000 patches corresponding to single pixels and to correctly evaluate the corresponding metrics, while using 8-cells CN (QGIS, FRAGSTATS using 8 cells) one single patch is detected. This difference influences the patch density (PD) and mean patch size (MPS) metrics because they depend on the number of patches. Conversely, edge density (ED) and landscape shape index (LSI) do not change using different CN values because the lengths of all edge segments and the total landscape area change in the same proportion.

**Table 2.** Landscape metrics for artificial test map 1 (chessboard).

	GRASS	FRAGSTATS		QGIS
<b>Cell neighborhood</b>	<i>4 cells</i>	<i>4 cells</i>	<i>8 cells</i>	<i>8 cells</i>
<b>N patches</b>	5000	5000	1	1
<b>Patch density [patches/100 ha]</b>	10,000	10,000	2	2
<b>Mean patch size [m<sup>2</sup>]</b>	0.01	0.01	50	50
<b>Edge density [m/ha]</b>	4000	4000	4000	4000
<b>Landscape shape index</b>	70.71	70.42	70.42	70.71

Results for the second artificial test map (Figure 4), with two large patches separated by a single straight edge. Table 3 shows how all the landscape metrics values are the same regardless of the value of CN used. This is obviously due to the fact that the map contains only two large compact patches with no fragmentation and a single edge.

**Table 3.** Landscape metrics for artificial test map 2 (two large patches).

	GRASS	FRAGSTATS		QGIS
<b>Cell neighborhood</b>	<i>4 cells</i>	<i>4 cells</i>	<i>8 cells</i>	<i>8 cells</i>
<b>N patches</b>	2	2	2	2
<b>Patch density [patches/100 ha]</b>	2.02	2.02	2.02	2.02
<b>Mean patch size [m<sup>2</sup>]</b>	49.5	49.5	49.5	49.5
<b>Edge density [m/ha]</b>	60.4	60.4	60.4	60.4
<b>Landscape shape index</b>	1.5	1.5	1.5	1.5

For the third artificial test map (Figure 5), with a large patch covering half of the map, Table 4 reports how all the landscape metrics do not change with different CN choices. The reason is the same as for the previous map.

**Table 4.** Landscape metrics for artificial test map 3 (a large patch covering half of the map).

	GRASS	FRAGSTATS		QGIS
<b>Cell neighborhood</b>	<i>4 cells</i>	<i>4 cells</i>	<i>8 cells</i>	<i>8 cells</i>
<b>N patches</b>	1	1	1	1
<b>Patch density [patches/100 ha]</b>	2	2	2	2
<b>Mean patch size [m<sup>2</sup>]</b>	50	50	50	50
<b>Edge density [m/ha]</b>	60	60	60	60
<b>Landscape shape index</b>	1.06	1.06	1.06	1.06

Table 5 shows how both choices for CN correctly detect the only patch, composed by a single pixel, present in the fourth artificial test map (Figure 6). All the landscape metrics values are the same for the different CN values.

**Table 5.** Landscape metrics for artificial test map 4 (a single patch of one pixel).

	GRASS	FRAGSTATS	QGIS
<b>Cell neighborhood</b>	<i>4 cells</i>	<i>4 cells</i>	<i>8 cells</i>
<b>N patches</b>	1	1	1
<b>Patch density [patches/100 ha]</b>	10,000	10,000	10,000
<b>Mean patch size [m<sup>2</sup>]</b>	0.01	0.01	0.01
<b>Edge density [m/ha]</b>	4000	4000	4000
<b>Landscape shape index</b>	1	1	1

Finally, results for the real forest map are reported in Table 6. Values provided by QGIS for mean patch size and edge density differ from those calculated by FRAGSTATS with the same CN value by 4 orders of magnitude, respectively of a factor  $10^4$  and  $10^{-4}$ . This is due to the fact that areas in QGIS are expressed in square meters, while in GRASS and FRAGSTATS, *ha* is used. Therefore, the values of the two mean patch size and edge density metrics for QGIS were rescaled using *ha* in Table 6. The same procedure will be applied to all the successive results from QGIS.

**Table 6.** Landscape metrics for the 2006 forest map of Val di Fassa, values for QGIS in *ha*.

	GRASS	FRAGSTATS	QGIS
<b>Cell neighborhood</b>	<i>4 cells</i>	<i>4 cells</i>	<i>8 cells</i>
<b>N patches</b>	43,594	43,594	21,094
<b>Patch density [patches/100 ha]</b>	368.3626	368.2583	178.1906
<b>Mean patch size [m<sup>2</sup>]</b>	0.2715	0.2715	0.5611
<b>Edge density [m/ha]</b>	728.6887	728.5863	728.6592
<b>Landscape shape index</b>	198.1790	198.1631	198.1792

It is evident from Table 6 that some metrics, such as number of patches, patch density, and mean patch size, change when different values of CN are used. Edge density (ED) and landscape shape index (LSI) values are independent from the choice for CN.

## 5. Aggregation Index

The literature review of Section 2 revealed that in many cases the value of CN is not explicitly stated. Sometimes it is possible to infer the value by the indication of the GIS software used for the processing but even then there can be some ambiguity, such as for the case of FRAGSTATS which can use two different values and the default depends on the software version.

For this reason, it was investigated whether it is possible to find a parameter indicating how the uncertainty on the CN value reflects on the significance of the landscape metrics. Tests on artificial maps (Section 4, Tables 2–5) showed how the more fragmented a map is the more the values of landscape metrics change for different values of CN. Therefore, the aggregation index (AI), defined in Section 3.2, can be used as an indicator of the importance of the knowledge of CN for the analysis of a map. Its value varies between 0 and 100; it is 0 when the patch class is completely disaggregated, i.e., there are no like adjacencies; it is 100 when the patch class is completely aggregated in a single compact patch. Since tests in Sections 3.3 and 3.4 have shown that for a given value of CN landscape metrics do not depend on the software used to evaluate them, all the values reported in this section were evaluated using FRAGSTATS.

The first tests to investigate a possible relationship between the AI and the variation of the values of landscape metrics using different CN values were carried out on the artificial maps described in

Section 3.3. Table 7 shows the values of the AI for the four artificial maps, the number of patches obtained using values of CN 4 and 8, and their variation as percentage.

**Table 7.** Aggregation index and difference in the number of patches using cell neighbourhood (CN) values of 4 and 8 for the 4 artificial maps.

	Art. Map 1	Art. Map 2	Art. Map 3	Art. Map 4
<b>Aggregation Index [%]</b>	0	99.49	99.92	N/A
<b>N patches (4 cells)</b>	5000	2	1	1
<b>N patches (8 cells)</b>	1	2	1	1
<b>Difference % N patches</b>	99.98%	0	0	0

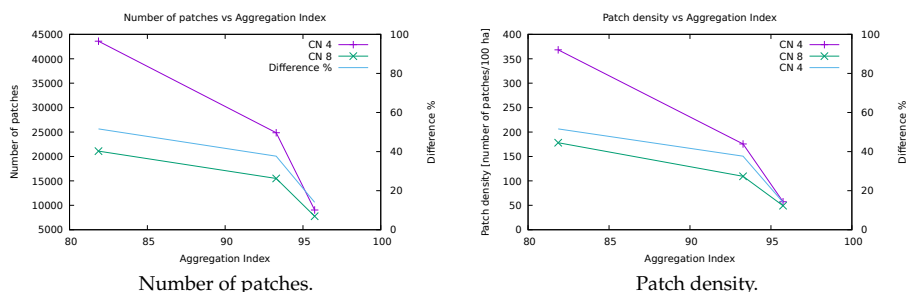
For the first raster, with a chessboard configuration, disaggregation is maximum and the AI is 0, while the difference in the number of patches for the two values of CN is nearly 100%. Conversely, the second and third artificial maps have an AI above 99% and the number of patches does not depend on CN (Tables 3 and 4). Finally, for the last artificial map the AI is undefined, since it is not possible to evaluate the AI for classes consisting of a single patch. These first tests show (Table 7) that for maps with low values of AI the use of different values of CN leads to large variations of even a simple metric such as the number of patches, while for maps with high AI values the choice for CN becomes irrelevant. The four artificial maps were built to represent extreme configurations; therefore, this hypothesis was tested on real maps. The three maps for the forest in Val di Fassa in 2006, described in Section 3.4, represent a sequence of maps with increasing patch compactness, as shown in Figure 8.

Table 8 reports landscape metrics evaluated using four cells CN and eight cells CN for the three maps. AI values do not depend on the choice of CN; therefore, it can be used to assess the effect of the variation of CN on landscape metrics.

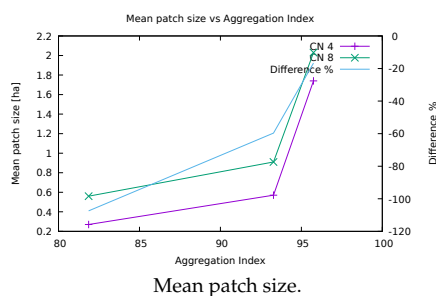
**Table 8.** Landscape metrics for the three forest maps of Val di Fassa with 4 cells and 8 cells CN.

	Original Map		Third Quartile		First Quartile	
	4 cells	8 cells	4 cells	8 cells	4 cells	8 cells
<b>N patches</b>	43,594	21,094	24,859	15,497	9037	7768
<b>Patch density [patches/100 ha]</b>	368.26	178.19	175.61	109.47	57.31	49.26
<b>Mean patch size [m<sup>2</sup>]</b>	0.27	0.56	0.57	0.91	1.74	2.03
<b>Edge Density [m/ha]</b>	728.59	728.59	271.82	271.82	173.02	173.02
<b>Landscape shape index</b>	198.16	198.16	80.83	80.83	54.30	54.30
<b>Aggregation index [%]</b>	81.86	81.86	93.28	93.28	95.75	95.75

Figure 10 shows the variation of the number of patches, patch density, and mean patch size metrics as a function of AI, both in the cases of 4 and 8 cells CN.



**Figure 10.** Cont.



**Figure 10.** Variation of number of patches, patch density, and mean patch size landscape metrics as a function of AI for CN values of 4 and 8, left scale, and their difference in percentage, right scale, evaluated on the map representing the forest coverage in Val di Fassa in 2006 and the two maps obtained by applying a low pass filter with a  $3 \times 3$  pixels window, assigning the first and the third quartile.

The difference between values obtained with the two values of CN decreases when the AI increases for all the tested indexes. In fact, the differences for the number of patches and the patch density decrease from 51.6% to 14.0%, while the difference for mean patch size, negative because the mean patch size increases with CN, varies from  $-107.4\%$  to  $-16.7\%$  (Table 9).

**Table 9.** Differences for number of patches, patch density (PD) (number of patches/100 ha) and mean patch size (MPS) (ha) for the two values of CN (4 or 8 cells) for the map representing the forest coverage in Val di Fassa in 2006 and the two maps obtained by applying a low pass filter with a  $3 \times 3$  pixels window, assigning the first and the third quartile.

	Original Map	Third Quartile	First Quartile
Aggregation index [%]	81.86	93.28	95.75
N patches (4 cells)	43,594	24,859	9037
N patches (8 cells)	21,094	15,497	7768
Difference % N patches	51.6%	37.6%	14.0%
Patch density (4 cells) [patches/100 ha]	368.26	175.61	57.31
Patch density (8 cells) [patches/100 ha]	178.19	109.47	49.26
Difference % PD	51.6%	37.6%	14.0%
Mean patch size (4 cells) [m <sup>2</sup> ]	0.27	0.57	1.74
Mean patch size (8 cells) [m <sup>2</sup> ]	0.56	0.91	2.03
Difference % MPS	$-107.4\%$	$-59.6\%$	$-16.7\%$

These tests show that using a different value for the CN leads to large differences for landscape metrics for fragmented maps, while the metrics values are similar for maps with more aggregated patches. Therefore, the aggregation index can be used as an indicator to assess the suitability of the values of the landscape metrics for making comparisons when the CN value is unknown.

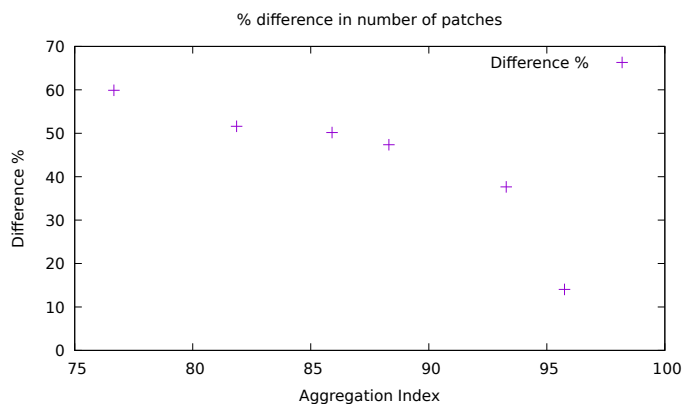
To further test this hypothesis, the number of patches was evaluated for three forest maps of Val di Fassa in 1954, 1974, and 1994 with the two values for CN. The results are reported in Table 10; for these maps the difference of number of patches is higher when the AI is lower.

**Table 10.** Differences of number of patches for the two values of CN (4 or 8 cells) for the three forest maps of Val di Fassa in 1954, 1974, and 1994.

	1954 Map	1974 Map	1994 Map
Aggregation index [%]	85.90	76.66	88.31
N patches (4 cells)	33,361	135,505	90,090
N patches (8 cells)	16,624	54,350	47,416
Difference % N patches	50.17%	59.89%	47.37%

For these maps the AI varies from 76.6% to 88.3%, with the difference in number of patches between the two CN values ranging from 47.37% (1994 map) to 59.89% (1974 map).

Figure 11 plots all the couples of values of AI and difference in number of patches using 4 and 8 cells CN in percentage for the six forest maps.



**Figure 11.** Difference in number of patches using 4 and 8 cells CN in percentage as a function of AI for the six forest maps.

When AI is lower than 95% the difference in the number of patches ranges from 37% to 60%, while for the highest value the difference is 14%.

## 6. Conclusions

This study provides an assessment of the importance of the knowledge of the cell neighborhood value when using landscape metrics. Tests on specially crafted maps have shown that CN can affect the values of some metrics with variations up to 5000%. For real maps the variation is not so dramatic but in some cases it can still reach 100% (Table 6). Therefore, the knowledge of the value used for the cell neighborhood (CN) parameter is fundamental, especially if metrics' values are compared across studies from different research groups using different software, even if the same classification scheme is used. Nevertheless, software users often seem to underestimate the importance of the effects that CN can have on landscape metrics and the selected CN value is often omitted.

When the CN is unknown, tests in Section 5 showed that it is possible to use the values of the aggregation index to assess the effect that the uncertainty on CN can have on some landscape metrics, thus providing an indication of the reliability of comparison with other metrics' evaluations. The determination of an analytical link between the AI and the significance of CN for landscape metrics evaluation is under way. This could lead to the identification of a threshold AI value below which the knowledge of the CN is essential. From the user perspective, the choice of cell neighborhood is implicit in the choice of the software. A 4-cells CN provides more evident trends for some metrics, see e.g., Figure 10. The main recommendation is to use the same cell neighborhood when comparing results. Moreover, researchers should clearly indicate the CN values when presenting their results.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2220-9964/8/12/586/s1>.

**Author Contributions:** Research conceptualization, Paolo Zatelli, Stefano Gobbi, Clara Tattoni and Marco Ciolli; methodology and software Paolo Zatelli, Marco Ciolli, Clara Tattoni and Stefano Gobbi; validation, Paolo Zatelli and Nicola Zorzi, Stefano Gobbi, Clara Tattoni and Marco Ciolli; writing—original draft preparation, Paolo Zatelli, Stefano Gobbi, Clara Tattoni and Marco Ciolli; writing—review and editing, Stefano Gobbi, Marco Ciolli, Nicola La Porta, Duccio Rocchini, Clara Tattoni, Maria Giulia Cantiani and Paolo Zatelli; funding acquisition, Marco Ciolli, Nicola La Porta and Paolo Zatelli.

**Funding:** Stefano Gobbi is funded by Doctoral Programme in Civil, Environmental and Mechanical Engineering established at the Department of Civil, Environmental and Mechanical Engineering at the University of Trento and Fondazione Edmund Mach.

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

## Abbreviations

The following abbreviations are used in this manuscript:

AI	aggregation index
CN	cell neighborhood
ED	edge density
GIS	geographic information system
GPL	Gnu Public License
FOSS	free and open source software
LSI	landscape shape index
MPS	mean patch size
NP	number of patches
PD	patch density

## Appendix A. Available Metrics

Feature tables for FRAGSTATS 4.2, GRASS 7.6 (r.li), and QGIS 3.4 (LecoS 3.0.0). Names of metrics change in different software systems, names in the following tables are from the FRAGSTATS 4.2 manual [45].

**Table A1.** Software features: cell neighborhood and available area and edge metrics for FRAGSTATS, GRASS (r.li), and QGIS (LecoS).

	FRAGSTATS	GRASS	QGIS
<b>Cell neighborhood</b>	4 or 8 cells	4 cells	8 cells
Patch Metrics			
Patch Area (AREA)	✓	✓	
Patch Perimeter (PERIM)	✓		
Radius of Gyration (GYRATE)	✓		
Class Metrics			
Total (Class) Area (CA)	✓	✓	
Percentage of Landscape (PLAND)	✓	✓	✓
Largest Patch Index (LPI)	✓		
Total Edge (TE)	✓		✓
Edge Density (ED)	✓	✓	✓
Patch Area Distribution (AREA_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Radius of Gyration Distribution (GYRATE_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Landscape Metrics			
Total Area (TA)	✓	✓	
Largest Patch Index (LPI)	✓		✓
Total Edge (TE)	✓		
Edge Density (ED)	✓	✓	
Patch Area Distribution (AREA_MN, _AM, _MD, _RA, _SD, _CV)	✓		✓
Radius of Gyration Distribution (GYRATE_MN, _AM, _MD, _RA, _SD, _CV)	✓		



**Table A2.** Software features: available shape metrics for FRAGSTATS, GRASS (r.li), and QGIS (LecoS).

	FRAGSTATS	GRASS	QGIS
Patch Metrics			
Perimeter-Area Ratio (PARA)	✓		
Shape Index (SHAPE)	✓	✓	✓
Fractal Dimension Index (FRAC)	✓		✓
Related Circumscribing Circle (CIRCLE)	✓		
Contiguity Index (CONTIG)	✓		
Class Metrics			
Perimeter-Area Fractal Dimension (PAFRAC)	✓		
Perimeter-Area Ratio Distribution (PARA_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Shape Index Distribution (SHAPE_MN, _AM, _MD, _RA, _SD, _CV)	✓	✓	
Fractal Index Distribution (FRAC_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Linearity Index Distribution (LINEAR_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Related Circumscribing Square Distribution (SQUARE_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Contiguity Index Distribution (CONTIG_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Landscape Metrics			
Perimeter-Area Fractal Dimension (PAFRAC)	✓		
Perimeter-Area Ratio Distribution (PARA_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Shape Index Distribution (SHAPE_MN, _AM, _MD, _RA, _SD, _CV)	✓	✓	
Fractal Index Distribution (FRAC_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Linearity Index Distribution (LINEAR_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Related Circumscribing Square Distribution (SQUARE_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Contiguity Index Distribution (CONTIG_MN, _AM, _MD, _RA, _SD, _CV)	✓		

**Table A3.** Software features: available core area metrics for FRAGSTATS, GRASS (r.li), and QGIS (LecoS).

	FRAGSTATS	GRASS	QGIS
Patch Metrics			
Core Area (CORE)	✓		
Number of Core Areas (NCA)	✓		
Core Area Index (CAI)	✓		
Class Metrics			
Total Core Area (TCA)	✓		✓
Core Area Percentage of Landscape (CPLAND)	✓		
Number of Disjunct Core Areas (NDCA)	✓		
Disjunct Core Area Density (DCAD)	✓		
Core Area Distribution (CORE_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Disjunct Core Area Distribution (DCORE_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Core Area Index Distribution (CAI_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Landscape Metrics			
Total Core Area (TCA)	✓		
Number of Disjunct Core Areas (NDCA)	✓		
Disjunct Core Area Density (DCAD)	✓		
Core Area Distribution (CORE_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Disjunct Core Area Distribution (DCORE_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Core Area Index Distribution (CAI_MN, _AM, _MD, _RA, _SD, _CV)	✓		

**Table A4.** Software features: available contrast metrics for FRAGSTATS, GRASS (r.li), and QGIS (LecoS).

	FRAGSTATS	GRASS	QGIS
Patch Metrics			
Edge Contrast Index (ECON)	✓		
Class Metrics			
Contrast-Weighted Edge Density (CWED)	✓	✓	
Total Edge Contrast Index (TECI)	✓		
Edge Contrast Index Distribution (ECON_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Landscape Metrics			
Contrast-Weighted Edge Density (CWED)	✓	✓	
Total Edge Contrast Index (TECI)	✓		
Edge Contrast Index Distribution (ECON_MN, _AM, _MD, _RA, _SD, _CV)	✓		

**Table A5.** Software features: available aggregation metrics for FRAGSTATS, GRASS (r.li), and QGIS (LecoS).

	FRAGSTATS	GRASS	QGIS
Patch Metrics			
Euclidean Nearest Neighbor Distance (ENN)	✓		✓
Proximity Index (PROX)	✓		
Similarity Index (SIMI)	✓		
Class Metrics			
Interspersion & Juxtaposition Index (IJI)	✓		
Percentage of Like Adjacencies (PLADJ)	✓		✓
Aggregation Index (AI)	✓		
Clumpiness Index (CLUMPY)	✓		
Landscape Shape Index (LSI)	✓		
Normalized Landscape Shape Index (nLSI)	✓		
Patch Cohesion Index (COHESION)	✓		✓
Number of Patches (NP)	✓	✓	✓
Patch Density (PD)	✓	✓	✓
Landscape Division Index (DIVISION)	✓		✓
Splitting Index (SPLIT)	✓		✓
Effective Mesh Size (MESH)	✓		✓
Euclidean Nearest Neighbor Distance Distribution (ENN_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Proximity Index Distribution (PROX_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Similarity Index Distribution (SIMI_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Connectance (CONNECT)	✓		
Landscape Metrics			
Contagion (CONTAG)	✓		
Interspersion & Juxtaposition Index (IJI)	✓		
Percentage of Like Adjacencies (PLADJ)	✓		
Aggregation Index (AI)	✓		
Landscape Shape Index (LSI)	✓		
Patch Cohesion Index (COHESION)	✓		
Number of Patches (NP)	✓		
Patch Density (PD)	✓		
Landscape Division Index (DIVISION)	✓		
Splitting Index (SPLIT)	✓		
Effective Mesh Size (MESH)	✓		
Euclidean Nearest Neighbor Distance Distribution (ENN_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Proximity Index Distribution (PROX_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Similarity Index Distribution (SIMI_MN, _AM, _MD, _RA, _SD, _CV)	✓		
Connectance (CONNECT)	✓		

**Table A6.** Software features: available diversity indexes for FRAGSTATS, GRASS (r.li), and QGIS (LeccoS).

	FRAGSTATS	GRASS	QGIS
Landscape Metrics			
Patch Richness Density (PRD)	✓	✓	
Relative Patch Richness (RPR)	✓		
Shannon's Diversity Index (SHDI)	✓	✓	✓
Simpson's Diversity Index (SIDI)	✓	✓	✓
Modified Simpson's Diversity Index (MSIDI)	✓		
Shannon's Evenness Index (SHEI)	✓		✓
Simpson's Evenness Index (SIEI)	✓		
Modified Simpson's Evenness Index (MSIEI)	✓		

## References

- Steiniger, S.; Hay, G.J. Free and open source geographic information tools for landscape ecology. *Ecol. Inform.* **2009**, *4*, 183–195. [CrossRef]
- Wang, X.; Blanchet, F.G.; Koper, N. Measuring habitat fragmentation: An evaluation of landscape pattern metrics. *Methods Ecol. Evol.* **2014**, *5*, 634–646. Available online: <https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.12198> (accessed on 12 September 2019). [CrossRef]
- Cushman, S.; McGarigal, K.; Neel, M. Parsimony in landscape metrics: Strength, universality, and consistency. *Ecol. Indic.* **2008**, *8*, 691–703. [CrossRef]
- Lustig, A.; Stouffer, D.B.; Doscher, C.; Worner, S.P. Landscape metrics as a framework to measure the effect of landscape structure on the spread of invasive insect species. *Landsc. Ecol.* **2017**, *32*, 2311–2325. [CrossRef]
- Sertel, E.; Topaloğlu, R.H.; Şallı, B.; Yay Algan, I.; Aksu, G.A. Comparison of Landscape Metrics for Three Different Level Land Cover/Land Use Maps. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 408. [CrossRef]
- Uuemaa, E.; Antrop, M.; Roosaare, J.; Marja, R.; Mander, Ü. Landscape Metrics and Indices: An Overview of Their Use in Landscape Research. *Living Rev. Landsc. Res.* **2009**, *3*. [CrossRef]
- Ferretti, F.; Sboarina, C.; Tattoni, C.; Vitti, A.; Zatelli, P.; Geri, F.; Pompei, E.; Ciolli, M. The 1936 Italian Kingdom Forest Map reviewed: A dataset for landscape and ecological research. *Ann. Silv. Res.* **2018**, *42*, 3–19. [CrossRef]
- Smiraglia, D.; Tombolini, I.; Canfora, L.; Bajocco, S.; Perini, L.; Salvati, L. The Latent Relationship Between Soil Vulnerability to Degradation and Land Fragmentation: A Statistical Analysis of Landscape Metrics in Italy, 1960–2010. *Environ. Manag.* **2019**, *64*, 154–165. [CrossRef]
- Grafius, D.R.; Corstanje, R.; Harris, J.A. Linking ecosystem services, urban form and green space configuration using multivariate landscape metric analysis. *Landsc. Ecol.* **2018**, *33*, 557–573. [CrossRef]
- Wang, S.; Yang, K.; Yuan, D.; Yu, K.; Su, Y. Temporal-spatial changes about the landscape pattern of water system and their relationship with food and energy in a mega city in China. *Ecol. Model.* **2019**, *401*, 75–84. [CrossRef]
- Li, H.; Wu, J. Use and misuse of landscape indices. *Landsc. Ecol.* **2004**, *19*, 389–399. LAND.0000030441.15628.d6. [CrossRef]
- McGarigal, K.; Cushman, S.; Ene, E. FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps. Computer Software Program Produced by the Authors at the University of Massachusetts, Amherst. 2012. Available online: <http://www.umass.edu/landeco/research/fragstats/fragstats.html> (accessed on 12 September 2019).
- Cushman, S.; States, U.; Service, F.; Mount, R. Behavior of class-level landscape metrics across gradients of class aggregation and area. *Landsc. Ecol.* **2015**, 435–455. [CrossRef]
- Saura, S.; Castro, S. Scaling functions for landscape pattern metrics derived from remotely sensed data: Are their subpixel estimates really accurate? *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 201–216. [CrossRef]
- Tanner, E.P.; Fuhlendorf, S.D. Impact of an agri-environmental scheme on landscape patterns. *Ecol. Indic.* **2018**, *85*, 956–965. [CrossRef]

16. Gong, J.-Z.; Liu, Y.-S.; Xia, B.-C.; Zhao, G.-W. Urban ecological security assessment and forecasting, based on a cellular automata model: A case study of Guangzhou, China. *Ecol. Model.* **2009**, *220*, 3612–3620. [[CrossRef](#)]
17. Abadie, A.; Gobert, S.; Bonacorsi, M.; Lejeune, P.; Pergent, G.; Pergent-Martini, C. Marine space ecology and seagrasses. Does patch type matter in *Posidonia oceanica* seascapes? *Ecol. Indic.* **2015**, *57*, 435–446. [[CrossRef](#)]
18. Lausch, A.; Herzog, F. Applicability of landscape metrics for the monitoring of landscape change: Issues of scale, resolution and interpretability. *Ecol. Indic.* **2002**, *2*, 3–15. [[CrossRef](#)]
19. Ramesh, T.; Kalle, R.; Downs, C.T. Sex-specific indicators of landscape use by servals: Consequences of living in fragmented landscapes. *Ecol. Indic.* **2015**, *52*, 8–15. [[CrossRef](#)]
20. Lustig, A.; Stouffer, D.B.; Roigé, M.; Worner, S.P. Towards more predictable and consistent landscape metrics across spatial scales. *Ecol. Indic.* **2015**, *57*, 11–21. [[CrossRef](#)]
21. Shen, W.; Jenerette, G.D.; Wu, J.; Gardner, R.H. Evaluating empirical scaling relations of pattern metrics with simulated landscapes. *Ecography* **2004**, *27*, 459–469. [[CrossRef](#)]
22. László, Z.; Dénes, A.L.; Király, L.; Tóthmérész, B. Biased parasitoid sex ratios: *Wolbachia*, functional traits, local and landscape effects. *Basic Appl. Ecol.* **2018**, *31*, 61–71. [[CrossRef](#)]
23. Bueno, A.S.; Peres, C.A. Patch-scale biodiversity retention in fragmented landscapes: Reconciling the habitat amount hypothesis with the island biogeography theory. *J. Biogeogr.* **2019**, *46*, 621–632. [[CrossRef](#)]
24. Cordeiro, E.M.; Macrini, C.M.; Sujii, P.S.; Schwarcz, K.D.; Pinheiro, J.B.; Rodrigues, R.R.; Brancalion, P.H.; Zucchi, M.I. Diversity, genetic structure, and population genomics of the tropical tree *Centrolobium tomentosum* in remnant and restored Atlantic forests. *Conserv. Genet.* **2019**, *20*, 1073–1085. [[CrossRef](#)]
25. Hernández-Martínez, J.; Morales-Malacara, J.B.; Alvarez-Anörve, M.Y.; Amador-Hernández, S.; Oyama, K.; Avila-Cabadilla, L.D. Drivers potentially influencing host-bat fly interactions in anthropogenic neotropical landscapes at different spatial scales. *Parasitology* **2019**, *146*, 74–88. [[CrossRef](#)] [[PubMed](#)]
26. Dennis, M.; Scaletta, K.L.; James, P. Evaluating urban environmental and ecological landscape characteristics as a function of land-sharing-sparing, urbanity and scale. *PLoS ONE* **2019**, *14*, e0215796. [[CrossRef](#)] [[PubMed](#)]
27. Kurta, A.; Auteri, G.G.; Hofmann, J.E.; Mengelkoch, J.M.; White, J.P.; Whitaker, J.O.; Cooley, T.; Melotti, J. Influence of a large lake on the winter range of a small mammal: Lake Michigan and the silver-haired bat (*Lasionycteris noctivagans*). *Diversity* **2018**, *10*, 24. [[CrossRef](#)]
28. Kosicki, J.Z. Are Landscape Configuration Metrics Worth Including When Predicting Specialist and Generalist Bird Species Density? A Case of the Generalised Additive Model Approach. *Environ. Model. Assess.* **2018**, *23*, 193–202. [[CrossRef](#)]
29. Lamamy, C.; Bombieri, G.; Zarzo-Arias, A.; González-Bernardo, E.; Penteriani, V. Can landscape characteristics help explain the different trends of Cantabrian brown bear subpopulations? *Mammal Res.* **2019**, 559–567. [[CrossRef](#)]
30. Ramachandran, R.M.; Roy, P.S.; Chakravarthi, V.; Sanjay, J.; Joshi, P.K. Long-term land use and land cover changes (1920–2015) in Eastern Ghats, India: Pattern of dynamics and challenges in plant species conservation. *Ecol. Indic.* **2018**, *85*, 21–36. [[CrossRef](#)]
31. Monti, F.; Nelli, L.; Catoni, C.; Dell’Omo, G. Nest box selection and reproduction of European Rollers in Central Italy: A 7-year study. *Avian Res.* **2019**, *10*, 1–12. [[CrossRef](#)]
32. Morales-Díaz, S.P.; Alvarez-Añorve, M.Y.; Zamora-Espinoza, M.E.; Dirzo, R.; Oyama, K.; Avila-Cabadilla, L.D. Rodent community responses to vegetation and landscape changes in early successional stages of tropical dry forest. *For. Ecol. Manag.* **2019**, *433*, 633–644. [[CrossRef](#)]
33. Lastrucci, L.; Cerri, M.; Coppi, A.; Dell’Olmo, L.; Ferranti, F.; Ferri, V.; Filippini, F.; Foggi, B.; Galardini, R.; Reale, L.; et al. Spatial landscape patterns and trends of declining reed-beds in peninsular Italy. *Plant Biosyst.* **2019**, *153*, 427–435. [[CrossRef](#)]
34. Statuto, D.; Cillis, G.; Picuno, P. GIS-based Analysis of Temporal Evolution of Rural Landscape: A Case Study in Southern Italy. *Nat. Resour. Res.* **2019**, *28*, 61–75. [[CrossRef](#)]
35. Colson, F.; Bogaert, J.; Filho, A.C.; Nelson, B.; Pinagé, E.R.; Ceulemans, R. The influence of forest definition on landscape fragmentation assessment in Rondônia, Brazil. *Ecol. Indic.* **2009**, *9*, 1163–1168. [[CrossRef](#)]
36. Aguilar, R.; Calviño, A.; Ashworth, L.; Aguirre-Acosta, N.; Carbone, L.M.; Albrieu-Llinás, G.; Nolasco, M.; Ghilardi, A.; Cagnolo, L. Unprecedented plant species loss after a decade in fragmented subtropical chaco serrano forests. *PLoS ONE* **2018**, *13*, 1–15. [[CrossRef](#)]

37. Albrieu-Llinás, G.; Espinosa, M.O.; Quaglia, A.; Abril, M.; Scavuzzo, C.M. Urban environmental clustering to assess the spatial dynamics of *Aedes aegypti* breeding sites. *Geospat. Health* **2018**, *13*, 135–142. [[CrossRef](#)]
38. Colson, F.; Bogaert, J.; Ceulemans, R. Fragmentation in the Legal Amazon, Brazil: Can landscape metrics indicate agricultural policy differences? *Ecol. Indic.* **2011**, *11*, 1467–1471. [[CrossRef](#)]
39. Gaucherel, C.; Burel, F.; Baudry, J. Multiscale and surface pattern analysis of the effect of landscape pattern on carabid beetles distribution. *Ecol. Indic.* **2007**, *7*, 598–609. [[CrossRef](#)]
40. Llauss, A.; Nogué, J. Indicators of landscape fragmentation: The case for combining ecological indices and the perceptive approach. *Ecol. Indic.* **2012**, *15*, 85–91. [[CrossRef](#)]
41. Tattoni, C.; Ciolli, M.; Ferretti, F. The fate of priority areas for conservation in protected areas: A fine-scale markov chain approach. *Environ. Manag.* **2011**, *47*, 263–278. [[CrossRef](#)]
42. McGarigal, K.; Marks, B.J. *FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure*; USDA For. Serv. Gen. Tech. Rep. PNW-351; U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station: Portland, OR, USA, 1995.
43. GRASS Development Team. Geographic Resources Analysis Support System (GRASS) Software. Open Source Geospatial Foundation. 2017. Available online: [grass.osgeo.org](http://grass.osgeo.org) (accessed on 12 September 2019).
44. QGIS Development Team. QGIS Geographic Information System. Open Source Geospatial Foundation Project. 2019. Available online: <http://qgis.osgeo.org> (accessed on 12 September 2019).
45. McGarigal, K. *FRAGSTATS HELP*. 2015. Available online: <https://www.umass.edu/landeco/research/fragstats/documents/fragstats.help.4.2.pdf> (accessed on 12 September 2019).
46. Neteler, M.; Bowman, M.; Landa, M.; Metz, M. GRASS GIS: A multi-purpose Open Source GIS. *Environ. Model. Softw.* **2012**, *31*, 124–130. [[CrossRef](#)]
47. Ciolli, M.; Federici, B.; Ferrando, I.; Marzocchi, R.; Sguerso, D.; Tattoni, C.; Vitti, A.; Zatelli, P. FOSS tools and applications for education in geospatial sciences. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 225. [[CrossRef](#)]
48. GRASS Development Team. GRASS GIS Performance. 2019. Available online: [https://grasswiki.osgeo.org/wiki/GRASS\\_GIS\\_Performance](https://grasswiki.osgeo.org/wiki/GRASS_GIS_Performance) (accessed on 15 September 2019).
49. Jung, M. LecoS—A python plugin for automated landscape ecology analysis. *Ecol. Inform.* **2016**, *31*, 18–21. [[CrossRef](#)]
50. Zaragoza, B.; Belda, A.; Linares, J.; Martínez-Pérez, J.E.; Navarro, J.T.; Esparza, J. A Free and Open Source Programming Library for Landscape Metrics Calculations. *Environ. Model. Softw.* **2012**, *31*, 131–140. [[CrossRef](#)]
51. Hesselbarth, M.H.K.; Scianini, M.; With, K.A.; Wiegand, K.; Nowosad, J. landscapemetrics: An open-source R tool to calculate landscape metrics. *Ecography* **2019**, *42*, 1648–1657. [[CrossRef](#)]
52. Ciolli, M.; Tattoni, C.; Ferretti, F. Understanding forest changes to support planning: A fine-scale Markov chain approach. *Dev. Environ. Model.* **2012**, *25*, 355–373. [[CrossRef](#)]
53. Tattoni, C.; Ciolli, M.; Ferretti, F.; Cantiani, M.G. Monitoring spatial and temporal pattern of Paneveggio forest (Northern Italy) from 1859 to 2006. *IForest* **2010**, *3*, 72–80. [[CrossRef](#)]
54. McGarigal, K. Landscape Pattern Metrics. In *Wiley StatsRef: Statistics Reference Online*; American Cancer Society: New York, NY, USA, 2014. Available online: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118445112.stat07723> (accessed on 12 September 2019). [[CrossRef](#)]
55. Baker, W.; Cai, Y. The r.le programs for multiscale analysis of landscape structure using the GRASS geographical information system. *Landsc. Ecol.* **1992**, *7*, 291–302. [[CrossRef](#)]
56. Gobbi, S.; Cantiani, M.G.; Rocchini, D.; Zatelli, P.; Tattoni, C.; Ciolli, M.; La Porta, N. Fine spatial scale modelling of Trentino past forest landscape (Trentinoland): A case study of FOSS application. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences—ISPRS Archives, Bucharest, Romania, 26–30 August 2019; Volume XLII-4/W14, pp. 71–78. doi:10.5194/isprs-archives-XLII-4-W14-71-2019. [[CrossRef](#)]



### 3.6 AFTER VAIA STORM

On October 29<sup>th</sup>, 2018 the storm afterward named Vaia hit the Northeastern part of Italy. The storm generated in the Baltic Sea and was then powered by cold streams from the North, which generated in Italy strong winds [81]. By the time the storm hit PAT and neighbour territory the winds reached spikes of 200km/h in speed. This winds, helped by local conditions, such as sides exposure and slopes, caused profound and extended damage to the PAT forests, manifested in windthrows. In map 84 (in red) the areas interested by windthrow over the map of forest of 2015 are reported. The areas of the windthrow were detected in the work of [18] and reported to the public in the ForWind project [33].

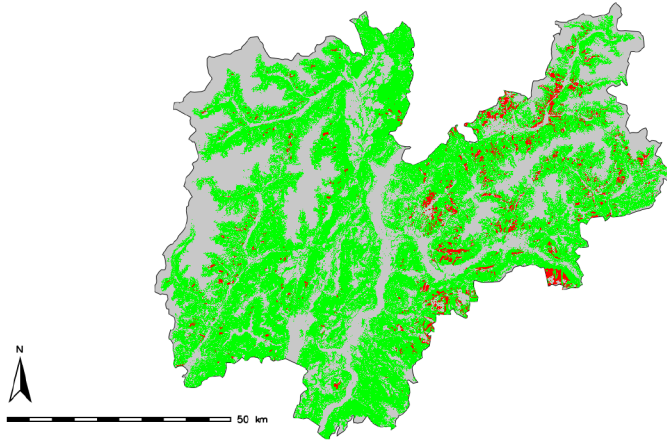


Figure 84: Areas interested by windthrow after Vaia storm (in red) over the 2015 forest map.

Table 49 reports on the extension of windthrow due to Vaia. Around 19000 ha of forests were destroyed, distributed heterogeneously over PAT, with more areas thrown in eastern part of this region. The third column of the table reports on the extension of the windthrow in areas which were invaded by the forests since 1954. A remarkable data that emerged was that around a third of the windthrow happened in these areas.

	Forest after Vaia storm	windthrow	windthrow in invaded areas
Area [ha]	341521	19495	5664

Table 49: Areas interested by Vaia storm.

The anomalous phenomenon destroyed an amount of forest biomass that is corresponding to around 10 years of annual prescribed wood cutting in PAT [35]. The current situation of the forest in PAT could not leave aside this catastrophic event. To recover the areas that were opened by the storm an incredible effort was undertaken by the local government to remove the fallen trees, which could be parasite carriers, and therefore generate diffuse diseases on the standing trees. After that the areas will be cleared by fallen trees the problems will be related to the fact that the terrain in this newly open areas will be unstable and it will be difficult for some year to the seed of new trees to settle. It is difficult to make a prevision of the time that will take to recover from this damage, because a lot of different variables could slow or accelerate the process.



Figure 85: The windthrow after Vaia storm in Passo Lavazzé, from [42]



### 3.7 FUTURE CHANGES SCENARIOS IN THE PANEVEGGIO PALE DI SAN MARTINO PARK

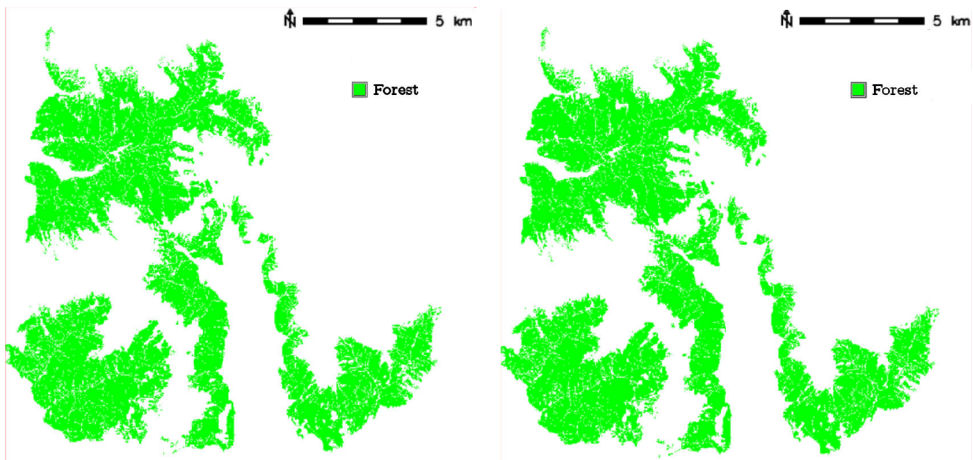
The aim of the ABM procedure implemented in the Paneveggio Pale di San Martino Park was to simulate a possible changes scenarios for forests in the park. The main assumption was that no phenomena other than the ones described by the ABM model will happen in the area, and that the dynamics of the forest growth happened during the phase of calibration of the model will continue in the future.

The year 0 for the simulation was the year 2015, and the final year was the year 2100.

The simulation was compared with a Markov chain procedure implemented in the same area. Markov chain is a provisional methods used to perform prevision of the evolution of complex systems. Contrary to ABM Markov Chains is based only on the probability of a single cell in a discrete system to change state, basing on purely stochastic and casual process. Markov chain has to be integrated in their prevision process by other input, in the case of [20] these input were given in the form of suitability maps that contains the information about the most likely cell to change state. The Markov chain procedure generated a map for the year 2050 and for the year 2100, and the same maps were generated for the ABM procedure for a comparison of the situation.

Year	Forest Area-ABM [ha]	Forest Area-MC [ha]
2015	7909	7909
2050	8908	9578
2100	9119	9988

Table 50: A possible evolution of forest extension in Paneveggio Pale di San Martino Park, a comparison between two methods for future changes scenarios prediction



(a) The situation predicted for the year 2050 in Paneveggio Pale di San Martino Park using ABM

(b) The situation predicted for the year 2100 in Paneveggio Pale di San Martino Park using ABM

Figure 86: The prediction of the evolution of forest in Paneveggio Pale di San Martino using ABM.

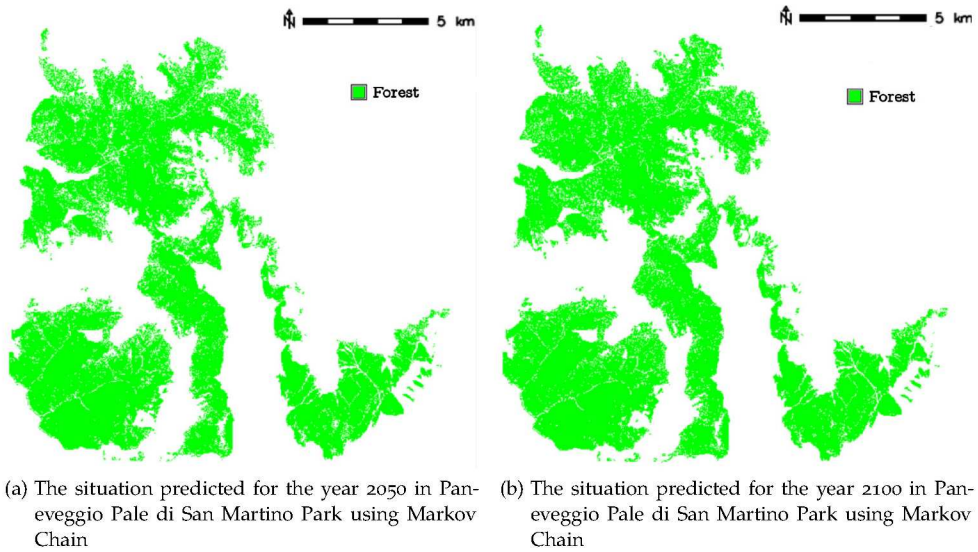


Figure 87: The prediction of the evolution of forest in Paneveggio Pale di San Martino using Markov Chain.

The results reported an apparently similar situation for the year 2050 and 2100. The models concur to simulate a forest expansion which behaves accordingly to what was observed in section 3.3. MC simulates a greater forest expansion than ABM, but the main differences are in the quality of forest growth. As the table 51 reports, a metric analysis was performed on both the MC simulation and the ABM situation.

Year	NP - ABM	NP - MC	Mean Area - ABM [ha]	Mean Area - MC [ha]
2050	406	7204	21.9	0.5
2100	663	9073	12.9	0.35

Table 51: A comparison between metrics calculated on both the simulated landscape using MC and ABM

MC tends to simulate a scattered afforestation, while ABM tends to cluster trees and make forest grow by closing the open areas near present trees. This is due to how ABM works: to simulate the born of a new tree the model checks the seed distribution, that is function of the distance of the evaluated cell to the nearest trees. On the other hand MC just checks for the probability of every cell to change their state, regardless their mutual position of the evaluated cells and the trees. By comparing the forest expansion explained in section 3.3 and 3.4 with the results of both these methods what seems to be the more realistic for this area of PAT is the ABM. It is worth reminding that the model could be furthermore enhanced, by adding data about the terrain humidity, the exposure of the sides, the slope and other information about the topography.

It is worth reminding as well that the ABM procedure was conceived and tested on a limited area of the study region, applying the same procedure to simulate the growth of the forest in different areas whit the same parameters could simulate an unrealistic evolution. The calibration therefore is a crucial and central step for evaluating future changes scenarios, as well as having a good understanding of the process that mainly drives forest expansion. For instance in our case the main forest class in Paneveggio are the Fir and Spruce forest, which

seeds are scattered by the wind. Other behaviours in the spreading of the seed (animal driven seed dispersal) are not well represented, because the distance from the nearby trees becomes an irrelevant variable. We can conclude that the ABM model simulated coherently the forest evolution in Paneveggio Pale di San Martino in the next 85 years with how forest evolved in the last 160 years.

### 3.8 EVALUATION OF ECOSYSTEM SERVICES

Ecosystem services are defined in the Millennium Development Goals (MDGs) as the benefits that an ecosystem provide to the population [86], which are divided in main classes:

- Support to life, or habitat provisioning;
- Provisioning of material goods;
- Climate and environmental regulation;
- Cultural values.

To evaluate the support to life provided by the forest it was evaluated the changes in ecotonal length as TE, reported in graph 68. This metrics decreases from the year 1954 toward the present day, a data that in the perspective of habitat provisioning evaluation can be interpreted as the lowering of spaces between forest and open areas. This represented a loss of habitat for many different species such as the *Tetrao urogallus* which lives mostly in the forest, but needs open spaces for its social life. On the other hand, other species, such as *Ursus Arctos* benefit from the compaction of the forests, because a more closed and connected forest landscape means a better shelter for movements, more opportunities to feed and built a lair.

For climate and environmental regulation the calculus here proposed is the Revised Universal Soil Loss Equation (RUSLE). This formula proposed by [123] is used to describe the soil loss due to rainfall events. The formula describes the soil loss as function of different factors:

$$A = R \cdot K \cdot L_s \cdot C \cdot P$$

Where:

- A is the annual soil loss due to erosion [t/ha year];
- R the rainfall erosivity factor;
- K the soil erodibility factor;
- $L_s$  the topographic factor derived from slope length and slope gradient;
- C the cover and management factor;
- P the erosion control practice factor.

This formula takes into account the local rain conditions in the R factor, and the soil erodibility in function of its composition and texture with the factor K. The factor C takes into account the land cover in each point where the RUSLE is calculated. Values of C near zero means that the land cover is more conservative, for instance the forest category provides a good protection against the soil loss, therefore associated to this land class is a C value near 0. The European Union provides a map of A calculated for its entire territory, with a 100x100m resolution [27]. The data provided by the EU also involve the singular factors that compose the RUSLE formula. Using the *r.mapcalc* it was possible to calculate the RUSLE in the area of forest by combining the single maps of each single factor. The result of this calculus is the soil erosion avoided by the forests. Figure 88 and table 52 reports on the avoided soil loss in the year of each dataset by the forest.

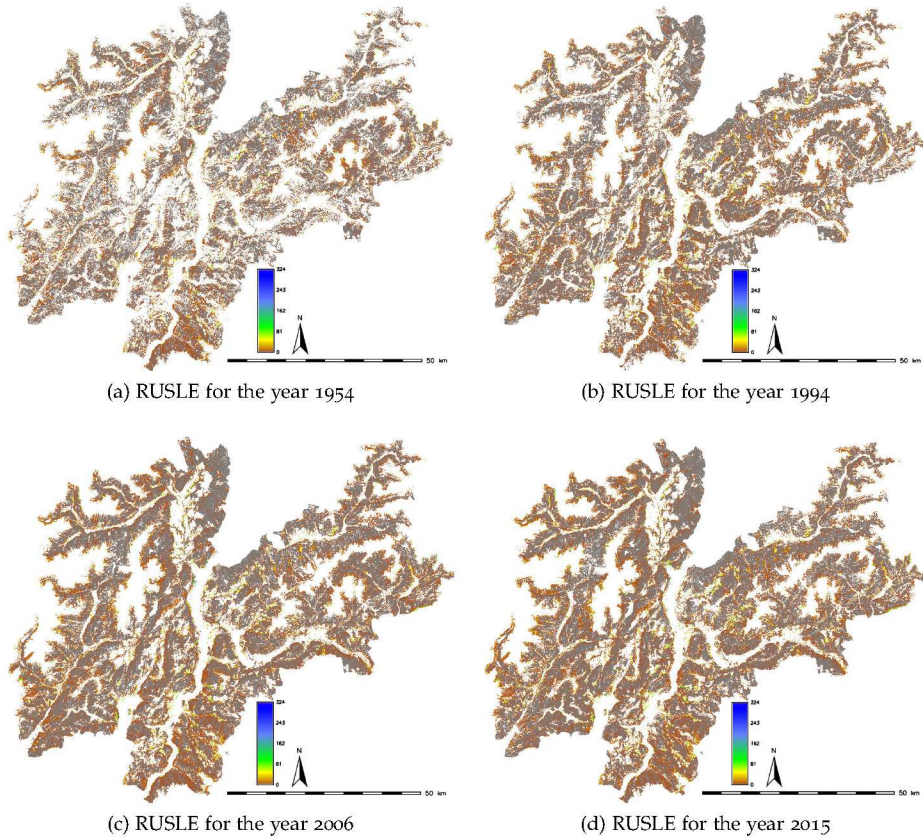


Figure 88: The value for the RUSLE formula evaluated in the different year of the current analysis.

Year	Avoided soil erosion in PAT [Mt/ha*y]
1954	60.5
1994	85.6
2006	96.5
2015	102.7

Table 52: Value of avoided soil erosion in areas occupied by forest across the dataset.

With the expansion of the forest obviously there was an increase of the protection from the soil loss, also the afforestation in potential excluded areas (figure 61) such as gullies and rocky areas provided a strong protection from soil loss. It is worth reminding then, that RUSLE only takes into account the soil loss due to sheet and rill erosion and ignores the effects of gully erosion and dispersive soils. For these reasons the RUSLE should be taken as an index of the quantity of the soil loss, and the avoided soil loss as an indication of the capability of the forest to maintain the soil balance.

Finally, as cultural values provided by the forest, it is possible to take into account the work of Tattoni [113] which analysed the data about the sales record of postcards in a small area of PAT, to relate the structure of the landscape depicted in postcards to the preferences of the general public. Postcards with overly-exploited landscape or with dense forest coverage were less likely chosen, over landscape with alternation of forest and open areas. This is an

index that even if the structure of the forest evolved toward a situation of a better diversity in tree species and forest classes the human perception of the forest is not detailed enough to recognize patterns of biodiversity. Therefore, the cultural perception of a 'better looking' mountain landscape by the touristic public could be more linked to a situation similar of what could have been in the year 1970's, when forest were less dense and there was more alternation between forest areas and open areas, an intermediate situation between 1954 and the present day.

As a general indication ecosystem services provided by an expansion of forest were quite obviously increased, while this expansion meant a loss in a certain type of biodiversity and landscape diversity. As most of the processes in nature even this phenomena which appears in a relatively small region of the southern Alps shows a double side, depending on the face one is looking to afforestation this could show a positive or negative side.

## CONCLUSIONS

---

The present work reconstructed the forest landscape in PAT territory over the last 160 years, analysed how the forest ecosystem has evolved and the effect it had on the rest of the mountain landscape in PAT. This involved a thorough and complete research between all the available data to find the most significant ones to perform the analysis of the variation of the forests. After careful analysis the data sources used in this work were: 1859 cadastral maps, then the 1936 IKFM, the aerial imagery of the years 1954, 1994, 2006 and 2015. The fil-rouge that bounded together this mapsets was their fine resolution, which was at least 2x2m for the 1954 set of imagery, while it reached a 0.5x0.5 resolution for the 2006 and 2015 set of imagery. These data were collected and organized in different locations, in order to ease their elaboration. This step of the research found some issue in the original datasets, such as a misplacement of some municipality in the 1859 datasets and the non-homogeneity of the histogram of the greyscale images (of 1954 and 1994). Along with this images a series of other useful data were collected, among which, the DEM of the study area, a map of the forest classes and the map of the potential forest area for PAT.

The images of 1954 were given in non geographical format, therefore they needed a process of orthorectification, to transform their original projection into a geographical projection. Along with this process an accuracy analysis of the coordinates of the resultant rectified images was performed, linking the position of the GCPs to a number of factors, such as slope of the terrain, altitude of the points in the images and position of the GCPs respect to the centre of the image. The best terrain features to be used as GCPs were identified as well. The output of this part of the research was a series of orthorectified imagery of a portion of the PAT territory in 1954 that together with the work of other authors provided a complete orthorectified image of the PAT territory.

The literature review about the best methods to extract the forest cover from aerial imageries revealed that the most effective one was the land use classification using an object based image classification approach (OBIA). This approach is preferable in fine scale images, because it allows the detection of single entities in an image. The algorithm can, exploits the geometrical features of each entity to perform a better classification. Along with the application of this process to the aerial imagery, a new module in GRASS GIS was developed with the aim to classify marks and symbols from historic and thematic map in a separate land use class which will be then blended with the background land class. This process allowed a better classification of the historic maps.

The OBIA procedure was automatized and applied to each map in each dataset and from the resulting land cover maps the category forest was extracted. These data were validated by checking the correct classification of land use in a number of random points for each dataset to check the accuracy of the classification. The validation process identified some limitations of the OBIA procedure: segmentation was an issue in this procedure, especially in those images that carried little information i.e. greyscale images. The algorithm that performed segmentation tent to perform over segmentation on some objects, an error which led to a misclassification on small areas. The maps of 1954, in particular, needed a revision of the classification which was integrated with a method based on the classification of pixels based on their value. It is worth underlying that these errors emerged from collateral analysis of the extracted data, and it is recommendable to always perform a visual check of the results of a classification procedure, especially if performed on a regional scale. If the classification reports wrong results the further analysis is affected and leads to misleading conclusion. OBIA

procedure, otherwise, proved to be effective, and it was easily automatable for all the aerial imageries used in the current work and provided an accurate result on fine scale multiband images, but it was easily adaptable to other maps, such as the H.C.P.A.T. or the IKFM.

Finally, after the extraction of the forest category from each dataset the research split in three main strands: the first was the analysis of the evolution of forest coverage, the second was the evaluation of some ecosystem services, the third was the setting up of a procedure to analyse possible scenarios for the forest growth.

The analysis of forest coverage underlined a situation where from 1859 up to the present day there has been a progressive and extended afforestation. Furthermore, this process has been studied using a set of landscape metrics, which are useful tools to understand how landscape changed through the years. Landscape metrics highlighted that forests in PAT expanded by closing open areas and by progressively compacting, rather than colonize sparsely new areas. This process was studied for each forest class and in different altitude ranges, the results concluded that afforestation was an homogeneous process in each altitude range and forest class. Along with the analysis of landscape metrics an issue was underlined in the current Literature, where a fundamental parameter for the calculus of landscape metrics, the Cell Neighbors (CN), is often not declared. The results of a landscape metrics analysis depend even on the shape of the CN and not declaring this parameter could lead to a misinterpretation of the results. This issue was analysed for the first time in the available Literature along the present work.

The study of ecosystem services highlighted a situation where afforestation favoured some processes, such as defence from soil erosion, but depauperated the biodiversity and the ecological habitats.

The afforestation process highlighted by this work was similar to what was described in other areas of the Alps [71, 107, 60, 8, 120, 26, 105]. The dynamics which led the forest expansion are similar: as agriculture and herding become less relevant in the mountain ecosystem forest starts to colonize the abandoned areas. This process follows a period of forest exploitation with a consequent deforestation [59]. This shows how the combined action of abandonment of rural areas [114] on a global scale and climate changes are affecting forest ecosystems in the Alps.

Finally, future change scenarios were studied in a protected area of PAT (Paneveggio Pale di San Martino Natural Park) using a procedure based on Agent Based Models (ABM). From this analysis it emerged that afforestation will continue in future years, if external factors do not intervene. The results of this procedure were then compared with the results of a different procedure for calculating future changes scenarios, namely Markov Chain. The two procedures were concordant in simulating afforestation.

A series of future development of the current research is being proposed. Firstly, the data about forest coverage are being used in the Climate KIC project Saturn, a project in cooperation between Fondazione Edmund Mach, Birmingham municipality, Climate KIC, Gothenburg municipality, Hub Innovazione Trentino, University of Trento set up to study the forest expansion in Trento Municipality and neighbours territories to study how afforestation has affected the abandoned agricultural grounds. After the tragic events of Vaia storm in 2018, an application of the ABM procedure can be done to simulate the reinvasion of the forests in those areas interested by windthrows, to assess the time needed for the forests to recover. Forest maps represent a valuable source of information to study ecological connectivity of forest and open areas and how they have evolved since the 1859. The data about forest coverage can be furthermore unbundled and the variation of the forest presence can be studied for each single municipality, or other sub-areas (e.g. protected areas) over the PAT territory. The aerial images of 1954 have been rectified for the entirety of PAT territory and have a number of possible applications, such as study of the situation of the infrastructure, urbanization or the



glacier area in the middle of the XX century. The evolution of ecosystem services provided by the forest can be evaluated: the provision of timber, the habitat provision for some specific species, the invasion of the forest in landslide areas and evapotranspiration are some examples. Finally the ABM procedure proved its effectiveness, and it could be used to perform a prevision on the evolution of some protected areas and to assess the damage of the Vaia storm.

The current work has fulfilled its aims, while providing in the meantime new tools and enhancement of existing tools for remote sensing and leaving as heritage a large dataset that will be used to deepen the knowledge of the territory of PAT, and, more widely, to study emerging patterns in afforestation in an alpine environment.

## BIBLIOGRAPHY

---

- [1] European Environment Agency. *Agriculture and the environment in the EU accession countries*. Technical report, European Environment Agency, Copenhagen, 2004.
- [2] I. Alberico, G. Cavuoto, V. Di Fiore, M. Punzo, D. Tarallo, N. Pelosi, L. Ferraro, and E. Marsella. Historical maps and satellite images as tools for shoreline variations and territorial changes assessment: the case study of volturno coastal plain (southern italy). *Journal of Coastal Conservation*, 22(5):919–937, 2018. doi: 10.1007/s11852-017-0573-x.
- [3] N.L. Alvarez-Berrios, D.J. Redo, T.M. Aide, M.L. Clark, and R. Grau. Land change in the greater antilles between 2001 and 2010. *Land*, 2(1):81–107, 2013. doi: <https://doi.org/10.3390/land2020081>.
- [4] G. Antoniol, C. Basco, M. Ceccarelli, M. Metz, and M. Lennert. GRASS Development Team r.texture. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6. <https://grass.osgeo.org/grass76/manuals/r.texture.html>, 2016. accessed on 25 July 2019.
- [5] M. Antrop. Why landscapes of the past are important for the future. *Landscape and Urban Planning*, 70(1-2):21–34, 2005. doi: 10.1016/j.landurbplan.2003.10.002.
- [6] Andrey Lessa Derci Augustynczik. Habitat amount and connectivity in forest planning models: Consequences for profitability and compensation schemes. *Journal of Environmental Management*, 283:111982, 2021. ISSN 0301-4797. doi: <https://doi.org/10.1016/j.jenvman.2021.111982>. URL <https://www.sciencedirect.com/science/article/pii/S030147972100044X>.
- [7] Tamanini B. Alcune considerazioni sui rimboschimenti, le coniferazioni, le conversioni. *Dendronatura*, 1:55–69, 01 1993.
- [8] D. Baldock, G. Beaufoy, F. Brouwer, and F. Godeschalk. *Abandonment or redeployment of agricultural land in Europe*, pages 70–78. IEEP London and LEI-DLO den Haag., 1996. ISBN 90-5242-368-7.
- [9] Saikat Basu, Sangram Ganguly, Ramakrishna R. Nemani, Supratik Mukhopadhyay, Gong Zhang, Cristina Milesi, Andrew Michaelis, Petr Votava, Ralph Dubayah, Laura Duncanson, Bruce Cook, Yifan Yu, Sassan Saatchi, Robert DiBiano, Manohar Karki, Edward Boyda, Uttam Kumar, and Shuang Li. A semiautomated probabilistic framework for tree-cover delineation from 1-m naip imagery using a high-performance computing architecture. *IEEE Transactions on Geoscience and Remote Sensing*, 53(10):5690–5708, 2015. doi: 10.1109/TGRS.2015.2428197.
- [10] M. Belgiu and L. Dragut. Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. *International Journal of remote sensing*, 96:67–75, 2014.
- [11] Clemens Blattert, Renato Lemm, Esther Thurig, Golo Stadelmann, Urs-Beat Brandli, and Christian Temperli. Long-term impacts of increased timber harvests on ecosystem services and biodiversity: A scenario study based on national forest inventory data. *Ecosystem Services*, 45:101150, 2020. ISSN 2212-0416. doi: <https://doi.org/10.1016/j.ecosyst.2020.101150>.

1016/j.ecoser.2020.101150. URL <https://www.sciencedirect.com/science/article/pii/S2212041620300929>.

- [12] Blazek, R. GRASS Development Team v.what.rast. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6. <https://grass.osgeo.org/grass74/manuals/v.what.rast>, 2019. accessed on 25 July 2019.
- [13] C.A. Bouman and M. Shapiro. A multiscale random field model for bayesian image segmentation. *IEEE Transactions on Image Processing*, 3(2):162–177, 1994. doi: 10.1109/83.277898.
- [14] D Buffoni, D. Leoni, and R. Bortolamedi. L’eredità cartografica catastale degli Asburgo in formato digitale. *Atti della Sesta conferenza italiana utenti ESRI*, pages 9–10, 2003.
- [15] A.P. Carleer, O. Debeir, and E. Wolff. Assessment of very high spatial resolution satellite image segmentations. *Photogrammetry Engineering Remote Sensing*, 71:1285–1294, 2005. doi: 10.14358/PERS.71.11.1285.
- [16] Giulia Cattani. *Metodologia per la valutazione della perdita di biodiversità dovuta al cambiamento del paesaggio agroforestale: il caso del Tesino*. University of Trento, master’s degree thesis, 2015.
- [17] Y.-Y. Chen, W. Huang, W.-H. Wang, J.-Y. Juang, J.-S. Hong, T. Kato, and S. Luyssaert. Reconstructing taiwan’s land cover changes between 1904 and 2015 from historical maps and satellite images. *Scientific Reports*, 9(1), 2019. doi: 10.1038/s41598-019-40063-1.
- [18] G Chirici, F Giannetti, D Travaglini, S Nocentini, S Francini, G D’Amico, E Calvo, D Fasolini, M Broll, F Maistrelli, J Tonner, M Pietrogiovanna, K Oberlechner, A Andriolo, R Comino, A Faidiga, I Pasutto, G Carraro, S Zen, F Contarin, L Alfonsi, A Wolynski, M Zanin, C Gagliano, S Tonolli, R Zoanetti, R Tonetti, R Cavalli, E Lingua, F Pirotti, S Grigolato, D Bellingeri, E Zini, D Gianelle, M Dalponte, E Pompei, A Stefani, R Motta, D Morresi, M Garbarino, G Alberti, F Valdevit, E Tomelleri, M Torresani, G Tonon, M Marchi, P Corona, and M Marchetti. Stima dei danni della tempesta "Vaia" alle foreste in Italia. *Forest@ - Rivista di Selvicoltura ed Ecologia Forestale*, 16(1):3–9, 2019. doi: 10.3832/efor3070-016. URL <http://foresta.sisef.org/contents/?id=efor3070-016>.
- [19] M Ciolli and E. Romagnoni. *Flora e vegetazione della Val di Leno*. Annali dell’Accademia di Scienze Forestali, 1996.
- [20] M. Ciolli, C. Tattoni, and F. Ferretti. Understanding forest changes to support planning: A fine-scale markov chain approach. *Developments in Environmental Modelling*, 25:355–373, 2012. doi: 10.1016/B978-0-444-59396-2.00021-3. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84885231036&doi=10.1016%2fB978-0-444-59396-2.00021-3&partnerID=40&md5=10bd3a529c11b1789a872320db2d48e0>.
- [21] M. Ciolli, B. Federici, I. Ferrando, R. Marzocchi, D. Sguerso, C. Tattoni, A. Vitti, and P. Zatelli. Foss tools and applications for education in geospatial sciences. *ISPRS International Journal of Geo-Information*, 6(7), 2017. doi: 10.3390/ijgi6070225.
- [22] M. Ciolli, M. Bezzi, G. Comunello, G. Laitempergher, S. Gobbi, C. Tattoni, and M.G. Cantiani. Integrating dendrochronology and geomatics to monitor natural hazards and landscape changes. *Applied Geomatics*, 11(1):39–52, 2019. doi: 10.1007/s12518-018-0236-0.

- [23] J.S. Clark, M. Silman, R. Kern, E. Macklin, and J. Hillerislambers. Seed dispersal near and far: Patterns across temperate and tropical forests. *Ecology*, 80(5): 1475–1494, 1999. doi: 10.1890/0012-9658(1999)080[1475:SDNAFP]2.o.CO;2. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-0032855350&doi=10.1890%2F0012-9658%281999%29080%5b1475%3aSDNAFP%5d2.0.C0%3b2&partnerID=40&md5=54f9337d8cb0e819df38eddb3879f751>. cited By 659.
- [24] S.A.O. Cousins, A.G. Auffret, J. Lindgren, and L. Trank. Regional-scale land-cover change during the 20th century and its consequences for biodiversity. *Ambio*, 44(1):17–27, 2015. doi: 10.1007/s13280-014-0585-9. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84920837313&doi=10.1007%2Fs13280-014-0585-9&partnerID=40&md5=18d8318a829169c9d0bf4dd7855e4f42>.
- [25] Andrew Crooks and Christian Castle. *The Integration of Agent-Based Modelling and Geographical Information for Geospatial Simulation*, pages 219–251. Springer, 01 2012. ISBN 978-90-481-8926-7. doi: 10.1007/978-90-481-8927-4\_12.
- [26] Lydie Didier. Invasion patterns of european larch and swiss stone pine in subalpine pastures in the french alps. *Forest Ecology and Management*, 145(1):67–77, 2001. ISSN 0378-1127. doi: [https://doi.org/10.1016/S0378-1127\(00\)00575-2](https://doi.org/10.1016/S0378-1127(00)00575-2). URL <https://www.sciencedirect.com/science/article/pii/S0378112700005752>. Structure of Mountain Forests-Assessment, Impacts, Managements, Modelling.
- [27] European Soil Data Centre (ESDAC). RUSLE for European Union. <https://esdac.jrc.ec.europa.eu/themes/rusle2015>, 2015.
- [28] A. Falcucci, L. Maiorano, and L. Boitani. Changes in land-use/land-cover patterns in italy and their implications for biodiversity conservation. *Landscape Ecology*, 22:81–107, 2007. doi: <https://doi.org/10.1007/s10980-006-9056-4>.
- [29] Jan Feranec, Gerard Hazeu, Susan Christensen, and Gabriel Jaffrain. Corine land cover change detection in europe (case studies of the netherlands and slovakia). *Land Use Policy*, 24(1):234–247, 2007. ISSN 0264-8377. doi: <https://doi.org/10.1016/j.landusepol.2006.02.002>. URL <https://www.sciencedirect.com/science/article/pii/S0264837706000068>.
- [30] F. Ferretti, C. Sboarina, C. Tattoni, A. Vitti, P. Zatelli, F. Geri, E. Pompei, and M. Ciolli. Italian kingdom forest map webgis. <http://carta1936.dicam.unitn.it>, 2018. Accessed: 2019-07-25.
- [31] F. Ferretti, C. Sboarina, C. Tattoni, A. Vitti, P. Zatelli, F. Geri, E. Pompei, and M. Ciolli. The 1936 italian kingdom forest map reviewed: A dataset for landscape and ecological research. *Annals of Silvicultural Research*, 42(1):3–19, 2018. doi: 10.12899/asr-1411.
- [32] A. Feurdean, C. Munteanu, T. Kuemmerle, A.B. Nielsen, S.M. Hutchinson, E. Ruprecht, C.L. Parr, A. Persoiu, and T. Hickler. Long-term land-cover/use change in a traditional farming landscape in romania inferred from pollen data, historical maps and satellite images. *Regional Environmental Change*, 17(8):2193–2207, 2017. doi: 10.1007/s10113-016-1063-7.
- [33] G. Forzieri, M. Pecchi, M. Girardello, A. Mauri, M. Klaus, C. Nikolov, M. Ruetschi, B. Gardiner, J. Tomaštk, D. Small, C. Nistor, D. Jonikavicius, J. Spinoni, L. Feyen, F. Giannetti, R. Comino, A. Wolynski, F. Pirotti, F. Maistrelli, I. Savulescu, S. Wurpillot-Lucas, S. Karlsson, K. Zieba-Kulawik, P. Strejczek-Jazwinska, M. Mokroš, S. Franz, L. Krejci, I. Haidu, M. Nilsson, P. Wezyk, F. Catani, Y.-Y. Chen, S. Luysaert, G. Chirici,

- A. Cescatti, and P. S. A. Beck. A spatially explicit database of wind disturbances in european forests over the period 2000–2018. *Earth System Science Data*, 12(1):257–276, 2020. doi: 10.5194/essd-12-257-2020. URL <https://essd.copernicus.org/articles/12/257/2020/>.
- [34] Apostolos Gaitanis, Kleomenis Kalogeropoulos, Vassilis Detsis, and Christos Chalkias. Monitoring 60 years of land cover change in the Marathon area, Greece. *Land*, 4(2): 337–354, 2015. ISSN 2073445X. doi: 10.3390/land4020337.
- [35] Cristina Gandolfo and Paola Comin. *Relazione sull’attività svolta nell’anno 2016*. Provincia Autonoma di Trento, servizio foreste e fauna, Trento, 2016. URL <http://www.forestefauna.provincia.tn.it>.
- [36] Yan Gao and Jean Mas. A comparison of the performance of pixel based and object based classifications over images with various spatial resolutions. *Online Journal of Earth Science*, 2:27–35, 01 2008.
- [37] Geoportale Provincia Autonoma di Trento. Metadata: Aree agricole. SCARICAMENTO/CARTOGRAFIA/SIS\_AGR/IMAGES/uni10\_tif\_nn\_comp.asp, 2008. URL <http://www.urbanistica-dati.provincia.tn.it>.
- [38] Geoportale Provincia Autonoma di Trento. Metadata: Aree bosco. [https://idt.provincia.tn.it/idt/vector/p\\_TN\\_f6bbcf22-8363-4438-a2cd-bce589eda3d9.zip](https://idt.provincia.tn.it/idt/vector/p_TN_f6bbcf22-8363-4438-a2cd-bce589eda3d9.zip), 2008. URL <http://www.urbanistica-dati.provincia.tn.it>.
- [39] Geoportale Provincia Autonoma di Trento. Metadata: Tipi forestali. [https://idt.provincia.tn.it/idt/vector/p\\_TN\\_dece9469-83b6-4cf7-85fe-71c5a60d0999.zip](https://idt.provincia.tn.it/idt/vector/p_TN_dece9469-83b6-4cf7-85fe-71c5a60d0999.zip), 2009.
- [40] Geoportale Provincia Autonoma di Trento. Metadata: DTM. <http://www.territorio.provincia.tn.it/portal/server.pt/community/lidar/847/lidar/23954>, 2016.
- [41] Geoportale Regione Lombardia. Metadata: Volo GAI 1954. Geoportale regione Lombardia, <http://www.geoportale.regione.lombardia.it/metadati> (2018), 2018.
- [42] Giornale del Trentino. Vaia in trentino, stimati danni per 47 milioni di euro. <https://www.giornaletrentino.it/cronaca/vaia-in-trentino-stimati-danni-per-47-milioni-di-euro-1.2118039>, 2019.
- [43] Giovanni Giovannini. *Paesaggi Agro-Forestali in Trentino tutela, ripristino e miglioramento degli ambienti tradizionali*. Provincia Autonoma di Trento, servizio foreste e fauna, Trento, 2017. ISBN 978-88-7702-433-6.
- [44] S. Gobbi, G. Maimeri, C. Tattoni, M.G. Cantiani, D. Rocchini, N. La Porta, M. Ciolli, and P. Zatelli. Orthorectification of a large dataset of historical aerial images: Procedure and precision assessment in an open source environment. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, volume 42 4W8, pages 53–59, 2018. doi: 10.5194/isprs-archives-XLII-4-W8-53-2018.
- [45] S. Gobbi, M.G. Cantiani, D. Rocchini, P. Zatelli, C. Tattoni, M. Ciolli, and N. La Porta. Fine spatial scale modelling of Trentino past forest landscape (Trentinoland): a case study of FOSS application. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, volume 42 4W14, pages 71–78, 26–30 August 2019, Bucharest, Romania, 2019. doi: 10.5194/isprs-archives-XLII-4-W14-71-2019.

- [46] Stefano Gobbi, Marco Ciolli, Nicola La Porta, Duccio Rocchini, Clara Tattoni, and Paolo Zatelli. New tools for the classification and filtering of historical maps. *ISPRS International Journal of Geo-Information*, 8(10), 2019. ISSN 2220-9964. doi: 10.3390/ijgi8100455. URL <https://www.mdpi.com/2220-9964/8/10/455>.
- [47] GRASS Development Team. Geographic Resources Analysis Support System (GRASS GIS) Web Site. <http://grass.osgeo.org>, 2019. Accessed: 2019-04-29.
- [48] Arnaud Grignard, Patrick Taillandier, Benoit Gaudou, Duc An Vo, Nghi Quang Huynh, and Alexis Drogoul. Gama 1.6: Advancing the art of complex agent-based modeling and simulation. In Guido Boella, Edith Elkind, Bastin Tony Roy Savarimuthu, Frank Dignum, and Martin K. Purvis, editors, *PRIMA 2013: Principles and Practice of Multi-Agent Systems*, pages 117–131, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg. ISBN 978-3-642-44927-7.
- [49] V. Grimm and S.F. Railsback. *Individual-Based Modeling and Ecology*. Princeton university press, 2013.
- [50] T. Grippa, M. Lennert, B. Beaumont, S. Vanhuysse, N. Stephenne, and E. Wolff. An open-source semi-automated processing chain for urban object-based classification. *Remote Sensing*, 9(4), 2017. doi: 10.3390/rs9040358.
- [51] Robert M. Haralick, K. Shanmugam, and Its’Hak Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621, 1973. doi: 10.1109/TSMC.1973.4309314.
- [52] Drew Heiderscheidt. The impact of world war one on the forests and soils of europe. *Ursidae: The Undergraduate Research Journal at the University of Northern Colorado*, 7(3), 2018.
- [53] Maximilian H.K. Hesselbarth, Marco Sciaini, Kimberly A. With, Kerstin Wiegand, and Jakub Nowosad. landscapemetrics: an open-source R tool to calculate landscape metrics. *Ecography*, 42(10):1648–1657, 2019. ISSN 16000587. doi: 10.1111/ecog.04617.
- [54] Danielle Hoja, Mathias Schneider, Rupert Mueller, Manfred Lehner, and Peter Reinartz De. Comparison of Orthorectification Methods Suitable for Rapid Mapping Using Direct Georeferencing and Rpc for Optical Satellite Data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII(Part B4): 1617–1624, 2008.
- [55] ISTAT. Population of the PAT from 1921 to 2011. <https://www.tuttitalia.it/trentino-alto-adige/provincia-autonoma-di-trento/statistiche/censimenti-popolazione/>, 2011.
- [56] ISTAT. Altitude ranges for Italian municipality. <https://www4.istat.it/it/archivio/156224>, 2018.
- [57] Brian Johnson and Zhixiao Xie. Unsupervised image segmentation evaluation and refinement using a multi-scale approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4):473–483, 2011. ISSN 09242716. doi: 10.1016/j.isprsjprs.2011.02.006. URL <http://dx.doi.org/10.1016/j.isprsjprs.2011.02.006>.
- [58] Brian Johnson, Milben Bragais, Isao Endo, Damasa Magcale-Macandog, and Paula Macandog. Image Segmentation Parameter Optimization Considering Within- and Between-Segment Heterogeneity at Multiple Scale Levels: Test Case for Mapping Residential Areas Using Landsat Imagery. *ISPRS International Journal of Geo-Information*, 4(4):2292–2305, 2015. doi: 10.3390/ijgi4042292.

- [59] Jed O Kaplan, Kristen M Krumhardt, and Niklaus Zimmermann. The prehistoric and preindustrial deforestation of europe. *Quaternary Science Reviews*, 28(27-28):3016–3034, 2009.
- [60] Andrej Kobler, Tomaz Cunder, and Janez Pirnat. Modelling spontaneous afforestation in postojna area, slovenia. *Journal for Nature Conservation*, 13(2):127–135, 2005. ISSN 1617-1381. doi: <https://doi.org/10.1016/j.jnc.2005.01.003>. URL <https://www.sciencedirect.com/science/article/pii/S1617138105000166>.
- [61] Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, and Laura Bieker. Recent Development and Applications of SUMO - Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements*, 5(3):128–138, 2012. ISSN 1942-261X. URL <http://elib.dlr.de/80483/>.
- [62] Amnon Krupnik. Accuracy prediction for ortho-image generation. *Photogrammetric Record*, 18(101):41–58, 2003. ISSN 0031868X. doi: 10.1111/0031-868X.t01-1-00001.
- [63] Kuhn, M. The R caret library. <https://topepo.github.io/caret/>, 2019. accessed on 25 July 2019.
- [64] Van-Minh Le, Yann Chevalyere, Tuong Ho, and Jean-daniel Zucker. Hybrid of linear programming and genetic algorithm for optimizing agent-based simulation. application to optimization of sign placement for tsunami evacuation. *The 2015 IEEE RIVF International Conference on Computing and Communication Technologies Research, Innovation, and Vision for Future (RIVF)*, pages 138–143, 02 2015. doi: 10.1109/RIVF.2015.7049889.
- [65] Lennert, M. GRASS Development Team Addon i.segment.uspo. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6. <https://grass.osgeo.org/grass72/manuals/addons/i.segment.uspo.html>, 2016. accessed on 25 July 2019.
- [66] Lennert, M. GRASS Development Team Addon v.class.mlR. Geographic Resources Analysis Support System (GRASS) Software, Version 7.3. <https://grass.osgeo.org/grass70/manuals/addons/v.class.mlR.html>, 2017. accessed on 25 July 2019.
- [67] Xiaoxiao Li and Guofan Shao. Object-based land-cover mapping with high resolution aerial photography at a county scale in midwestern usa. *Remote Sensing*, 6(11):11372–11390, 2014. ISSN 2072-4292. doi: 10.3390/rs61111372. URL <https://www.mdpi.com/2072-4292/6/11/11372>.
- [68] T.M. Lillesand, R.W. Kiefer, and J.W. Chipman. *Remote sensing and image interpretation*. Wiley, 2004.
- [69] Yi-Fan Liu, Yu Liu, Zhi-Hua Shi, Manuel Lopez-Vicente, and Gao-Lin Wu. Effectiveness of re-vegetated forest and grassland on soil erosion control in the semi-arid loess plateau. *CATENA*, 195:104787, 2020. ISSN 0341-8162. doi: <https://doi.org/10.1016/j.catena.2020.104787>. URL <https://www.sciencedirect.com/science/article/pii/S0341816220303374>.
- [70] Martin Loidl, Christoph Traun, and Gudrun Wallentin. Spatial patterns and temporal dynamics of urban bicycle crashes - a case study from salzburg (austria). *Journal of Transport Geography*, 52:38–50, 2016. ISSN 0966-6923. doi: <https://doi.org/10.1016/j.jtrangeo.2016.02.008>. URL <https://www.sciencedirect.com/science/article/pii/S0966692316000302>.

- [71] D MacDonald, J.R Crabtree, G Wiesinger, T Dax, N Stamou, P Fleury, J Gutierrez Lazpita, and A Gibon. Agricultural abandonment in mountain areas of europe: Environmental consequences and policy response. *Journal of Environmental Management*, 59(1):47–69, 2000. ISSN 0301-4797. doi: <https://doi.org/10.1006/jema.1999.0335>. URL <https://www.sciencedirect.com/science/article/pii/S030147979903353>.
- [72] Giulia Maimeri. *Variazione della copertura forestale in val di Fassa tramite analisi GIS multi-temporale*. University of Trento, master's degree thesis, 2018.
- [73] Somsakun Maneerat and Eric Daude. A spatial agent-based simulation model of the dengue vector *aedes aegypti* to explore its population dynamics in urban areas. *Ecological Modelling*, 333:66–78, 2016. ISSN 0304-3800. doi: <https://doi.org/10.1016/j.ecolmodel.2016.04.012>. URL <https://www.sciencedirect.com/science/article/pii/S0304380016301387>.
- [74] A.E. Maxwell, M.P. Strager, T.A. Warner, C.A. Ramezan, A.N. Morgan, and C.E. Pauley. Large-area, high spatial resolution land cover mapping using random forests, geobia, and naip orthophotography: Findings and recommendations. *Remote Sensing*, 11(12), 2019. doi: 10.3390/rs11121409.
- [75] K. McGarigal. FRAGSTATS HELP. <https://www.umass.edu/landeco/research/fragstats/>, 2015. Accessed: 2019-09-12.
- [76] K. McGarigal, SA Cushman, and E. Ene. FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. <http://www.umass.edu/landeco/research/fragstats/fragstats.html>, 2012. Accessed: 2019-09-12.
- [77] Metz, M. GRASS Development Team g.gui.image2target. Geographic Resources Analysis Support System (GRASS) Software, Version 7.3. <https://grass.osgeo.org/grass70/manuals/addons/v.class.mlR.html>, 2017. accessed on 25 July 2019.
- [78] Metz, M. and Lennert, M. GRASS Development Team Addon r.object.geometry. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6. <https://grass.osgeo.org/grass72/manuals/addons/r.object.geometry.html>, 2016. accessed on 25 July 2019.
- [79] Metz, M. and Lennert, M. GRASS Development Team Addon i.segment.stats. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6. <https://grass.osgeo.org/grass76/manuals/addons/i.segment.stats.html>, 2016. accessed on 25 July 2019.
- [80] Momsen, E. and Metz, M. GRASS Development Team i.segment. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6. <https://grass.osgeo.org/grass76/manuals/i.segment.html>, 2016. accessed on 25 July 2019.
- [81] Renzo Motta, Davide Ascoli, Piermaria Corona, Marco Marchetti, and Giorgio Vacchiano. Silviculture and wind damages. the storm "vaia". *Forest - Rivista di Selvicoltura ed Ecologia Forestale*, 15:94–98, 12 2018. doi: 10.3832/efor2990-015.
- [82] Netler, M. GRASS Development Team v.db.join. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6.



- <https://grass.osgeo.org/grass76/manuals/v.db.join.html>, 2019. accessed on 25 July 2019.
- [83] M. Odasso. *Tipi forestali in Trentino*, volume 192. Trento, centro di Ecologia Alpina, 2002.
- [84] M. Odasso, M. Miori, and C. Gandolfo. *I tipi forestali del Trentino: descrizione e aspetti dinamici*. Provincia Autonoma di Trento, Servizio Fauna e Foreste, 2018. ISBN 978-88-7702-460-2.
- [85] OpenStreetMap. OpensStreet Map for the Italian Territory. <https://www.openstreetmap.org/copyright>, 2021.
- [86] Francesco Orsi, Marco Ciolli, Eeva Primmer, Liisa Varumo, and Davide Geneletti. Mapping hotspots and bundles of forest ecosystem services across the european union. *Land Use Policy*, 99:104840, 2020. ISSN 0264-8377. doi: <https://doi.org/10.1016/j.landusepol.2020.104840>. URL <https://www.sciencedirect.com/science/article/pii/S0264837719313547>.
- [87] P. Piussi. Carta del limite potenziale del bosco in trentino. Technical report, Servizio Foreste Caccia e Pesca, Provincia Autonoma di Trento, 1992.
- [88] Provincia Autonoma di Trento. Aree protette della provincia di trento. <http://www.areeprotette.provincia.tn.it/>, 2018. Accessed: 2021-7-14.
- [89] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL <https://www.R-project.org/>.
- [90] J Radoux and P Bogaert. Good practices for object-based accuracy assessment. *Remote Sensing*, 9(7), 2017. doi: 10.3390/rs9070646. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85022338740&doi=10.3390/rs9070646&partnerID=40&md5=cb8876f57fa4cf28f3c99a320eac6b9>.
- [91] Revolti, R. Produzione e restauro della cartografia catastale in Trentino. <https://drive.google.com/file/d/oBoabPZ1IrxqdcnUwXzB1T2RvVoE/view>, 2017. Accessed: 2019-04-29.
- [92] Duccio Rocchini, Markus Metz, Alessandro Frigeri, Luca Delucchi, Matteo Marcantonio, and Markus Neteler. Robust rectification of aerial photographs in an open source environment. *Computers and Geosciences*, 39:145–151, 02 2012. doi: 10.1016/j.cageo.2011.06.002.
- [93] M.D.A. Rounsevell, I. Reginster, M.B. Araujo, T.R. Carter, N. Dendoncker, F. Ewert, J.I. House, S. Kankaanpaa, R. Leemans, M.J. Metzger, C. Schmit, P. Smith, and G. Tuck. A coherent set of future land use change scenarios for europe. *Agriculture, Ecosystems and Environment*, 114(1):57–68, 2006. ISSN 0167-8809. doi: <https://doi.org/10.1016/j.agee.2005.11.027>. URL <https://www.sciencedirect.com/science/article/pii/S0167880905005347>. Scenario-Based Studies of Future Land Use in Europe.
- [94] J Schiewe. Segmentation of high-resolution remotely sensed data-concepts, applications and problems. *International Archive Photogrammetry Remote Sensing Spatial Information Science*, 34:380–385, 2002.
- [95] Nina Schneeberger, Matthias Burgi, and P.D. Felix Kienast. Rates of landscape change at the northern fringe of the swiss alps: Historical and recent tendencies. *Landscape and Urban Planning*, 80(1):127–136, 2007. ISSN 0169-2046. doi: <https://doi.org/10.1016/>

- j.landurbplan.2006.06.006. URL <https://www.sciencedirect.com/science/article/pii/S0169204606001125>.
- [96] P. Serra, X. Pons, and D. Sauri. Land-cover and land-use change in a mediterranean landscape: A spatial analysis of driving forces integrating biophysical and human factors. *Applied Geography*, 28(3):189–209, 2008. ISSN 0143-6228. doi: <https://doi.org/10.1016/j.apgeog.2008.02.001>. URL <https://www.sciencedirect.com/science/article/pii/S014362280800012X>.
- [97] Servizio Catasto della Provincia Autonoma di Trento. History of the cadaster historical map for the Province of Trento. [http://www.catasto.provincia.tn.it/cenni\\_storici/pagina8.html](http://www.catasto.provincia.tn.it/cenni_storici/pagina8.html), 2019. Accessed: 2019-04-29.
- [98] Servizio Catasto della Provincia Autonoma di Trento. Relazioni del Servizio Catasto della Provincia di Trento (Report of the Cadaster Service of the Province of Trento). <http://www.tpg-tnbz.it/newseventi/relazioni-del-servizio-catasto-della-provincia-di-trento>, 2019. Accessed: 2019-04-29.
- [99] Servizio Catasto della Provincia Autonoma di Trento. HistoricalKat, documentation of the Trentino cadastral archive. <http://historicalkat.provincia.tn.it>, 2019. Accessed: 2019-04-29.
- [100] T. Sitzia. *Ecologia e gestione dei boschi di neoformazione nel paesaggio Trentino*. Servizio Foreste e Fauna, Provincia Autonoma di Trento, Trento, 2009.
- [101] T. Sitzia, P. Semenzato, and G. Trentanovi. Natural reforestation is changing spatial patterns of rural mountain and hill landscapes: A global overview. *Forest Ecology and Management*, 259(8):1354–1362, 2010. doi: 10.1016/j.foreco.2010.01.048.
- [102] Franck Taillandier and Carole Adam. Games ready to use: A serious game for teaching natural risk management. *Simulation & Gaming*, 49(4):441–470, 2018. doi: 10.1177/1046878118770217. URL <https://doi.org/10.1177/1046878118770217>.
- [103] Franck Taillandier, Patrick Taillandier, Esra Tepeli, Denys Breyse, Rasool Mehdizadeh, and Fadi Khartabil. A multi-agent model to manage risks in construction project (smacc). *Automation in Construction*, 58:1–18, 2015. ISSN 0926-5805. doi: <https://doi.org/10.1016/j.autcon.2015.06.005>. URL <https://www.sciencedirect.com/science/article/pii/S0926580515001235>.
- [104] Patrick Taillandier, Benoit Gaudou, Arnaud Grignard, Quang Nghi Huynh, Nicolas Marilleau, Philippe Caillou, Damien Philippon, and Alexis Drogoul. Building, composing and experimenting complex spatial models with the GAMA platform. *GeoInformatica*, 23(2):299–322, 2019. ISSN 13846175. doi: 10.1007/s10707-018-00339-6.
- [105] Ulrike Tappeiner and Alexander Cernusca. Alpine meadows and pastures after abandonment. results of the austrian mab-programme and the ec-step project integralp. *Pirineos*, 141:97–118, 1993.
- [106] Erich Tasser and Ulrike Tappeiner. Impact of land use changes on mountain vegetation. *Applied Vegetation Science*, 5(2):173–184, 2002. doi: <https://doi.org/10.1111/j.1654-109X.2002.tb00547.x>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1654-109X.2002.tb00547.x>.

- [107] Erich Tasser, Janette Walde, Ulrike Tappeiner, Alexandra Teutsch, and Werner Nogglér. Land-use changes and natural reforestation in the eastern central alps. *Agriculture, Ecosystems & Environment*, 118(1):115–129, 2007. ISSN 0167-8809. doi: <https://doi.org/10.1016/j.agee.2006.05.004>. URL <https://www.sciencedirect.com/science/article/pii/S0167880906001575>.
- [108] C. Tattoni, M. Ciolli, F. Ferretti, and M.G. Cantiani. Monitoring spatial and temporal pattern of paneveggio forest (northern italy) from 1859 to 2006. *IForest*, 3(MAY):72–80, 2010. doi: 10.3832/ifor0530-003.
- [109] C. Tattoni, M. Ciolli, and F. Ferretti. The fate of priority areas for conservation in protected areas: A fine-scale markov chain approach. *Environmental Management*, 47(2): 263–278, 2011. doi: 10.1007/s00267-010-9601-4.
- [110] C. Tattoni, E. Ianni, D. Geneletti, P. Zatelli, and M. Ciolli. Landscape changes, traditional ecological knowledge and future scenarios in the alps: A holistic ecological approach. *Science of the Total Environment*, 579:27–36, 2017. doi: 10.1016/j.scitotenv.2016.11.075.
- [111] C Tattoni, E Soardi, F Prosser, M Odasso, P Zatelli, and M. Ciolli. Fruit availability for migratory birds: a gis approach. *PeerJ*, 7, 2019. doi: 10.7717/peerj.6394.
- [112] Clara Tattoni, Marco Ciolli, F Ferretti, and Maria Giulia Cantiani. Monitoring spatial and temporal pattern of Paneveggio forest (northern Italy) from 1859 to 2006. *IFOREST*, 3(3):72–80, 2010. doi: 10.3832/ifor0530-003.
- [113] Clara Tattoni, Gianluca Grilli, Jorge Arana, and Marco Ciolli. The landscape change in the alps - what postcards have to say about aesthetic preference. *Sustainability*, 13 (13), 2021. ISSN 2071-1050. doi: 10.3390/su13137426. URL <https://www.mdpi.com/2071-1050/13/13/7426>.
- [114] UN. 68% of the world population projected to live in urban areas by 2050, says un. <https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html>, 2018.
- [115] V. Vaithyanathan and U. Rajappa. A review on clustering techniques in image segmentation. *International Journal of Applied Engineering Research*, 8(20 SPEC. ISSUE):2685–2688, 2013. ISSN 09739769.
- [116] Gudrun Wallentin. Spatial simulation: A spatial perspective on individual-based ecology, a review. *Ecological Modelling*, 350:30–41, 2017. ISSN 0304-3800. doi: <https://doi.org/10.1016/j.ecolmodel.2017.01.017>. URL <https://www.sciencedirect.com/science/article/pii/S0304380016306640>.
- [117] Gudrun Wallentin and Christian Neuwirth. Dynamic hybrid modelling: Switching between ab and sd designs of a predator-prey model. *Ecological Modelling*, 345:165–175, 2017. ISSN 0304-3800. doi: <https://doi.org/10.1016/j.ecolmodel.2016.11.007>. URL <https://www.sciencedirect.com/science/article/pii/S0304380016303714>.
- [118] Gudrun Wallentin, Ulrike Tappeiner, Josef Strobl, and Erich Tasser. Understanding alpine tree line dynamics: An individual-based model. *Ecological Modelling*, 218(3):235–246, 2008. ISSN 0304-3800. doi: <https://doi.org/10.1016/j.ecolmodel.2008.07.005>. URL <https://www.sciencedirect.com/science/article/pii/S0304380008003451>.
- [119] Wallentin, G. UNIGIS module: Spatial Simulation. [https://unigis-salzburg.github.io/Opt\\_Spatial-Simulation](https://unigis-salzburg.github.io/Opt_Spatial-Simulation), 2020.

- [120] Pierre Walther. Land abandonment in the swiss alps: a new understanding of a land-use problem. *Mountain Research and Development*, pages 305–314, 1986.
- [121] Wen, T. GRASS Development Team r.kappa. Geographic Resources Analysis Support System (GRASS) Software, Version 7.6. <https://grass.osgeo.org/grass74/manuals/r.kappa>, 2019. accessed on 25 July 2019.
- [122] WG Wilson, AM Deroos, and E McCauley. Spatial instabilities within the diffusive lotka-volterra system: individual-based simulation results. *Theoretical population biology*, 43(1):91–127, 1993.
- [123] Walter H Wischmeier and Dwight David Smith. *Predicting rainfall-erosion losses from cropland east of the Rocky Mountains: Guide for selection of practices for soil and water conservation*. Agricultural Research Service, US Department of Agriculture, 1965.
- [124] Annalisa Zanella, Clara Tattoni, and Marco Ciolli. Studio della variazione temporale della quantità e qualità del bestiame nel parco di paneveggio pale di san martino e influenza sui cambiamenti del paesaggio forestale. *Dendronatura*, 1:24–33, 01 2010.
- [125] P. Zatelli, S. Gobbi, C. Tattoni, N. La Porta, and M. Ciolli. Object-based image analysis for historic maps classification. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, volume 42 4W14, pages 247–254, 2019. doi: 10.5194/isprs-archives-XLII-4-W14-247-2019.
- [126] Paolo Zatelli, Stefano Gobbi, Clara Tattoni, Maria Giulia Cantiani, Nicola La Porta, Duccio Rocchini, Nicola Zorzi, and Marco Ciolli. Relevance of the cell neighborhood size in landscape metrics evaluation and free or open source software implementations. *ISPRS International Journal of Geo-Information*, 8(12), 2019. ISSN 2220-9964. doi: 10.3390/ijgi8120586. URL <https://www.mdpi.com/2220-9964/8/12/586>.

## ACKNOWLEDGEMENTS

This has been a long and exhausting journey. While you are a student you cannot imagine what it means carrying out and finishing a PhD. And never I ever would be able to finish without the help of all the people around me. Sharing an achievement or being helped in a tough moment requires both to be present for someone, and I cannot share strongly enough my gratitude to everyone who was there for me.

Firstly, I would like to thank my supervisors, for their incredible patience and dedication and all the passion they were able to transmit to me over this four-year-long cooperation. A big thank you to Prof. Paolo Zatelli, who followed me and showed me a universe made of maps, images and reference systems, and never lost his nerve, regardless the situation.

I cannot stress enough how grateful I am to you, my incredible Fiancé who patiently helped me, and stayed on my side through good and dark times, who awaited for me every night I could not be with her, who cheered me up, who shared with me countless moments when I felt overwhelmed and overjoyed. She is the reason, the fire which kept me walking toward the end of this path. Using her own words "Onaha means family, and family means no one is get left behind or forgotten", and she made me feel every time like we are a family.

Silently waiting for me, each time I needed or wanted a rest from the university and the research world they were, my parents, who never failed to be on my side. They are my models, and a lifetime thanking them would not be sufficient to express what I feel in this moment. To appreciate forests in Trentino, you have to live them, and what better way to live a place than to explore the less travelled path, or see things from a different prospective? The Sengio Rosso crew, incredible people and incredible climbers, companions of three years of climbing and adventuring, I would like to shout a big thank you to all of them. Finally, thanks to everyone made this four year long path an adventure worth living.

Over recent decades, forest land cover is dramatically changing in European mountains and in the Alps in particular. Since the 1950s the progressive urbanization of the valleys and the abandonment of mountain and rural dwellers has intensified. More than 60% of the Trentino land, is covered by forest and mainly by high forest. This human migration have brought to a progressive shrinking of meadows and pastures due to the natural forest expansion causing a dramatic change in the landscape, the consequences of which affect biodiversity, social and cultural dynamics and landscape perception as well as ecosystem services. The objective of this research focused on the application and experimentation of advanced GIS and modeling techniques to compare aerial imagery, historical maps and data and remote sensed images to understand the past landscape changes and their dynamics in Trentino and to build future scenarios based on long-term set of observations. The research produced a fine scale dataset representing past forest landscape for the Trentino territory. The analysis of these output data revealed a progressive afforestation process which interested homogeneously all the Trentino territory. A future forest landscape scenarios at a detailed scale (10 m) was as well produced, to simulate the future of the forest in a protected area of Trentino, to outline if the afforestation process will continue. Along with these main output of the research, new tools for image processing and evaluation of forest changes were developed.

Stefano Gobbi is a PhD student at the Department of Civil, Environmental and Mechanical Engineering (University of Trento, Italy) where he received his MSc in environmental engineering (2017). After the degree his career as PhD student revolved around the study of the history of forest in Provincia Autonoma di Trento, especially from a remote sensing point of view. His research interest include numerical cartography, remote sensing, forest management GIS tools development. He published several articles on these topics, in particular on the development of different methods for aerial imagery and map processing.