UNIVERSITÀ DEGLI STUDI DI SASSARI

CORSO DI DOTTORATO DI RICERCA

Scienze Agrarie

Curriculum Agrometeorologia ed Ecofisiologia dei Sistemi Agrari e Forestali

Ciclo XXX

Coupling remote sensing with wildfire spread modeling in Mediterranean areas

Olga Muñoz Lozano

Coordinatore del Corso	Prof. Ignazio Floris
Referente di Curriculum	Prof. Maurizio Mulas
Docente Guida	Prof. Donatella Spano
Docente Tutor	Dr. Michele Salis

ANNO ACCADEMICO 2016 - 2017

UNIVERSITÀ DEGLI STUDI DI SASSARI

CORSO DI DOTTORATO DI RICERCA

Scienze Agrarie

Curriculum Agrometeorologia ed Ecofisiologia dei Sistemi Agrari e Forestali

Ciclo XXX

Coupling remote sensing with wildfire spread modeling in Mediterranean areas

Olga Muñoz Lozano

Coordinatore del Corso	Prof. Ignazio Floris
Referente di Curriculum	Prof. Maurizio Mulas
Docente Guida	Prof. Donatella Spano
Docente Tutor	Dr. Michele Salis

ANNO ACCADEMICO 2016 - 2017

Index

List of Abbreviations

AAD	Absolute Average Deviation
AGEA	AGenzia per le Erogazioni in Agricoltura (Agricultural Supplies Agency)
AIC	Akaike Information Criteria
ASCII	American Standard Code for Information Interchange
CBD	Canopy Bulk Density
CBH	Canopy Base Height
CFR	Crown Fire Activity
CUS	Custom
CV	Coefficient of Variation
DBH	diameter at breast height
DEM	Digital Elevation Model
EEA	European Environmental Agency
Elev	Elevation
FLI	Fireline Intensity
FMC	dead fuel moisture content
FML	Flame Length
H1, H2 and H3	first, second and third strata of mean height respectively
HPA	Heat per unit Area
IQ	InterQuartile
K	Cohen's Kappa coefficient
LCP	Landscape file
LiDAR	Light Detection And Ranging
LMM	Linear Mixed Models
LUM	Land Use Map
MAD	Median of the Absolute Deviations
ML	Maximum likelihood

MTT	Minimum Travel Time
NDII	Normalized Difference Infrared Index
NDVI	Normalized Difference Vegetation Index
NIR	Near InfraRed
NN	Neural Networks
OA	Overall Accuracy
RCI	Reaction Intensity
RF	Random Forests
ROS	Rate of Spread
SAVI	Soil Adjusted Vegetation Index
SC	Sørensen's coeffcient
SD1, SD2 and SD3	first, second and third strata of standard deviation respectively
SDR	Spread Direction
SH	Stand Height
ST	Standard
SVM	Support Vector Machine
SWIR	Short Wave InfraRed
TOA	Time of Arrival
UTM	Universal Transverse of Mercator
WGS-84	World Geodetic System 1984

Abstract

Wildfires are a threat to the ecosystems and in the future this threat could become stronger due to climate change. Spatially explicit fire spread models are effective tools to study fire behavior and wildfire risk. However, to run fire spread simulations, one of the most important inputs is represented by fuel models and this information is not always available. In the last decades, remote sensing technologies have offered valuable information for the classification and characterization of fuels. For this reason, in this work we created accurate maps of main fuel types for Mediterranean areas combining multispectral and LiDAR data. This information improves the current available information, which derives from the Land Use Map of Sardinia. We also enhanced the characterization of canopy fuel models using LiDAR data producing canopy layers ready to be used for wildfire spread modeling. Finally, we compared the variation in simulated wildfire spread and behavior determined by the use of fine-scale maps v. lower resolution maps. In these simulations, we assessed also the effect of using LiDAR-derived canopy layers as well. The results showed more accurate outputs when using our custom fuel and canopy layers produced in this work. In conclusion, this work suggests that the use of LiDAR and satellite imagery data can contribute to improve estimates of modeled wildfire behavior.

Introduction

Wildfires threat ecosystems worldwide (Cole and Landres 1996; Pausas and Keeley 2009) and especially in Mediterranean areas (Syphard *et al.* 2009). For instance, in 2016 in Europe the 68% of fires occurred in Mediterranean countries and these fires accounted for 93% of the overall burned area of Europe (San-Miguel-Ayanz *et al.* 2017).

In the future, wildfires impacts on ecosystems are expected to increase due to climate change (mainly higher temperatures and more frequent heat waves, Flannigan *et al.* 2000, 2009; Liu *et al.* 2010; Arca *et al.* 2012; Kovats *et al.* 2014; Kurnik *et al.* 2017; Lozano *et al.* 2017). Studies carried out to investigate the effects of recent warming on wildfire season length and behavior confirmed that higher temperatures bring with them prolonged fire seasons and longer fire events, as well as more fire ignitions and larger fires (Piñol *et al.* 1998; Westerling *et al.* 2006; Turco *et al.* 2014; Urrutia-Jalabert *et al.* 2018). Indeed, (Kovats *et al.* 2014) predicted warmer and drier weather (especially in summer) with increased fire risk in Mediterranean areas.

In a context of likely increase of wildfire-derived damages (particularly from large fires, which account for the most of burned area although limited in number), the advances in fire behavior analysis and in the assessment of fire risk will play a key role for fire management and research.

A number of papers reported that spatially explicit fire spread models are effective tools to study wildfire behavior and risk (e.g.: Calkin et al. 2011; Miller and Ager 2013). Most of fire spread models are based on physical principles and empirical observations (Duff *et al.* 2013). The use of fire simulators was proved to be an effective and powerful tool not only for Northern America ecosystems, (Thompson *et al.* 2011; Ager *et al.* 2013, 2014) but also in the Mediterranean area (Salis *et al.* 2014; Alcasena *et al.* 2015; Kalabokidis *et al.* 2015).

To run fire spread simulations, one of the most important inputs is represented by fuel distribution. The spread of fire is affected for some fuel factors such as crown bulk density, crown base height, canopy height, percent of canopy cover, surface area-to-

volume ratio, vertical and horizontal continuity, dead and live fuel load, and size classes of fuel elements (Riaño *et al.* 2003). For that reason, the vegetation is operationally classified into different fuel types following a scheme of fuel properties that groups the vegetation classes with similar combustion behavior (Pyne *et al.* 1996). The accuracy of the fuel map used as input in fire spread simulations highly affects the results obtained. Moreover, the consistency and accuracy of the input data layers are very important for realistic predictions of fire growth (Finney 1998; Keane *et al.* 1998).

Traditionally, fuel types have been mapped by means of aerial photography and extensive fieldwork, which is highly expensive and time consuming (Riaño *et al.* 2003; Arroyo *et al.* 2008; Marino *et al.* 2016). Remote sensing data provide an alternative as source of fuel information, since they can provide spatial information on land cover. Fuel type mapping from satellite imagery has been attempted by several authors (Lasaponara and Lanorte 2007a; b; Mallinis *et al.* 2008; Otukei and Blaschke 2010). However, the main limitation of satellite images is their inability to estimate heights and to estimate vertical distribution of forest stands (Arroyo *et al.* 2008). These factors are critical not only for fuel type discrimination, but also for assessing some fuel characteristics needed for fire spread modeling such as fuel load, canopy cover, tree height, crown base height, and crown bulk density (Riaño *et al.* 2003). Light Detection And Ranging (LIDAR) allows overcoming these limitations (Arroyo *et al.* 2008). The ability of penetrating the canopy layer is leading authors to include LiDAR data as an essential source of information for fuel types mapping and fuel characterization (Riaño *et al.* 2003; Mutlu *et al.* 2008; Erdody and Moskal 2010; García *et al.* 2011; González-Ferreiro *et al.* 2014; Hermosilla *et al.* 2014; Marino *et al.* 2016; Ruiz *et al.* 2018).

In this context, the aims of the following three chapters are 1) to create accurate maps of main fuel types for Mediterranean areas combining multispectral and LiDAR data; 2) to improve the characterization of canopy fuel models using LiDAR data; and 3) to compare the variation in simulated wildfire spread and behavior determined by the use of fine-scale v. lower resolution maps, and to assess the effect of using our custom canopy layers in the simulations. The final objective is to develop a methodology for creating those maps and canopy layers which could be reproduced for other areas or for new available data.

REFERENCES

Ager AA, Buonopane M, Reger A, Finney MA (2013) Wildfire Exposure Analysis on the National Forests in the Pacific Northwest, USA. *Risk Analysis* **33**, 1000–1020. doi:10.1111/j.1539-6924.2012.01911.x.

Ager AA, Day MA, McHugh CW, Short K, Gilbertson-Day J, Finney MA, Calkin DE (2014) Wildfire exposure and fuel management on western US national forests. *Journal of Environmental Management* **145**, 54–70. doi:10.1016/j.jenvman.2014.05.035.

Alcasena FJ, Salis M, Ager AA, Arca B, Molina D, Spano D (2015) Assessing Landscape Scale Wildfire Exposure for Highly Valued Resources in a Mediterranean Area. *Environmental Management* **55**, 1200–1216. doi:10.1007/s00267-015-0448-6.

Arca B, Pellizzaro G, Duce P, Salis M, Bacciu V, Spano D, Ager AA, Scoccimarro E (2012) Potential changes in fire probability and severity under climate change scenarios in mediterranean areas. In: Spano D, Bacciu V, Salis M, Sirca C (eds) 'International Conference on Fire Behaviour and Risk', Alghero (Italy). 92–98.

Arroyo LA, Pascual C, Manzanera A (2008) Fire models and methods to map fuel types: The role of remote sensing. *Forest Ecology and Management* **256**, 1239–1252. doi:10.1016/j.foreco.2008.06.048.

Calkin DE, Thompson MP, Finney MA, Hyde KD (2011) A real-time risk assessment tool supporting wildland fire decision-making. *Journal of Forestry* **109**, 274–280.

Cole DN, Landres PB (1996) Threats to wilderness ecosystems: impacts and research needs. *Ecological Applications* **6**, 168–184. doi:10.2307/2269562.

Duff TJ, Chong DM, Tolhurst KG (2013) Quantifying spatio-temporal differences between fire shapes: Estimating fire travel paths for the improvement of dynamic spread models. *Environmental Modelling & Software* **46**, 33–43. doi:10.1016/j.envsoft.2013.02.005.

Erdody TL, Moskal LM (2010) Fusion of LiDAR and imagery for estimating forest canopy fuels. *Remote Sensing of Environment* **114**, 725–737. doi:10.1016/j.rse.2009.11.002.

Finney MA (1998) FARSITE : Fire Area Simulator — Model Development and Evaluation. USDA Forest Service, Rocky Mountain Research Station, RMRS-RP-4. (Ogden, UT) http://www.firemodels.org/content/view/52/72/.

Flannigan MD, Krawchuk M a, de Groot WJ, Wotton BM, Gowman LM (2009) Implications of changing climate for global wildland fire. *International Journal of Wildland Fire* **18**, 483–507. doi:10.1071/WF08187.

Flannigan MD, Stocks BJ, Wotton BM (2000) Climate change and forest fires. *Science of The Total Environment* **262**, 221–229. doi:10.1016/S0048-9697(00)00524-6.

García M, Riaño D, Chuvieco E, Salas J, Danson FM (2011) Multispectral and LiDAR data fusion for fuel type mapping using Support Vector Machine and decision rules. *Remote Sensing of Environment* **115**, 1369–1379. doi:10.1016/j.rse.2011.01.017.

González-Ferreiro E, Diéguez-Aranda U, Crecente-Campo F, Barreiro-Fernández L, Miranda D, Castedo-dorado F (2014) Modelling canopy fuel variables for Pinus radiata D . Don in NW Spain with low-density LiDAR data. **23**, 350–362. doi:10.1071/WF13054.

Hermosilla T, Ruiz LA, Kazakova A, Coops NC, Moskal LM (2014) Estimation of forest structure and canopy fuel parameters from small-footprint full-waveform LiDAR data. *International Journal of Wildland Fire* **23**, 224–233. doi:10.1071/WF13086.

Kalabokidis K, Palaiologou P, Gerasopoulos E, Giannakopoulos C, Kostopoulou E, Zerefos C (2015) Effect of Climate Change Projections on Forest Fire Behavior and Values-at-Risk in Southwestern Greece. *Forests* **6**, 2214–2240. doi:10.3390/f6062214.

Keane RE, Garner JL, Schmidt KM, Menakis JP, Finney MA (1998) Development of input data layers for the FARSITE fire growth model for the Selway-Bitterroot Wilderness Complex, USA. USDA Forest Service, Rocky Mountain Research Station, RMRS-GTR-3. (Fort Collins, CO)

Kovats S, Valentini R, Bouwer LM, Georgopoulou E, Jacob D, Martin E, Rounsevell M, Soussana J-F (2014) Chapter 23. Europe. In: Barros *et al.* (eds) Climate Change 2014: Impacts, Adaptation and Vulnerability Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, (Cambridge) doi:10.1017/CBO9781107415324.004.

Kurnik B, Van Der Linden P, Mysiak J, Swart R, Füssel H-M, Christiansen T, Cavicchia L, Gualdi S, Mercogliano P, Rianna G, Kramer K, Michetti M, Salis M, Schelhaas M-J, Leitner M, Vanneuville W, Macadam I (2017) Chapter 3. Weather- and climate-related natural hazards in Europe. In: Climate change adaptation and disaster risk reduction in Europe - Enhancing coherence of the knowledge base, policies and practices. Publications Office of the European Union, (Luxembourg) doi:10.2800/938195.

Lasaponara R, Lanorte A (2007a) Remotely sensed characterization of forest fuel types by using satellite ASTER data. *International Journal of Applied Earth Observation and Geoinformation* **9**, 225–234. doi:10.1016/j.jag.2006.08.001.

Lasaponara R, Lanorte A (2007b) On the capability of satellite VHR QuickBird data for fuel type characterization in fragmented landscape. *Ecological Modelling* **204**, 79–84. doi:10.1016/j.ecolmodel.2006.12.022.

Liu Y, Stanturf J, Goodrick S (2010) Trends in global wildfire potential in a changing climate. *Forest Ecology and Management* **259**, 685–697. doi:10.1016/j.foreco.2009.09.002.

Lozano OM, Salis M, Ager AA, Arca B, Alcasena FJ, Monteiro AT, Finney MA, Giudice L Del, Scoccimarro E, Spano D (2017) Assessing Climate Change Impacts on Wildfire Exposure in Mediterranean Areas. *Risk Analysis* **37**, 1799–2022. doi:10.1111/risa.12739.

Mallinis G, Mitsopoulos ID, Dimitrakopoulos AP, Gitas IZ, Karteris M (2008) Local-Scale Fuel-Type Mapping and Fire Behavior Prediction by Employing High-Resolution Satellite Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **1**, 230–239. doi:10.1109/JSTARS.2008.2011298.

Marino E, Ranz P, Luis J, Ángel M, Esteban J, Madrigal J (2016) Remote Sensing of Environment Generation of high-resolution fuel model maps from discrete airborne laser scanner and Landsat-8 OLI : A low-cost and highly updated methodology for large areas. *Remote Sensing of Environment* **187**, 267–280. doi:10.1016/j.rse.2016.10.020.

Miller C, Ager AA (2013) A review of recent advances in risk analysis for wildfire management. *International Journal of Wildland Fire* **22**, 1–14. doi:10.1071/WF11114.

Mutlu M, Popescu SC, Stripling C, Spencer T (2008) Mapping surface fuel models using lidar and multispectral data fusion for fire behavior. **112**, 274–285. doi:10.1016/j.rse.2007.05.005.

Otukei JR, Blaschke T (2010) Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation* **12**, 27–31. doi:10.1016/j.jag.2009.11.002.

Pausas JG, Keeley JE (2009) A Burning Story: The Role of Fire in the History of Life. *BioScience* **59**, 593–601. doi:10.1525/bio.2009.59.7.10.

Piñol J, Terradas J, Lloret F (1998) Climate warming, wildfire hazard, and wildfire occurrence in coastal eastern Spain. *Climatic Change* **38**, 345–357. doi:10.1023/A:1005316632105.

Pyne SJ, Andrews PL, Laven RD (1996) 'Introduction to wildland fire.' (John Wiley and Sons: New York)

Riaño D, Meier E, Allgöwer B, Chuvieco E, Ustin SL (2003) Modeling airborne laser scanning data for the spatial generation of critical forest parameters in fire behavior modeling. *Remote Sensing of Environment* **86**, 177–186. doi:10.1016/S0034-4257(03)00098-1.

Ruiz LÁ, Recio JA, Crespo-Peremarch P, Sapena M (2018) An object-based approach for mapping forest structural types based on low density LiDAR and multispectral imagery. *Geocarto International* **33**, 443–457. doi:10.1080/10106049.2016.1265595.

Salis M, Ager AA, Finney MA, Arca B, Spano D (2014) Analyzing spatiotemporal changes in wildfire regime and exposure across a Mediterranean fire-prone area. *Natural Hazards* **71**, 1389–1418. doi:10.1007/s11069-013-0951-0.

San-Miguel-Ayanz J, Durrant T, Boca R, Libertà G, Branco A, De Rigo D, Ferrari D, Maianti P, Artés Vivancos T, Schulte E, Loffler P (2017) Forest Fires in Europe, Middle East and North Africa 2016. Publications Office of the European Union, doi:10.2760/17690.

Syphard AD, Radeloff VC, Hawbaker TJ, Stewart SI (2009) Conservation Threats Due to Human-Caused Increases in Fire Frequency in Mediterranean-Climate Ecosystems. *Conservation Biology* **23**, 758–769. doi:10.1111/j.1523-1739.2009.01223.x.

Thompson MP, Calkin DE, Finney MA, Ager AA, Gilbertson-Day JW (2011) Integrated national-scale assessment of wildfire risk to human and ecological values. *Stochastic Environmental Research and Risk Assessment* **25**, 761–780.

Turco M, Llasat M-C, von Hardenberg J, Provenzale A (2014) Climate change impacts on wildfires in a Mediterranean environment. *Climatic Change* **125**, 369–380. doi:10.1007/s10584-014-1183-3.

Urrutia-Jalabert R, González ME, González-Reyes Á, Lara A, Garreaud R (2018) Climate variability and forest fires in central and south-central Chile. *Ecosphere* **9**,. doi:10.1002/ecs2.2171.

Westerling AL, Hidalgo HG, Cayan DR, Swetnam TW (2006) Warming and Earlier Spring Increase Western U.S. Forest Wildfire Activity. *Science* **313**, 940–943. doi:10.1126/science.1128834.

Chapter 1: Mapping fuel types combining Multispectral and LiDAR data of Sardinia

1. INTRODUCTION

Accurate fuel maps are not only a critical input for fire spread models, but also to plan fire prevention, management and suppression activities. The availability of accurate and spatially explicit information on fuel properties is critical in order to improve fire management decision-support systems since fuels affect fire ignition and propagation (Ottmar and Alvarado 2004; Chuvieco *et al.* 2009).

The high spatial and temporal variability of fuels makes field survey methods very expensive and time consuming for obtaining realistic characterization, and often limited for fuel mapping. Hence methods based on aerial photography and remotely sensed data have risen in the last years (Arroyo *et al.* 2008; Bajocco *et al.* 2015). Most studies based on satellite imagery have been performed at a coarse resolution or in very small areas with a very high resolution (Lasaponara and Lanorte 2007a; b, Mallinis *et al.* 2008, 2014; Otukei and Blaschke 2010).

As far as LiDAR (Light Detection and Ranging) is concerned, though these data have proved to be suitable to estimate some fuel properties, few studies have evaluated their usefulness to create fuel maps because of the difficulties in identifying land cover classes only from this data source (Yan *et al.* 2015).

The most important limitation of optical images is their inability to assess vegetation height, which is a critical variable to discriminate fuel types. The integration of LiDAR data allows overcoming this limitation (Arroyo *et al.* 2008).

In the last years some works have combined LiDAR and satellite imagery to map fuels, but the identification of individual species stills remain a complex task. Varga and Asner (2008) developed a new fire fuel index through the fusion of hyperspectral and LiDAR data to model the three-dimensional volume of grass fuels. Koetz *et al.* (2008) used LiDAR and hyperspectral data to map different land cover types including roads, buildings and vegetation but with only three vegetation types (ground fuel, shrubs and

tree canopy). Also García *et al.* (2011) combined LiDAR and multispectral data in this case, for fuel mapping, but since they followed the Prometheus fuel type system to classify the different fuel types, only structural characteristics were considered such as the average height of vegetation or the average height difference between shrubs and trees. Recently, Reese *et al.* (2014) classified alpine vegetation combining optical satellite data and LiDAR derived data, and Marino *et al.* (2016) obtained fuel model maps from discrete airborne laser scanner and Landsat-8 OLI (30-m resolution).

Regarding Sardinia, there are no fuel type maps for the whole island at fine scale. Until now, the maps used to derive fuel type information are coarse resolution land use maps (e.g. Corine Land Cover, EEA 2011 or Land Use Map of Sardinia, Autonomous Region of Sardinia 2008).

The overall aim of the work is to improve the available information about the fuel types spatial distribution by creating accurate fuel maps at fine scale from remotely sensed data, namely combining LiDAR with multispectral data.

2. METHODS

2.1. Study area

The study is performed in the island of Sardinia (Italy), which is located between 38° 44' and 41° 05' N latitude and 8° 7' and 9° 49' E longitude. The availability of LiDAR data for the island is limited to the coastal areas and, within these boundaries, two study areas covering approximately the same extent (4,000 ha approximately) were selected to carry out this work (Fig. 1.1). These areas are located on the eastern coast (Siniscola) and on the southeastern coast (Muravera),

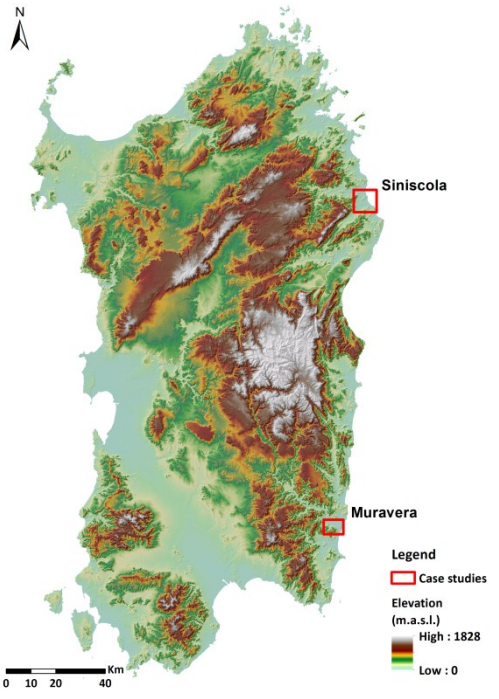

Fig. 1.1. Location of the case studies in Sardinia

and were selected to include in the study different types of vegetation structures that can be often found in Sardinia (i.e.: broadleaf forests and Mediterranean maquis).

2.2. Classification method

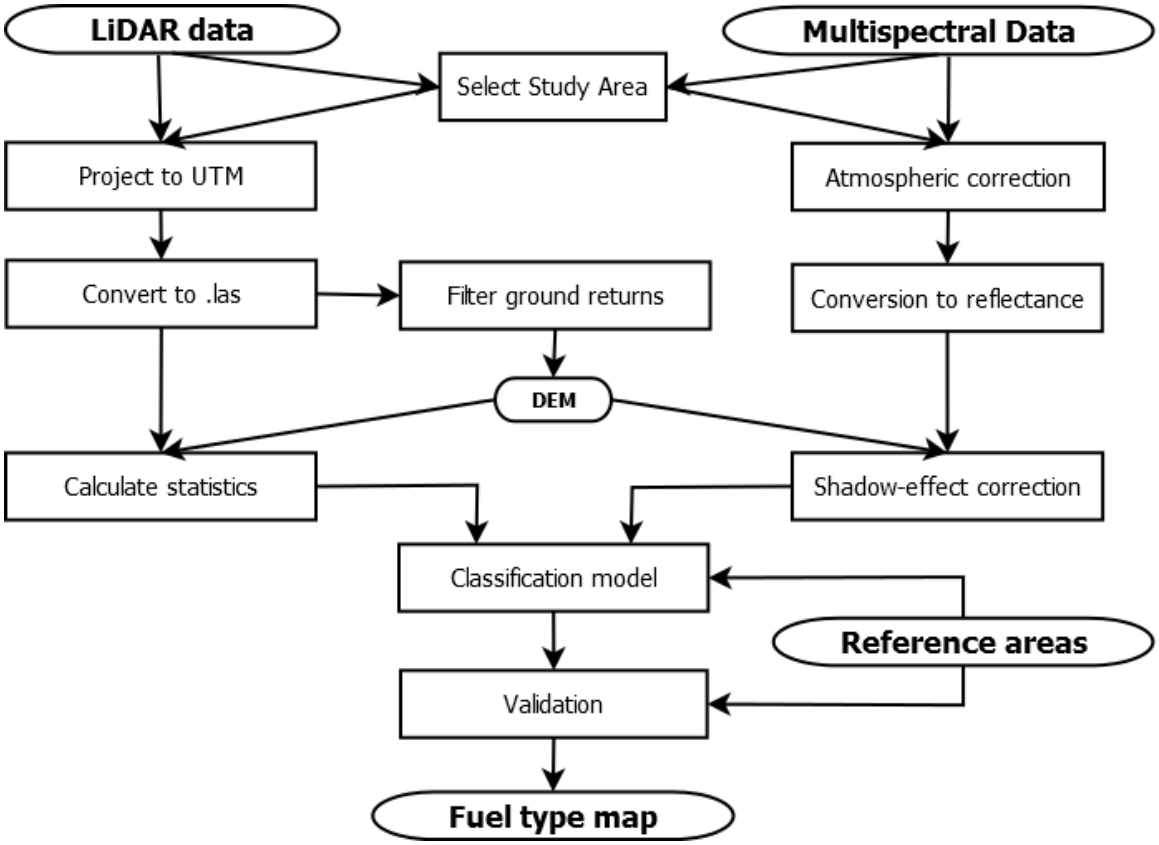

Fig. 1.2. Work flow diagram showing the main steps for producing a fuel map. DEM = Digital Elevation Model

The fuel type classification was carried out following the scheme showed in figure 1.2. After processing remote sensing data (see sections 2.4 and 2.5), we obtained a series of layers with a number of variables (bands of multispectral data, spectral indices and height distribution statistics and canopy cover from LiDAR). Then a set of reference areas was defined to determine their actual land cover type. For this analysis, we used as source of reference data, an orthophoto co-registered with simultaneously acquired LiDAR data (0.1m-resolution).

The calibration and validation of the classification model has been done by assigning the corresponding LiDAR-derived metrics and Spot-5 reflectance values to the reference areas (where the fuel type is already known). We used 70% of reference areas to fit the model and 30% to validate it.

We evaluated four different algorithms which are commonly used to classify remote sensed data: **Maximum Likelihood** (ML), **Neural Networks** (NN), **Support Vector Machine** (SVM) **and Random Forests** (RF).

The **maximum likelihood (ML)** algorithm is one of the most widespread parametric methods that have been traditionally used for classification of remote sensing data (Martin *et al.* 1998; Walter 2004; Shalaby and Tateishi 2007). Since a normal distribution within each class is assumed, a probability function of a pixel belonging to a certain class can be calculated according to the training data's values. In this way, the ML classifier assigns the pixel to the class which maximizes the probability function (Chuvieco 2010).

However, since parametric methods implies assumptions such as the normal distribution of data, alternative non-parametric methods like **artificial neural networks (NN)** have been employed to classify remote sensing images (Frizzelle and Moody 2001; Qiu and Jensen 2004; Yuan *et al.* 2009). Neural networks are able to learn from training data creating a complex classification scheme which is used to classify the rest of observations (Chuvieco 2010). They try to mimic the neural storage and analytical operations of the brain where "neurons" are interconnected through weighted relationships (Frizzelle and Moody 2001).

Another non-parametric method used in remote sensing classification is **suppport vector machine (SVM)** which is also a machine learning algorithm (Pal and Mather 2005; García *et al.* 2011; Mountrakis *et al.* 2011). Using the training data, the algorithm attempts to find a hyperplane that separates the dataset into a discrete predefined classes minimizing misclassifications (Mountrakis *et al.* 2011).

Recently **random forests (RF)** have been also introduced for remote sensed data analysis, especially for LiDAR data (Falkowski *et al.* 2009; Yu *et al.* 2011; Valbuena *et al.* 2016), but it has been also used for classifying optical data (Pal 2005; Reese *et al.* 2014). It is a non-parametric method which consists of a combination of decision trees classifiers where different samples are randomly chosen from the training data to construct each individual tree (Breiman 2001). This algorithm searches only a random subset of the variables in order to minimize the correlation between the classifiers in the ensemble.

These different classification were carried out using the R package 'Rasclass' (R Core Team 2016; Wiesmann and Quinn 2016). Finally we built four fuel maps for each study area following the different classification model obtained from the outputs of each algorithm.

2.3. Fuel types

In this work, the fuel types that we discriminated with the proposed methodology are those shown in table 1.1. We selected these fuel types because they have been already characterized and/or tested in other studies in Sardinia (Arca *et al.* 2007, 2009, Salis *et al.* 2013, 2016, 2018; Ager *et al.* 2014; Alcasena *et al.* 2015). To compare our outputs with the information available until now, we also reclassified the Land Use Map (LUM) of Sardinia (Autonomous Region of Sardinia 2008) for the same fuel types. The original classes of the LUM of Sardinia are based on the Corine Land Cover classification (EEA 2011). The LUM of Sardinia was elaborated using different sources such as: ortophoto AGEA (AGenzia per le Erogazioni in Agricoltura) 2003, ortophoto 2004, images Ikonos 2005-06, images Landsat 2003 and images Aster 2004. As reference data, 4000 sample site were set throughout the island.

Table 1.1. Fuel types to be mapped and corresponding classes of the Land Use Map (LUM) of sardinia.

Code	Fuel type	LUM Classes
1	Buildings (non-fuel)	143, 1111, 1112, 1121, 1122, 1211, 1212, 1224, 1421
2	Roads (non-fuel)	123, 1221, 1322, 1421
3	Water	3315, 5111, 5112, 5122, 5211, 5212, 5231
4	Bare ground	131, 133, 1321, 3311
5	Sparse vegetation	333, 3313
6	Mixed agricultural	242, 243, 2112, 2121, 2123
7	Vineyard and orchard	221, 222, 223, 2411, 2413
8	Herbaceous vegetation	321, 2111
9	Garrigue	244, 411, 421, 3232, 3241
10	Mediterranean maquis	3221, 3222, 3231
11	Conifer forests	313, 3121, 3242
12	Broadleaf forests	3111, 31121, 31122
13	Mixed forests	141, 313

2.4. LiDAR data processing

LiDAR data of the study area in ASCII (.xyz) format were provided by the Autonomous Region of Sardinia (Servizio osservatorio del paesaggio e del territorio, sistemi informativi territoriali). These LiDAR data were recorded in different periods from 21st October 2008 to 10th May 2009 using as laser equipment an Altm GEMINI sensor and obtaining a minimum point density of 0.8 m^{-2}.

First step to process the LiDAR data (Fig. 1.2) was to project the point data to UTM (zone 32N) with datum WGS-84 and convert them to a LAS format using LAStools (Isenburg 2015). Then we filtered the ground points creating a digital elevation model (DEM) which was used to calibrate the height of the points and calculate the metrics with FUSION (McGaughey 2014). The descriptive statistics computed by the Gridmetrics command of FUSION are shown in tables 1.2, 1.3 and 1.4. To discriminate different land cover types, we left out variables with constant values, because they did not add extra information. We also omitted absolute variables, since point density is not spatially uniform. For the above reason, we only considered relative variables.

Table 1.2. Variables related to the metrics of heights obtained from FUSION gridmetrics (McGaughey 2014).

Variable name	Description
Elev minimum	minimum
Elev maximum	maximum
Elev mean	mean
Elev mode	mode
Elev stddev	standard deviation
Elev variance	variance
Elev CV	coefficient of variation
Elev IQ	interquartile range
Elev skewness	skewness
Elev kurtosis	kurtosis
Elev AAD	absolute average deviation
Elev L1, L2, L3 and L4	L-moments
Elev L CV	L-moment of coefficient of variation
Elev L skewness	L-moment of skewness
Elev L kurtosis	L-moment of kurtosis
p01, p05, p10, p20… p95, p99	percentiles
Elev MAD median	median of the absolute deviations from the overall median
Elev MAD mode	median of the absolute deviations from the overall mode
Elev quadratic mean	quadratic mean
Elev cubic mean	cubic mean
Canopy relief ratio	(mean height- min height) / (max height– min height)
Variable name	Description
Elev minimum	minimum
Elev maximum	maximum
Elev mean	mean
Elev mode	mode
Elev stddev	standard deviation
Elev variance	variance
Elev CV	coefficient of variation
Elev IQ	interquartile range
Elev skewness	skewness
Elev kurtosis	kurtosis
Elev AAD	absolute average deviation
Elev L1, L2, L3 and L4	L-moments
Elev L CV	L-moment of coefficient of variation
Elev L skewness	L-moment of skewness
Elev L kurtosis	L-moment of kurtosis
p01, p05, p10, p20… p95, p99	percentiles
Elev MAD median	median of the absolute deviations from the overall median
Elev MAD mode	median of the absolute deviations from the overall mode
Elev quadratic mean	quadratic mean
Elev cubic mean	cubic mean
Canopy relief ratio	(mean height- min height) / (max height– min height)

Table 1.3. Variables related to the metrics of canopy cover obtained from FUSION gridmetrics (McGaughey 2014).

Variable name	Description
Canopy cover	Percentage first returns above 2.00 m
allcover	Percentage all returns above 2.00 m
afcover	(All returns above 2.00 m) / (Total first returns) * 100
abovemean	Percentage first returns above mean
abovemode	Percentage first returns above mode
allabovemean	Percentage all returns above mean
allabovemode	Percentage all returns above mode
afabovemean	(All returns above mean) / (Total first returns) * 100
afabovemode	(All returns above mode) / (Total first returns) * 100

Table 1.4. Variables related to the metrics of strata obtained from FUSION gridmetrics (McGaughey 2014). These variables were computed for each strata (below 0.5 m; from 0.5 to 1 m; from 1 to 2 m; from 2 to 3 m; from 3 to 5 m; from 5 to 10 m and above 10 m).

Variable name	Description
Elev strata return proportion	(Total return count for the strata)/(all returns)
Elev strata min	Minimum elevation for the strata
Elev strata max	Maximum elevation for the strata
Elev strata mean	Average elevation for the strata
Elev strata mode	Mode of elevations for the strata
Elev strata median	Median elevation for the strata
Elev strata stddev	Standard deviation of elevations within the the strata
Elev strata CV	Coefficient of variation for elevations within the the strata
Elev strata skewness	Skewness of elevations within the the strata
Elev strata kurtosis	Kurtosis of elevations within the the strata

Each variable was converted into a raster ASCII file to be used as input for the classification.

2.5. Multispectral data processing

The selection of a satellite image presenting a good compromise between spectral and spatial resolution is a key point for data processing. Two Spot-5 satellite images (Table 1.5) acquired over both study areas in 2009 were used. The relatively high spatial

resolution (i.e. 10m) satisfies the research objectives, generating a high-resolution scene model. Furthermore, it facilitates the fusion with the LiDAR data, allowing to have a satisfactory number of points (more than 80 points) included within each pixel. With a very high resolution (such as 2 m) we would not be able to compute LiDAR statistics for each pixel. Finally, the information content of this data is likely much higher compared to the data sources available until now in Sardinia.

Table 1.5. Technical characteristics of Spot-5 satellite. NIR is Near Infrared and SWIR is short wave infrared.

Bands	**Spectral range (µm)**	**Spatial resolution (m)**
B1 (Green)	0.50 - 0.59	10
B2 (Red)	0.61 - 0.68	10
B3 (NIR)	0.78 - 0.89	10
B4 (SWIR)	1.58 - 1.75	20
Pan	0.48 - 0.71	2

Spot-5 images were already orthorectified and therefore the first step was to remove the atmospheric effect and to convert the values to reflectance (Fig. 1.2). To carry out this step, we used ATCOR-2 (Richter and Schläpfer 2012) which is a model that corrects the image according to a set of standard atmospheric profiles (Chuvieco 2010).

A topographic correction has been then performed to remove the effect of shadowing due to the slope and aspect (Chuvieco 2010). Since our study areas cover also some forests, we decided to use the correction developed by Soenen *et al.* (2005), which consider the effect of the vertical growth of trees. The formulation for this correction is:

$$\rho_n = \rho \times \left(\frac{\cos\alpha \cdot \cos\theta_s + C}{\cos\gamma + C} \right)$$

where α is the terrain slope, γ is the incidence angle, θ_s is the solar zenith angle and C is an empirical constant calculated for every band separately from the parameters of the regression of reflectance and the cosine of the incidence angle. Namely, C is the intercept to slope ratio:

$$\rho = b + m \times \cos\gamma$$

$$C = \frac{b}{m}$$

In addition to the information from the different bands, some spectral indices have been also calculated since they have been proved to be good indicators of different vegetation species. These indices are: Normalized Difference Vegetation Index (NDVI), Soil Adjusted Vegetation Index (SAVI), and Normalized Difference Infrared Index (NDII). We computed these indices using the raster calculator tool (ArcGIS) following the equation:

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

where 'RED' and 'NIR' define the spectral reflectance measurements acquired in the visible (red) and near-infrared regions, respectively.

$$SAVI = \frac{NIR - RED}{NIR + RED + L} \times (1 + L)$$

where 'NIR' is the reflectance value of the near infrared band, 'RED' is reflectance of the red band, and 'L' is the correction factor for soil brightness. The value of L varies depending on the amount or cover of green vegetation: with very high vegetation, L=0, while in areas with no green vegetation, L=1. When L=0, then SAVI = NDVI. Generally, the default value used for the most of scientific papers is L=0.5.

$$NDII = \frac{NIR - SWIR}{NIR + SWIR}$$

where 'SWIR' and 'NIR' define the spectral reflectance measurements acquired in the short-wave infrared and near-infrared regions, respectively.

2.6. Variable selection

The maximum likelihood algorithm is a parametric method, thus the best explaining not correlated variables should be selected. Therefore, first step was to create four sets of explanatory variables (databases) with the data associated to the reference areas: (1) height metrics (Table 1.2); (2) cover metrics (Table 1.3); (3) strata metrics (Table 1.4); and (4) multispectral image bands and spectral indices (Table 1.5). Then, for each set of explanatory variables we calculated the correlation matrix and we deleted one of the pair of variables with high correlation coefficient. Finally, from the remaining not highly correlated variables, we selected those common to both study areas. We used the same set of variables for the classification with the SVM algorithm, because, even if it has no theoretical limitations, it works worse when variables increase (Mountrakis *et al.* 2011).

The other algorithms that we used (NN, and RF) are non-parametric methods. Theoretically, these algorithms should not be affected by distributional assumptions and correlation problems, but some studies found different results (Strobl *et al.* 2007, 2008). In any case, they have no limitations regarding the number of variables. Therefore, for these algorithms, we tried the classification with both approaches, that is: 1) the subset of variables and 2) all variables.

2.7. Statistical methods

The accuracy of the classification experiments was estimated using the remaining 30% of the reference areas which were not used for the training phase. For each classification output, we calculated the confusion matrix, the overall accuracy and the Cohen's Kappa coefficient (Congalton 1991; Senseman *et al.* 1995).

The overall accuracy is the simplest measure of agreement and is calculated as follows:

$$OA = \frac{\sum_{i=1}^{r} x_{ii}}{\sum_{i=1}^{r} x_{i+}}$$

where x_{ii} are the agreement cases (values in diagonal of the confusion matrix) and x_{i+} are the total number of reference values. When the agreement is very high OA coefficient values are close to one.

The Kappa coefficient (K) is a bivariate agreement coefficient that becomes zero for chance agreement, one for perfect agreement, and negative for less than chance agreement (Foody 2006; Chuvieco 2010). K values are calculated as follows:

$$K = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+}x_{+1})}{N^2 - \sum_{i=1}^{r}(x_{i+}x_{+i})}$$

where r is the number of rows in the error matrix, x_{ii} is the number of observations in row i and column i, x_{i+} and x_{+i} are the marginal totals of row i and column i, respectively and N is the total number of observations.

Total accuracy coefficients does not reveal if error was evenly distributed among classes or if some classes were very accurately classified whereas other classes were completely misclassified. Therefore, we calculated the user's and producer's accuracy as a measure of accuracy for each fuel type. User's accuracy is related to the error of commission whereas producer's accuracy is related to the error of omission. An error of commission is said to occur when a class is mapped incorrectly where it does not exist. Then an error is 'committed', meaning that a class is over-mapping. On the other hand, the error of omission refers to reference areas that were left out (or omitted) from the correct class in the classified map.

Finally, an accuracy assessment procedure was also followed for evaluating the quality of the information of the LUM map. We built the confusion matrix and calculated the different accuracy measures. For this purpose, we randomly selected 30% of reference areas and we crossed these data with LUM data.

3. RESULTS

We defined as reference areas for the different classes a total of 2,105 pixels for the study area of Siniscola and 673 pixels for the study area of Muravera. We should set

more reference areas in the Siniscola study area due to the scattered anthropic areas which were intermingled with the vegetation. The frequency of reference areas was according to the pre-estimated cover extent of each class (Table 1.6).

Table 1.6. Fuel types and their corresponding number of pixels used as the reference data in each study area.

Code	Fuel type	Reference pixels	
		Siniscola	Muravera
1	Buildings (non-fuel)	276	41
2	Roads (non-fuel)	373	74
3	Water	78	86
4	Bare ground	159	42
5	Sparse vegetation	232	39
6	Mixed agricultural	138	49
7	Vineyard and orchard	59	54
8	Herbaceous vegetation	141	63
9	Garrigue	116	44
10	Mediterranean maquis	129	69
11	Conifer forests	154	53
12	Broadleaf forests	118	59
13	Mixed forests	132	
	Total	2105	673

For each study area we used the four different algorithms (ML, NN, SVM and RF) with a set of less correlated variables to perform the classifications. The selected variables were: the Spot-5 bands (Band 1-Green, Band 2-Red, Band 3-NIR and Band 4-SWIR); the calculated spectral indices NDVI and NDII; the LiDAR height metrics coefficient of variation, kurtosis, 5^{th} percentile, 95^{th} percentile and skewness; the canopy cover; and from the LiDAR strata metrics the minimum height of the first stratum, and the proportion of returns of the 6^{th} stratum.

Moreover, for the NN and RF algorithms we tried also the classifications with all variables.

3.1. Overall performance

Results of all classifications showed rather high values of accuracy with overall accuracy ranging from 77.23% to 91.60% and Kappa coefficient ranging from 0.75 to 91 (Table 1.7).

The algorithm obtaining the highest accuracy values was RF in both study areas, especially when the classification was performed with the subset of variables (overall accuracy about 91.6% and Kappa coefficient about 0.91). Using NN with all variables we obtained the lower accuracy values for both study areas (overall accuracy about 77.5% and Kappa coefficient about 0.75).

Overall accuracy values were always higher than Kappa coefficient values for all classifications.

Comparing the results of the two study areas, the accuracy results were similar with differences lower than 1% in most of the classifications. Using the RF algorithm with all variables the accuracy values of Siniscola were approximately 3% higher than those of Muravera.

In table 1.7 we have also included the values related to the validation carried out for the Land Use Map (LUM) of Sardinia. As can be observed these accuracy values are much lower than those obtained from our classifications.

Table 1.7. Accuracy results for the two study areas using the different classification algorithms and set of variables. ML = Maximum Likelihood; NN = Neural Networks; SVM = Support Vector Machine; RF = Random Forests; LUM = Land Use Map.

Study area	Method	Variables	Overall accuracy	Kappa coefficient
Muravera	ML	subset	83.17%	0.82
	NN	subset	80.69%	0.79
		all	77.23%	0.75
	SVM	subset	84.16%	0.83
	RF	subset	91.58%	0.91
		all	88.12%	0.87
	LUM		60.40%	0.57
Siniscola	ML	subset	84.15%	0.83
	NN	subset	80.03%	0.78
		all	77.65%	0.75
	SVM	subset	86.21%	0.85
	RF	subset	91.60%	0.91
		all	90.97%	0.90
	LUM		53.18%	0.49

3.2. Performance per classes

In the case of Siniscola study area, almost all algorithms coincided in the categories covering the minor extent in the study area, which were 'Water', 'Bare ground' and 'Mixed forest' (Table 1.8, Fig. 1.3, and Appendix A). The highest proportion of land, correspond to fuel types 'Garrigue', 'Herbaceous vegetation' and 'Mixed agricultural' even if there was not much agreement among the classification outputs on this point.

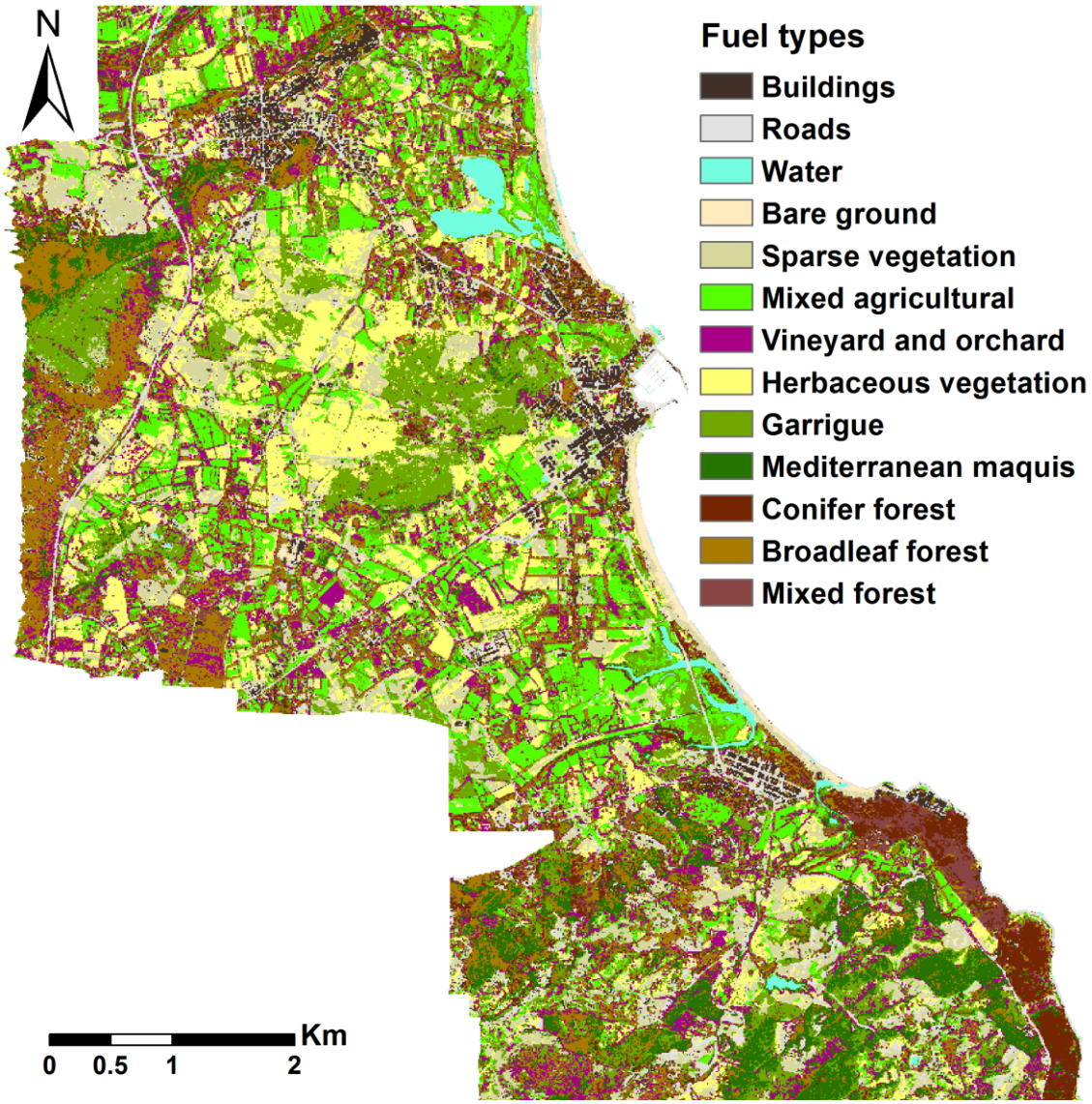

Fig. 1.3. Classification output map for the study area of Siniscola using the Random Forests (RF) algorithm with the subset of variables.

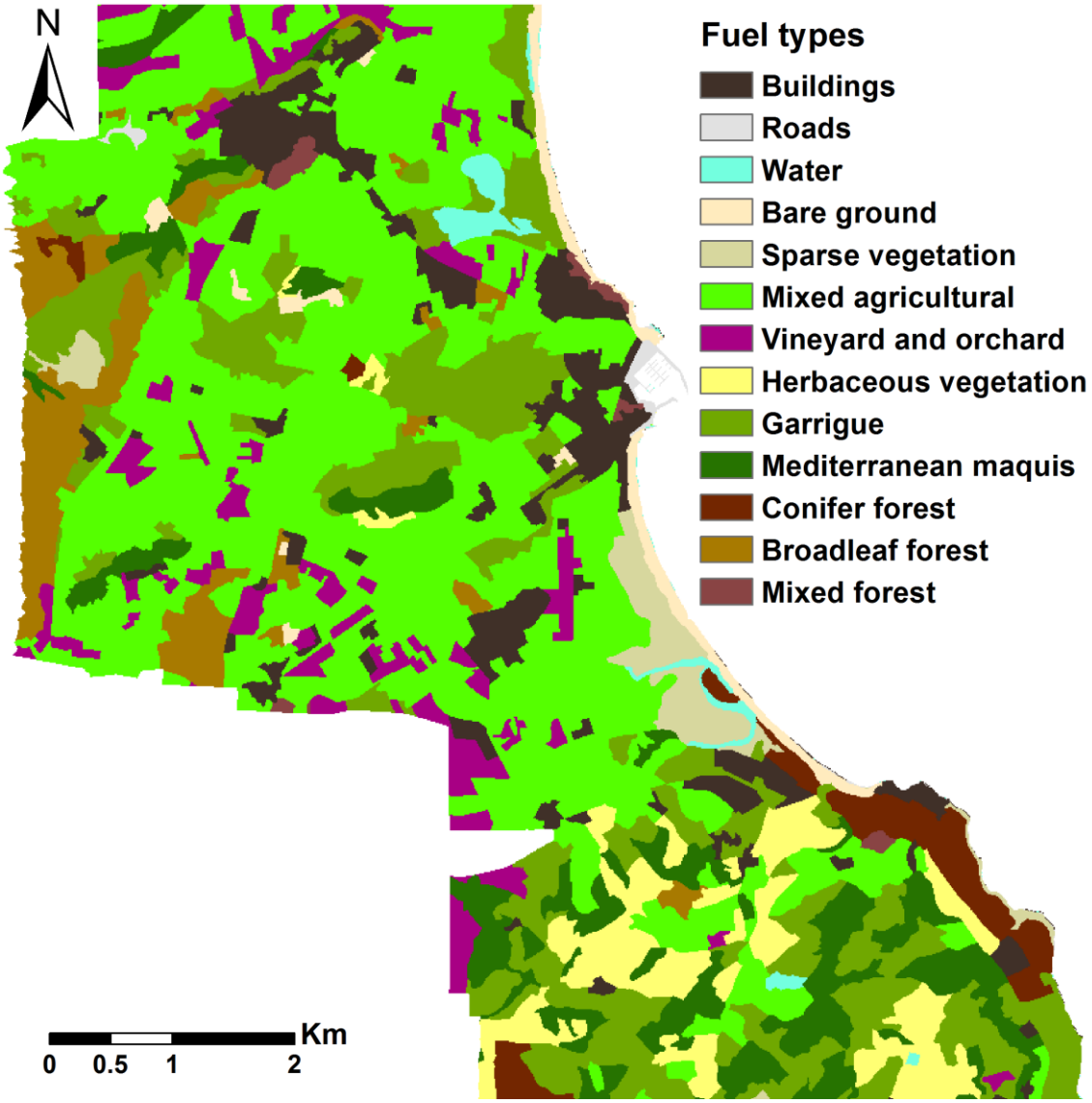

Fig. 1.4. Reclassification in the investigated classes of the Land Use Map (LUM) of Sardinia for the study area of Muravera.

Regarding the LUM of Sardinia (Table 1.8 and Fig. 1.4), the 'Mixed agricultural' is the fuel type covering the largest extent followed in a much lesser extent by 'Garrigue'. This map showed as the least frequent fuel types 'Roads' and 'Mixed forest'.

Table 1.8. In the study area of Siniscola, percentage of area covered by each fuel type according to each algorithm and set of variables. ML = Maximum Likelihood; NN = Neural Networks; SVM = Support Vector Machine; RF = Random Forests; LUM = Land Use Map.

Class	ML	NN		SVM	RF		LUM
	subset	all	subset	subset	all	subset	
Buildings	6.16	7.77	4.80	4.48	4.68	3.94	7.56
Roads	6.31	8.31	6.95	6.25	11.33	5.95	0.27
Water	0.86	2.72	2.32	1.06	1.31	1.19	1.01
Bare ground	1.52	5.99	1.03	0.71	0.72	0.94	1.64
Sparse vegetation	10.26	7.36	8.41	11.72	9.90	10.58	2.32
Mixed agricultural	8.52	19.98	8.54	9.36	13.67	11.45	41.42
Vineyard and orchard	12.49	4.27	1.83	8.50	7.59	11.49	5.42
Herbaceous vegetation	14.62	10.96	16.10	12.76	10.62	12.92	5.94
Garrigue	17.26	9.72	25.30	21.10	12.54	17.32	17.45
Mediterranean maquis	4.90	9.84	6.82	8.26	11.15	6.71	8.56
Conifer forest	2.79	3.25	5.14	4.78	3.58	3.63	2.74
Broadleaf forest	12.97	7.46	9.88	8.01	9.55	12.05	5.16
Mixed forest	1.34	2.37	2.87	3.01	3.36	1.82	0.51

Also in the study area of Muravera there was a high degree of classification correspondence among maps regarding the fuel types least frequent which were 'Bare ground', 'Water' and 'Herbaceous vegetation' (Tanle 1.9, Fig. 1.5, and Appendix B). Fuel types covering the largest extent were 'Mediterranean maquis', 'Broadleaf forest' and 'Garrigue'. The fuel type 'Mixed forest' is not included in this case because was not found in this study area.

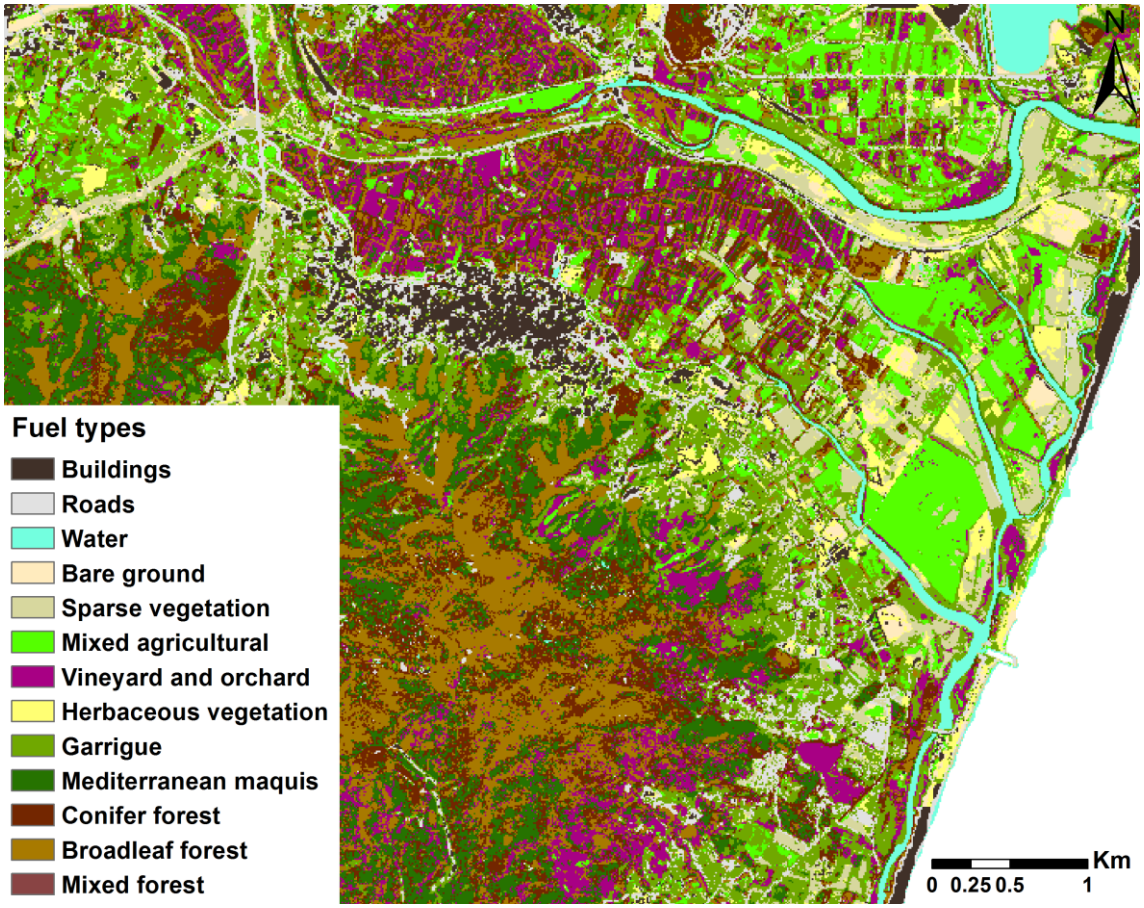

Fig. 1.5. Classification output map for the study area of Muravera using the Random Forests (RF) algorithm with the subset of variables.

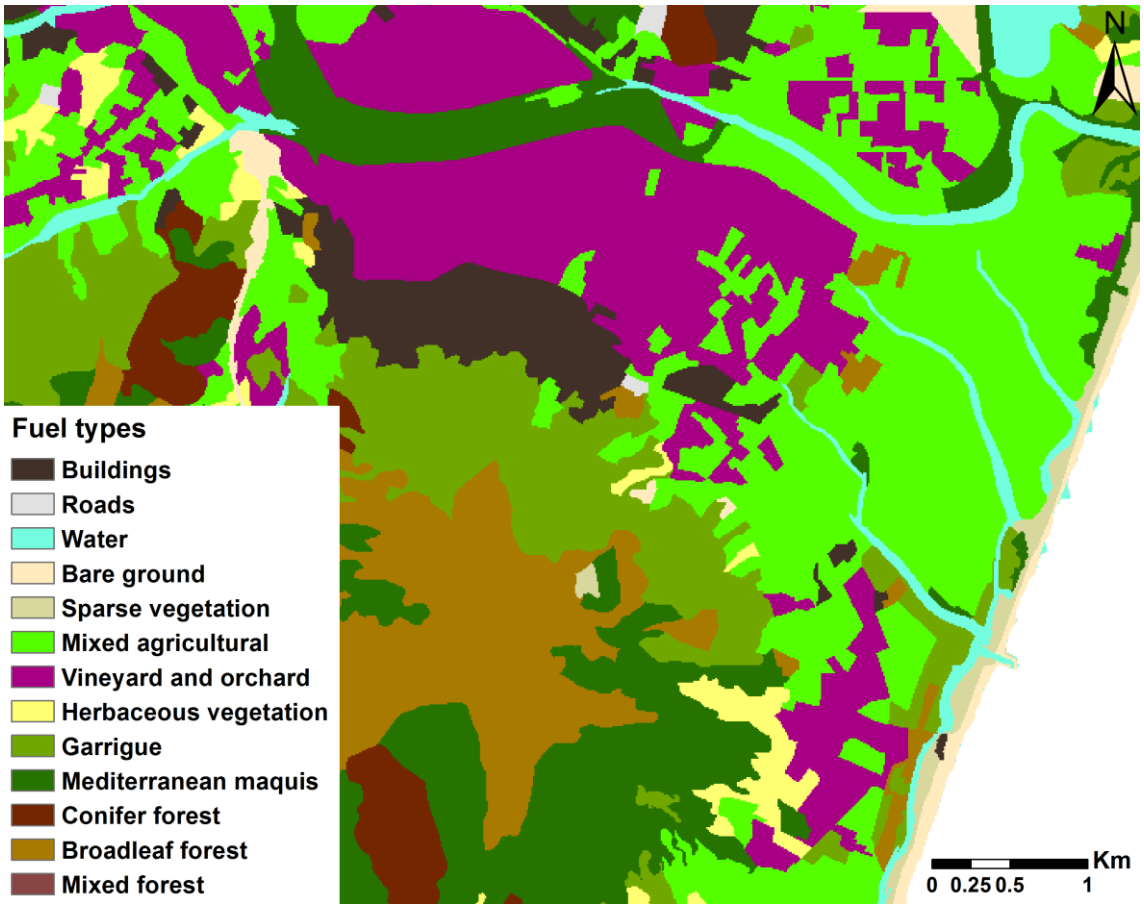

Fig. 1.6. Fuel types of the Muravera study area, according to the Land Use Map (LUM) of Sardinia

In this area, the LUM of Sardinia (Table 1.9 and Fig. 1.6) showed 'Mixed agricultural' as the most frequent fuel type followed by the 'Mediterranean maquis' whereas 'Roads' and 'Sparse vegetation' were present only in a 0.20 and 0.88%, respectively, present in this area.

Table 1.9. In the study area of Muravera, percentage of area covered by each fuel type according to each algorithm and set of variables. ML = Maximum Likelihood; NN = Neural Networks; SVM = Support Vector Machine; RF = Random Forests; LUM = Land Use Map.

Class	ML		NN	SVM		RF	LUM
	subset	all	subset	subset	all	subset	
Buildings	6.24	5.48	1.97	2.18	3.29	3.34	4.60
Roads	5.91	9.20	5.94	5.29	8.25	6.81	0.20
Water	1.53	2.88	2.48	2.32	2.66	2.35	3.09
Bare ground	1.22	3.42	0.91	0.81	3.59	0.88	1.72
Sparse vegetation	1.91	3.69	3.76	3.23	5.97	5.25	0.88
Mixed agricultural	5.71	4.35	9.19	7.18	3.63	6.48	22.89
Vineyard and orchard	10.25	14.31	18.09	15.36	9.35	9.41	16.53
Herbaceous vegetation	2.67	3.12	3.50	3.22	1.66	2.53	2.01
Garrigue	21.57	4.95	10.27	16.58	10.46	17.42	12.85
Mediterranean maquis	15.54	21.48	16.33	20.16	22.89	15.65	18.96
Conifer forest	5.91	5.95	15.57	10.32	12.39	13.14	2.71
Broadleaf forest	21.53	21.17	12.00	13.33	15.87	16.75	13.57

Regarding the user's and producer's accuracy, 'Water' was the class which higher values (Tables 1.10 and 1.11, and Appendices C and D). However, regarding the producer's accuracy the most accurately classified fuel types were very different in the two study area: in Siniscola they were 'Conifer forest', 'Garrigue' and 'Mixed agricultural' while in Muravera they were 'Herbaceous vegetation', 'Mixed agricultural' and 'Sparse vegetation'. Results regarding the fuel types with higher error of omission were very uneven. In Siniscola the lowest value of producer's accuracy was in the fuel

type ‘Vineyard and orchard’ with the NN algorithm using the subset of variables (9.09%), but with the RF algorithm using all variables or with the ML algorithm we obtained a 100% of producer’s accuracy. In Muravera again the NN algorithm with all variables showed the worst performance in this case for the ‘Conifer forest’ with a producer’s accuracy value of 0% while RF with all variables reached a 100% for the same fuel type.

In case of user’s accuracy, in both areas the fuel types ‘Mixed Agricultural’ and ‘Buildings’ showed the highest values, whereas ‘Vineyard and orchard’ in Siniscola (40%) and ‘Conifer forest’ in Muravera (0%) showed the lowest values when they were classified using the NN algorithm with the subset of variables and all variables respectively.

In both, producer’s and user’s accuracy, RF showed high values and very balanced for all fuel types, especially when performed with the subset of variables. Conversely, NN showed most uneven results for the different fuel types.

Also in this case we included data from the validation carried out with the Sardinian LUM. The results are very unbalanced showing for some fuel types very high accuracy values, such as ‘Broadleaf forest’ and ‘Water’ in both areas whereas other fuel types such as ‘Conifer forest’ showed very different values in each study area. There were also some fuel types for which producer’s and user’s accuracy was very unbalanced, especially in the Siniscola study area (‘Sparse vegetation’ and ‘Roads’).

Table 1.10. Producer's and user's accuracy (%) from the classification outputs obtained using the different algorithms in the study area of Siniscola. ML = Maximum Likelihood; NN = Neural Networks; SVM = Support Vector Machine; RF = Random Forests; LUM = Land Use Map.

Class	Accuracy	ML	NN		SVM	RF		LUM
		subset	subset	all	subset	subset	all	
Buildings	Producer's	89.77	85.71	79.12	90.11	89.77	88.64	94.68
	User's	91.86	96.30	90.00	94.25	96.34	93.98	69.53
Roads	Producer's	81.48	91.96	80.36	91.96	94.44	92.59	15.84
	User's	91.67	89.57	86.54	87.29	90.27	90.91	100.00
Water	Producer's	100.00	100.00	92.00	100.00	100.00	100.00	81.48
	User's	100.00	92.59	82.14	100.00	95.83	100.00	100.00
Bare ground	Producer's	83.64	75.00	84.09	70.45	87.27	74.55	60.78
	User's	64.79	71.74	56.06	75.61	84.21	83.67	51.67
Sparse vegetation	Producer's	54.79	72.60	53.42	73.97	82.19	84.93	1.28
	User's	70.18	72.60	68.42	75.00	80.00	70.45	100.00
Mixed agricultural	Producer's	91.67	84.85	96.97	90.91	94.44	94.44	97.73
	User's	100.00	100.00	65.31	100.00	100.00	100.00	29.86
Vineyard and orchard	Producer's	100.00	9.09	81.82	68.18	85.71	100.00	84.00
	User's	77.78	40.00	81.82	88.24	100.00	93.33	72.41
Herbaceous vegetation	Producer's	94.74	79.31	58.62	75.86	81.58	78.95	12.90
	User's	66.67	63.89	73.91	70.97	88.57	93.75	57.14
Garrigue	Producer's	97.30	86.84	84.21	100.00	97.30	100.00	65.71
	User's	87.80	68.75	80.00	90.48	92.31	94.87	30.26
Mediterranean maquis	Producer's	84.78	90.48	88.10	95.24	97.83	95.65	86.11
	User's	100.00	79.17	88.10	83.33	95.74	97.78	55.36
Conifer forests	Producer's	95.24	90.70	88.37	95.35	100.00	100.00	50.00
	User's	93.02	68.42	64.41	85.42	97.67	100.00	38.89
Broadleaf forests	Producer's	84.85	65.71	74.29	65.71	90.91	93.94	94.12
	User's	80.00	62.16	89.66	95.83	93.75	100.00	84.21
Mixed forests	Producer's	76.32	61.36	65.91	90.91	94.74	100.00	0.00
	User's	82.86	90.00	90.63	83.33	94.74	95.00	0.00

Table 1.11. Producer's and user's accuracy (%) from the classification outputs obtained using the different algorithms in the study area of Muravera. ML = Maximum Likelihood; NN = Neural Networks; SVM = Support Vector Machine; RF = Random Forests; LUM = Land Use Map.

Class	Accuracy	ML	NN		SVM	RF		LUM
		subset	subset	all	subset	subset	all	
Buildings	Producer's	85.71	69.23	69.23	76.92	85.71	92.86	100.00
	User's	92.31	90.00	52.94	100.00	100.00	100.00	100.00
Roads	Producer's	80.00	75.00	54.17	79.17	85.00	65.00	0.00
	User's	84.21	72.00	72.22	86.36	85.00	72.22	0.00
Water	Producer's	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	User's	100.00	96.43	100.00	100.00	100.00	100.00	70.27
Bare ground	Producer's	86.67	84.62	76.92	100.00	80.00	80.00	41.67
	User's	92.86	78.57	66.67	86.67	85.71	92.31	35.71
Sparse vegetation	Producer's	88.89	69.23	84.62	76.92	100.00	100.00	0.00
	User's	61.54	90.00	100.00	71.43	81.82	75.00	0.00
Mixed agricultural	Producer's	88.24	100.00	90.48	90.48	88.24	82.35	60.00
	User's	100.00	84.00	95.00	100.00	93.75	93.33	22.50
Vineyard and orchard	Producer's	75.00	85.71	100.00	64.29	87.50	87.50	100.00
	User's	80.00	63.16	73.68	75.00	100.00	93.33	66.67
Herbaceous vegetation	Producer's	80.00	90.48	90.48	95.24	96.00	92.00	44.00
	User's	95.24	90.48	82.61	86.96	92.31	85.19	84.62
Garrigue	Producer's	92.31	35.71	64.29	64.29	100.00	76.92	28.57
	User's	54.55	50.00	81.82	56.25	81.25	66.67	28.57
Mediterranean maquis	Producer's	63.16	69.23	69.23	84.62	100.00	94.74	47.06
	User's	80.00	69.23	81.82	73.33	100.00	94.74	72.73
Conifer forests	Producer's	68.75	83.33	0.00	83.33	87.50	100.00	100.00
	User's	91.67	76.92	0.00	66.67	87.50	84.21	84.21
Broadleaf forests	Producer's	90.91	76.47	94.12	76.47	81.82	81.82	78.26
	User's	62.50	92.86	57.14	92.86	81.82	100.00	94.74

4. DISCUSSION

High accuracies were reached by combining multispectral and LiDAR data in the fuel type classification, especially if we compare these results with the LUM of Sardinia which showed very low accuracy coefficients.

The higher values of overall accuracy respect to the Kappa coefficient are common since latter is a more conservative measure than the overall classification accuracy. According to both accuracy coefficients (overall accuracy and Kappa coefficient), the best method for classifying this kind of data in these fuel types is the RF algorithm, followed by the SVM algorithm. The least performing algorithm was the NN contrary to what Frizzelle and Moody (2001) found in their study of characterization of land cover from multispectral data. In their work they obtained better results with the NN algorithm than the ML maybe because they used few variables and the classified only eight land cover categories. However, in agreement with our results, García *et al.* (2011) found in their study combining also LiDAR and Multispectral data, that SVM had higher potential for combing different data sources than ML. Also in accordance with our results, in the study of Pal and Mather (2005) RF showed accuracy coefficients slightly better than SVM, even if also in this case the classification was carried out using only multispectral data for seven classes.

The most similar work to our study is the one carried out by Reese *et al.* (2014) since they used Spot-5 and LiDAR data to classify 12 classes of vegetation using the RF algorithm. However, they made it for alpine vegetation which includes mainly different types of herbaceous vegetation and shrubs and two classes of broadleaf forests. Their accuracies coefficients were much lower than ours (63.1% of overall accuracy).

Also the classifications and the LUM of Sardinia differ in the fuel distribution of the study areas (Figs. 1.3, 1.4, 1.5, and 1.6), probably due to the coarser resolution of the latter.

Regarding the performance per fuel type, besides 'Water', which showed very high accuracies in both areas, the rest of fuel types showed very different values in each study area. This is probably because of the very different spatial distribution of fuel types in each area, which makes easier to define certain fuel types in each area.

However we selected these two study areas for this reason (different distribution of fuels characteristics of the island) and even if the results differ, in both cases showed high values of accuracy. Also in our study RF is the algorithm showing the most balanced performance regarding the results per fuel types with high values of both, producer's and user's accuracy.

5. CONCLUSIONS

Regarding the objective of this study, we can conclude that it is possible to create high accuracy fuel maps by combining multispectral and LiDAR data and the resulting maps improve the available information until now.

Even if in this case we applied the methodology only to a small study area, the same methodology could be extended to other areas. Moreover, if new LiDAR data will be available for the whole island in the next years, we could update the fuel maps by following this methodology.

In this study, the RF algorithm carried out with a subset of variables showed the best performance for the fuel type classification. However, RF with all variables performs a reasonably accurate classification (even if the results are not good as with selected variables). Therefore, since the selection of variables suppose extra-work and time, for a quick classification or if the classification should be done for lots of different areas, this methodology could be applied.

Further work should be focus on the improvement of the variable selection for automatize the process and optimize the results.

Acknowledgments

We would like to acknowledge the Autonomous Region of Sardinia for providing the LiDAR data and the orthofotos. In addition, we also thank the Forest Service of Sardinia for the Spot imagery.

References

Ager AA, Preisler HK, Arca B, Spano D, Salis M (2014) Wildfire risk estimation in the Mediterranean area. *Environmetrics* **25**, 384–396. doi:10.1002/env.2269.

Alcasena FJ, Salis M, Ager AA, Arca B, Molina D, Spano D (2015) Assessing Landscape Scale Wildfire Exposure for Highly Valued Resources in a Mediterranean Area. *Environmental Management* **55**, 1200–1216. doi:10.1007/s00267-015-0448-6.

Arca B, Bacciu V, Pellizzaro G, Salis M, Ventura A, Duce P, Brundu G (2009) Fuel model mapping by Ikonos imagery to support spatially explicit fire simulators. In: Spano *et al*. (eds) 'International Conference on Fire Behaviour and Risk', Matera. pp 92–98.

Arca B, Duce P, Pellizzaro G, Bacciu V, Salis M, Spano D (2007) Evaluation of Farsite Simulator in Mediterranean maquis. *International Journal of Wildland Fire* **16**, 563–572. doi:10.1071/WF06070.

Arroyo LA, Pascual C, Manzanera A (2008) Fire models and methods to map fuel types: The role of remote sensing. *Forest Ecology and Management* **256**, 1239–1252. doi:10.1016/j.foreco.2008.06.048.

Autonomous Region of Sardinia (2008) Carta dell'Uso del Suolo in scala 1:25.000. http://www.sardegnageoportale.it/index.php?xsl=1598&s=291548&v=2&c=8831&t=1.

Bajocco S, Dragoz E, Gitas I, Smiraglia D, Salvati L, Ricotta C (2015) Mapping forest fuels through vegetation phenology: The role of coarse-resolution satellite time-series. *PLoS ONE* **10**, 1–14. doi:10.1371/journal.pone.0119811.

Breiman L (2001) Random forests. *Machine Learning* **45**, 5–32. doi:10.1023/A:1010933404324.

Chuvieco E (2010) 'Teledetección ambiental. La observación de la Tierra desde el Espacio.' (Ariel Ciencia: Barcelona)

Chuvieco E, Wagtendok J, Riaño D, Yebra M, Ustin SL (2009) Estimation of fuel conditions for fire danger assessment. In: Chuvieco (ed) 'Earth observation of wildland fires in Mediterranean ecosystems' pp. 83–96. (Springer: Berlin Heidelberg)

Congalton RG (1991) A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* **37**, 35–46. doi:10.1016/0034-4257(91)90048-B.

EEA (2011) Version 15 of raster data on land cover for the corine land cover 2006 inventory. European Environmental Agency, http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2006-raster-1.

Falkowski MJ, Evans JS, Martinuzzi S, Gessler PE, Hudak AT (2009) Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. *Remote Sensing of Environment* **113**, 946–956. doi:10.1016/j.rse.2009.01.003.

Foody GM (2006) What is the difference between two maps? A remote senser's view. *Journal of Geographical Systems* **8**, 119–130. doi:10.1007/s10109-006-0023-z.

Frizzelle BG, Moody A (2001) Mapping Continuous Distributions of Land Cover : A Comparison of Maximum-Likelihood. Estimation and Artificial Neural Networks. *Photogrammetric Engineering & Remote Sensing* **67**, 693–705.

García M, Riaño D, Chuvieco E, Salas J, Danson FM (2011) Multispectral and LiDAR data fusion for fuel type mapping using Support Vector Machine and decision rules. *Remote Sensing of Environment* **115**, 1369–1379. doi:10.1016/j.rse.2011.01.017.

Isenburg M (2015) LAStools - Efficient Tools for Lidar Processing. Version 140329. http://lastools.org.

Koetz B, Morsdorf F, Van Der Linden S, Curt T, Allgöwer B (2008) Multi-source land cover classification for forest fire management based on imaging spectrometry and LiDAR data. *Forest Ecology and Management* **256**, 263–271. doi:10.1016/j.foreco.2008.04.025.

Lasaponara R, Lanorte A (2007a) Remotely sensed characterization of forest fuel types by using satellite ASTER data. *International Journal of Applied Earth Observation and Geoinformation* **9**, 225–234. doi:10.1016/j.jag.2006.08.001.

Lasaponara R, Lanorte A (2007b) On the capability of satellite VHR QuickBird data for fuel type characterization in fragmented landscape. *Ecological Modelling* **204**, 79–84. doi:10.1016/j.ecolmodel.2006.12.022.

Mallinis G, Galidaki G, Gitas I (2014) A Comparative Analysis of EO-1 Hyperion, Quickbird and Landsat TM Imagery for Fuel Type Mapping of a Typical Mediterranean Landscape. *Remote Sensing* **6**, 1684–1704. doi:10.3390/rs6021684.

Mallinis G, Mitsopoulos ID, Dimitrakopoulos AP, Gitas IZ, Karteris M (2008) Local-Scale Fuel-Type Mapping and Fire Behavior Prediction by Employing High-Resolution Satellite Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **1**, 230–239. doi:10.1109/JSTARS.2008.2011298.

Marino E, Ranz P, Luis J, Ángel M, Esteban J, Madrigal J (2016) Remote Sensing of Environment Generation of high-resolution fuel model maps from discrete airborne laser scanner and Landsat-8 OLI : A low-cost and highly updated methodology for large areas. *Remote Sensing of Environment* **187**, 267–280. doi:10.1016/j.rse.2016.10.020.

Martin M., Newman S., Aber J., Congalton R. (1998) Determining forest species composition using high spectral resolution remote sensing data. *Remote Sensing of Environment* **65**, 249–254. doi:10.1016/S0034-4257(98)00035-2.

McGaughey RJ (2014) FUSION/LDV: Software for LIDAR Data Analysis and Visualization. 154.

Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* **66**, 247–259. doi:10.1016/j.isprsjprs.2010.11.001.

Ottmar RD, Alvarado E (2004) Linking Vegetation Patterns to Potential Smoke Production and Fire Hazard 1. USDA Forest Service, Pacific Northwest Research Station, General Technical Report PSW-GTR-193. (Seattle, WA)

Otukei JR, Blaschke T (2010) Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation* **12**, 27–31. doi:10.1016/j.jag.2009.11.002.

Pal M (2005) Random forest classifier for remote sensing classification. *International Journal of Remote Sensing* **26**, 217–222. doi:10.1080/01431160412331269698.

Pal M, Mather PM (2005) Support vector machines for classification in remote sensing. *International Journal of Remote Sensing* **26**, 1007–1011. doi:10.1080/01431160512331314083.

Qiu F, Jensen JR (2004) Opening the black box of neural networks for remote sensing image classification. *International Journal of Remote Sensing* **25**, 1749–1768. doi:10.1080/01431160310001618798.

R Core Team (2016) R: A language and environment for statistical computing. http://www.r-project.org/.

Reese H, Nyström M, Nordkvist K, Olsson H (2014) Combining airborne laser scanning data and optical satellite data for classification of alpine vegetation. *International Journal of Applied Earth Observation and Geoinformation* **27**, 81–90. doi:10.1016/j.jag.2013.05.003.

Richter R, Schläpfer D (2012) ATCOR-2/3 User Guide, Version 8.2.0. 223.

Salis M, Ager AA, Arca B, Finney MA, Bacciu V (2013) Analyzing wildfire exposure to human and ecological values in Sardinia, Italy. *International Journal of Wildland Fire* **22**, 549–565. doi:10.1071/WF11060.

Salis M, Del Giudice L, Arca B, Ager AA, Alcasena-urdiroz F, Lozano O, Bacciu V, Spano D (2018) Modeling the effects of different fuel treatment mosaics on wild fi re spread and behavior in a Mediterranean agro-pastoral area. *Journal of Environmental Management* **212**, 490–505. doi:10.1016/j.jenvman.2018.02.020.

Salis M, Laconi M, Ager AA, Alcasena FJ, Arca B, Lozano O, Fernandes De Oliveira A, Spano D (2016) Evaluating alternative fuel treatment strategies to reduce wildfire losses in a Mediterranean area. *Forest Ecology and Management* **368**, 207–221. doi:10.1016/j.foreco.2016.03.009.

Senseman GM, Bagley CF, Tweddale SA (1995) Accuracy Assessment of the Discrete Classification of Remotely-Sensed Digital Data for Landcover Mapping. US Army Corps of Engineers, Construction Engineering Research Laboratories. CERL-TR-EN-95/04

Shalaby A, Tateishi R (2007) Remote sensing and GIS for mapping and monitoring land cover and land-use changes in the Northwestern coastal zone of Egypt. *Applied Geography* **27**, 28–41. doi:10.1016/j.apgeog.2006.09.004.

Soenen SA, Peddle DR, Coburn CA (2005) SCS+C: A modified sun-canopy-sensor topographic correction in forested terrain. *IEEE Transactions on Geoscience and Remote Sensing* **43**, 2148–2159. doi:10.1109/TGRS.2005.852480.

Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A (2008) Conditional variable importance for random forests. *BMC Bioinformatics* **9**, 1–11. doi:10.1186/1471-2105-9-307.

Strobl C, Boulesteix AL, Zeileis A, Hothorn T (2007) Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics* **8**,. doi:10.1186/1471-2105-8-25.

Valbuena R, Maltamo M, Packalen P (2016) Classification of forest development stages from national low-density lidar datasets: a comparison of machine learning methods. *Revista de Teledetección* **0**, 15. doi:10.4995/raet.2016.4029.

Varga TA, Asner GP (2008) Hyperspectral and LiDAR remote sensing of fire fuels in Hawaii Volcanoes National Park. *Ecological Applications* **18**, 613–623. doi:10.1890/07-1280.1.

Walter V (2004) Object-based classification of remote sensing data for change detection. *ISPRS Journal of Photogrammetry and Remote Sensing* **58**, 225–238. doi:10.1016/j.isprsjprs.2003.09.007.

Wiesmann D, Quinn D (2016) Rassclass: Supervised Raster Image Classification. R package version 0.2.2.

Yan WY, Shaker A, El-Ashmawy N (2015) Urban land cover classification using airborne LiDAR data: A review. *Remote Sensing of Environment* **158**, 295–310. doi:10.1016/j.rse.2014.11.001.

Yu X, Hyyppä J, Vastaranta M, Holopainen M, Viitala R (2011) Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS Journal of Photogrammetry and Remote Sensing* **66**, 28–37. doi:10.1016/j.isprsjprs.2010.08.003.

Yuan H, Van Der Wiele CF, Khorram S (2009) An automated artificial neural network system for land use/land cover classification from landsat TM imagery. *Remote Sensing* **1**, 243–265. doi:10.3390/rs1030243.

Chapter 2: Characterizing Mediterranean canopy fuel properties from LiDAR data

1. INTRODUCTION

The knowledge of the canopy fuel characteristics is for several aspects related to fire management, as for instance for predicting crown fire occurrence and behavior (Scott and Reinhardt 2002). Therefore, the characterization of the canopy should be a preliminary work necessary for forest managers in case of wildfires as well as for using fire spread models.

Wildfire spread models are efficient tools not only for simulating single fires but they have been also used for assessing potential wildfire risk, optimizing fuel treatments, etc (Finney 2006; Jahdi *et al.* 2015; Lozano *et al.* 2017; Salis *et al.* 2018). Fuel distribution maps are very important for wildfire spread modeling, but they are not the only necessary input. Widely used fire spread models such as FARSITE (Finney 1998) and FlamMap (Finney 2006) require spatially explicit estimation of canopy fuel characteristics to simulate also the crown fire. These required variables are: (1) Stand Height (SH) which is the average height of the dominant tree layer; (2) Canopy Base Height (CBH) which is the average height of the bottom of the tree crowns in the stand; and (3) Canopy Bulk Density (CBD) which is the density of the crown biomass above the shrub layer (Keane *et al.* 1998).

As well as for the fuel maps, it is very difficult characterize canopy fuel structure based only on field surveys and therefore indirect estimation methods are needed (Hevia *et al.* 2016). Optical remote sensing methods can be used to assess some vegetation characteristics over wide areas for relatively low costs (Arroyo *et al.* 2008; Claggett *et al.* 2010; Mallinis *et al.* 2013). However, these techniques are unable to accurately characterize canopy structural attributes while light detection and ranging (LiDAR) pulses penetrate the canopy and allows characterizing the forest structure (Wasser *et al.* 2013). In fact, Riaño et al. (2003) modeled airborne laser scanning data for producing forest parameters used in fire behavior modeling. They used high density pulse LiDAR data which is only available for small areas because of its high cost. However, even

with low point density LiDAR data, some studies proved that it is possible to characterize some canopy fuel properties for conifer species, but for broadleaf areas there are not many studies (Erdody and Moskal 2010; Brubaker *et al.* 2014; González-Ferreiro *et al.* 2014; Hermosilla *et al.* 2014; Hevia *et al.* 2016). In 2005, Andersen, Mcgaughey, and Reutebuch (2005) estimated conifer canopy fuel parameters using high-density LIDAR data. Hermosilla et al. (2014) estimated forest structure and canopy fuel parameters in a conifer forest from small-footprint full-waveform LiDAR data. In other studies, canopy fuel variables have been modelled using in this case medium or low-density LiDAR data but only in conifer stands (Hall *et al.* 2005; Zhao *et al.* 2011; González-Ferreiro *et al.* 2014, 2017). Brubaker et al. (2014) estimated only the canopy height using low density LiDAR but in a deciduous forest.

The present study aims at improving the characterization of broadleaf forest fuel models using low-density LiDAR data in a Mediterranean area. Namely we wanted to obtain canopy input data, which are also required by wildfire spread models such as Stand Height, Crown Base Height and Crown Bulk Density.

2. METHODS

2.1. Study area

The study area (between 39° 23' and 39° 25' N latitude and 9° 32' and 9° 35' E longitude , Fig. 2.1) is located in Muravera which is a municipality of south-eastern Sardinia (Italy). This area is close to the sea and therefore there are not very high altitudes (up to 572 m.a.s.l.). The area covers about 400 ha and was selected because it includes the largest extent of broadleaf evergreen forest within Sardinia areas for which LiDAR data are available. The forest includes as main species *Quercus ilex* L. and other secondary species such as *Olea europaea* L. var. *sylvestris*, *Juniperus oxycedrus* L., *Phillyrea latifolia* L., *Arbutus unedo* L., *Pistacia lentiscus* L., *Quercus suber* L.

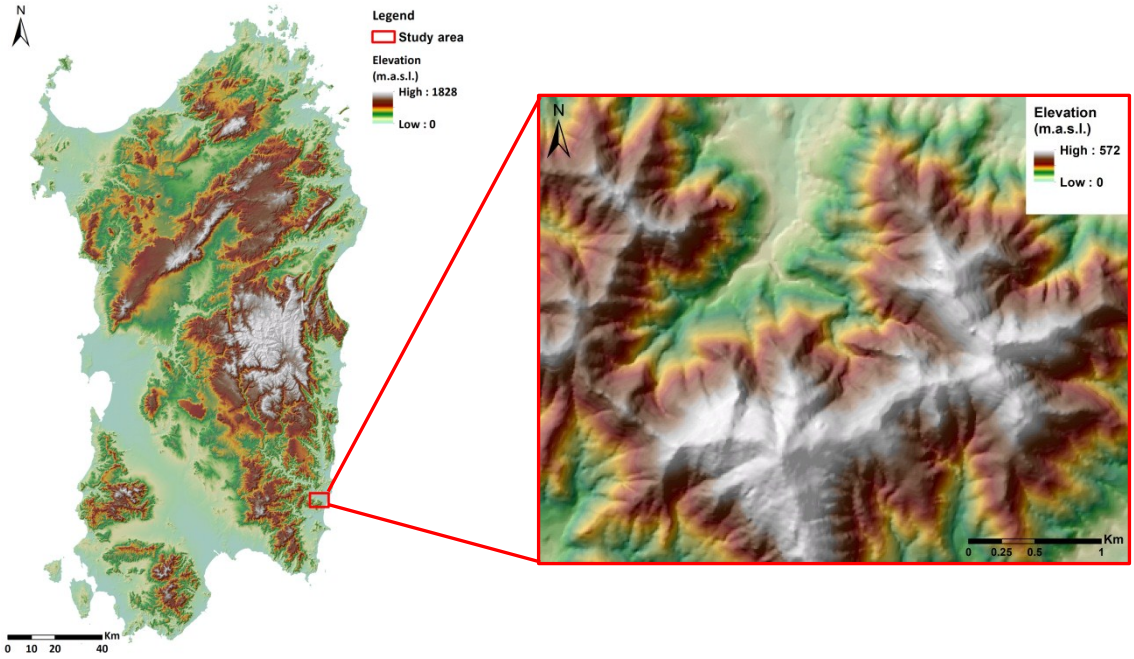

Fig. 2.1. Location of the study area of Muravera in south-eastern Sardinia

2.2. LiDAR data

Lidar data processed in chapter 1 corresponding to the Muravera case study were used for this chapter as well.

2.3. Field data

The field work provided the data to fit the model. The sampling design is presented in the following lines.

Plot shape

We decided to use circular plots because they are easier to georeference since only a single point is needed. Furthermore, in circular plots the perimeter-to-area ratio is minimum, therefore also the negative impact of the edge-effect on the LiDAR metrics is minimum (Frazer *et al.* 2011).

Plot size

Some studies suggested that plot size should be the same as cloud point raster pixel size (Magnussen and Boudewyn 1998; Condés *et al.* 2013). In our study we were not able to change the pixel size of the cloud point raster since we wanted to combine our outputs with those obtained in chapter 1 (determined by the resolution of the multispectral image which is 10 m). Therefore, our field sampling plots had to be 100 m^2. However, it was demonstrated that from larger field plots more accurate results may be achieved (Gobakken and Næsset 2009; Frazer *et al.* 2011; Zolkos *et al.* 2013; Ruiz *et al.* 2014; Hansen *et al.* 2015). Thus, we decided to use 10 m radius plots (314 area m^2 approximately). Ruiz et al. (2014) assessed the combined effect of field plot size and LiDAR density on the estimation of four forest structure variables (volume, total biomass, basal area and canopy cover). For a 314 m^2 field plot with a density of 0.8 points/m^2, their results showed coefficients of determination (R^2) ranging from 0.78 to 0.86 for the different variables. Namely, for the canopy cover they suggested plot areas of 300-400 m^2.

Plot number

Accuracy of outputs is not only related to plot size but also to plot number (Zeide 1980). It is possible to minimize the number of plots if the sampling design covers the whole variability (Condés *et al.* 2013). For this reason, the stratified sampling based on LiDAR data ensures that the entire data range of the predictor variables is sampled, and, consequently, the predictions by LiDAR-derived regression models are better (Hawbaker *et al.* 2009; Frazer *et al.* 2011; Condés *et al.* 2013). The locations of our plots were derived from a stratified sampling based on two LiDAR variables: mean height and standard deviation of height. The stratification was carried out first by the mean height (three strata), and then each stratum was further stratified into three strata according to the standard deviation (Fig. 2.2). Then we build a map with these combined strata (Fig. 2.3). A total of 27 plots were set and the number of plots per stratum was assigned according to their frequency in the map (Table 2.2).

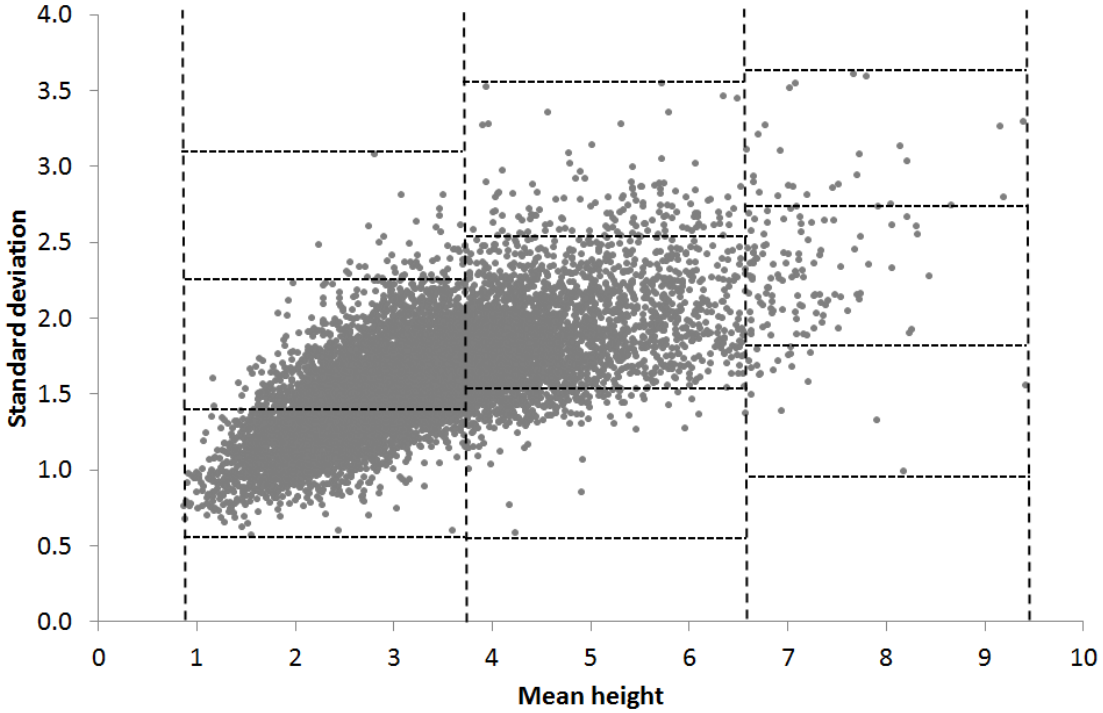

Fig. 2.2. Scatterplot showing the values of the mean height and standard deviation (derived from LiDAR data) for each pixel of the study area. Data were stratified first by the mean height (x axis) and then by the standard deviation of LiDAR pulse height (y axis). Dash lines show the breaks among different strata.

Table 2.2. Number of field plots assigned to each combined stratum. H1, H2 and H3 means the first, second and third strata of mean height respectively; and SD1, SD2 and SD3 are the first, second and third strata of standard deviation respectively (inside each mean height interval)

Strata	Number of plots
H1 - SD1	6
H1 - SD2	8
H1 - SD3	1
H2 - SD1	2
H2 - SD2	6
H2 - SD3	1
H3 - SD1	1
H3 - SD2	1
H3 - SD3	1
Total	**27**

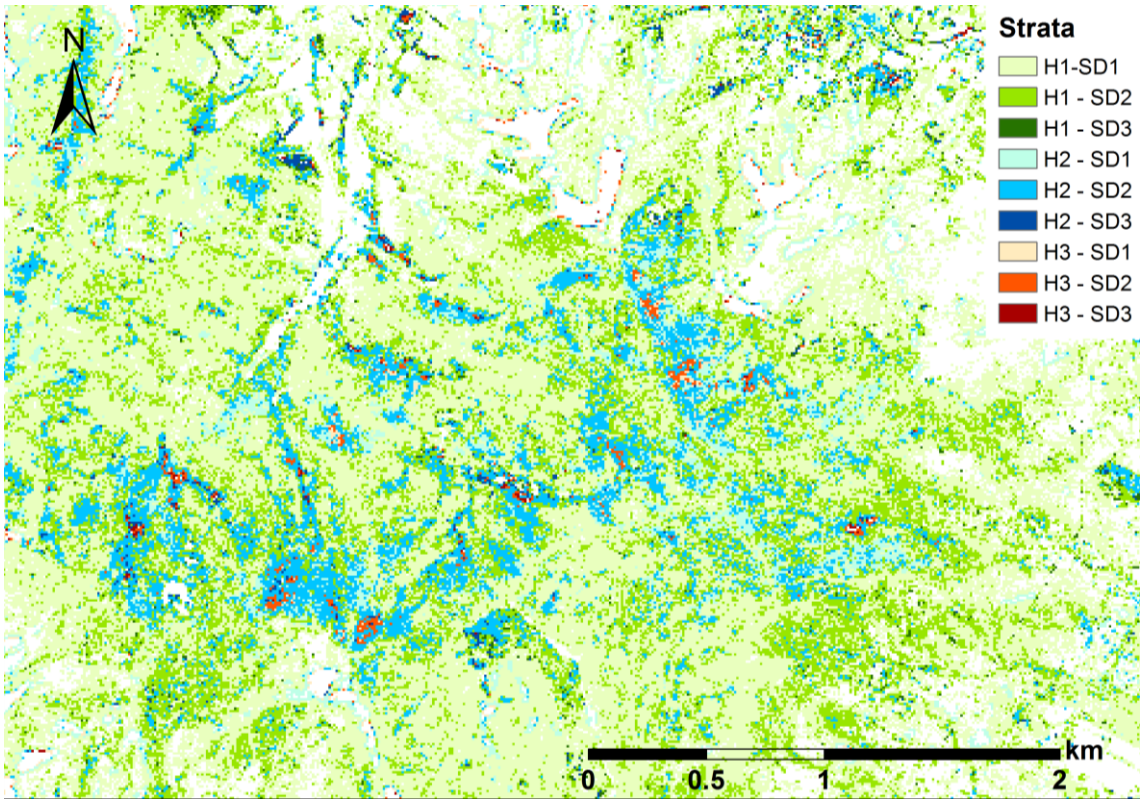

Fig. 2.3. Map of the combining strata resulting from the stratification of LiDAR data according to the mean height and standard deviation: H1, H2 and H3 means the first, second and third strata of mean height respectively; and SD1, SD2 and SD3 are the first, second and third strata of standard deviation respectively (inside each mean height interval).

Parameters measured in each plot

We used handheld GPS navigator to reach the sampling plots. Then, we logged the coordinates of the center of the circular plot using a dual-frequency, geodetic, real-time RTK receiver (Leica GPS900) with a horizontal precision of 10mm + 1ppm and a vertical precision of 20mm + 1ppm.

At each plot, we recorded the vegetation species and measured the diameters at breast height (DBH) of all trees (when DBH was higher than 7.5 cm). We verified the distance of each tree to the center of the plot by means of a laser vertex and a transponder (Haglof Vertex Laser VL400).

We additionally measured the total height, the crown base height and crown projection diameters (two cross diameters) of four trees in the plot (the most northern, most eastern, most western and most southern tree of each plot).

To measure tree diameters we used a tree caliper, for total heights we used a handheld laser rangefinder/clinometer (TruPulse 360, Laser Technology), for crown base height we used a 3-m folding rule, and for crown projection diameters we used a 10-m measuring tape.

Since we were only allowed to carry out non-destructive inventories in the study area, we had to indirectly calculate the Crown Bulk Density (CBD). Firstly, we calculated the crown volume as an ellipsoid which diameters were the two cross diameters of the crown projection and the third diameter was the difference between the total height and the crown base height. Then we calculated the CBD using the equations for the crown dry weight proposed by Tabacchi et al. (2011).

2.4. Modeling canopy structure inside the sampling plots (allometric models)

Field work provided diameter measures for all trees within the plot, but only few measures for height, CBH and canopy volume. To obtain these parameters for unmeasured trees we built Linear Mixed Models (LMMs) implemented on the R package lme4 (Bates *et al.* 2015). LMMs include random effects allowing to avoid pseudo-replication and to evaluate conditional effects on parameter estimates (Bolker *et al.* 2008). We used tree diameter as fixed effect for tree height and canopy volume, whereas for CBH we used tree height as fixed effect because preliminary analysis reported no relationship between tree diameter and CBH. Instead, as random effect we used Plot ID over both the intercept and slope. The Akaike Information Criteria (AIC) was used for model selection (Burnham and Anderson 1998). Models were fitted using the maximum likelihood algorithm because other quasi-likelihood methods (e.g. restricted maximum likelihood) are not adequate for inference (e.g. AIC, Bolker et al. 2008). To assess the variance explained of the selected models we used the R^2 for mixed models of Nakagawa and Schielzeth (2013). The R^2 of Nakagawa and Schielzeth reports both the variance explained by the model (conditional) and the fixed part only (marginal). The R^2 of Nakagawa and Schielzeth was implemented using the R package MuMIn.

2.5. Variable selection

We extracted from the LiDAR data the descriptive statistics corresponding to the field plots using the CloudMetrics tool in FUSION (McGaughey 2014). Even if we left out variables with constant values and relative variables, the database still had 116 variables (Cloudmetrics output metrics are the same as those from Gidmetrics, see Chapter 1, 2.4. LiDAR data processing). Since we wanted to test linear equations as canopy structure models, we had to reduce the number of variables. The first step was to split the database in seven databases as follows: percentiles of heights, other metrics of heights, canopy cover metrics and the metrics of strata grouped in four groups (strata below 1 meter, strata between 1 and 3 meters, strata between 3 and 10 meters and strata above 10 meters). Then we run a principal components analysis to each group and we selected the most explanatory variables (original variables and not principal components) of each group. At the end, we created a database with 33 variables (Table 2.2)

Table 2.2. Selected variables to fit the canopy structure models

Variable name	Description
Elev minimum	minimum
Elev mean	mean
Elev mode	mode
Elev stddev	standard deviation
Elev CV	coefficient of variation
Elev skewness	skewness of elevations
Elev kurtosis	kurtosis of elevations
Elev P05	5th percentile
Elev P05	25th percentile
Elev P70	70th percentile
Elev P95	95th percentile
Canopy relief ratio	(mean height- min height) / (max height– min height)
Canopy cover	Percentage first returns above 2.00 m
allcover	Percentage all returns above 2.00 m
allabovemean	Percentage all returns above mean
allabovemode	Percentage all returns above mode
Elev strata below 0.50 return proportion	(Total return count for the strata)/(all returns)
Elev strata 0.50 to 1.00 return proportion	(Total return count for the strata)/(all returns)
Elev strata 0.50 to 1.00 min	Minimum elevation for the strata
Elev strata 0.50 to 1.00 mean	Average elevation for the strata
Elev strata 0.50 to 1.00 mode	Mode of elevations for the strata
Elev strata 0.50 to 1.00 skewness	Skewness of elevations within the the strata
Elev strata kurtosis	Kurtosis of elevations within the the strata
Elev strata 1.00 to 2.00 return proportion	(Total return count for the strata)/(all returns)
Elev strata 1.00 to 2.00 min	Minimum elevation for the strata
Elev strata 1.00 to 2.00 mean	Average elevation for the strata
Elev strata 1.00 to 2.00 skewness	Skewness of elevations within the the strata
Elev strata return 2.00 to 3.00 proportion	(Total return count for the strata)/(all returns)
Elev strata 2.00 to 3.00 mean	Average elevation for the strata
Elev strata 2.00 to 3.00 stddev	Standard deviation of elevations within the the strata
Elev strata return 3.00 to 5.00 proportion	(Total return count for the strata)/(all returns)
Elev strata return 5.00 to 10.00 proportion	(Total return count for the strata)/(all returns)
Elev strata return above 10.00 proportion	(Total return count for the strata)/(all returns)

2.6. Canopy structure modeling for the whole study area

Combining the collected field data with the 33-variables database from LiDAR data we wanted to test different models for the following canopy variables: average stand height, average crown base height and average crown volume. We performed a stepwise selection for each response variable to identify the most explanatory variables and then we tried different linear regressions combining these variables. After the stepwise analysis and trying different variables and combinations of variables, we selected the models prioritizing the minimum number of variables, the lowest AIC, the highest adjusted correlation coefficient (R^2) and the significance of coefficients.

For the best performing models we calculated the variance inflation factors to test the multicollinearity and we plotted some diagnostics graphics (Residuals v. predicted values, Regression Influence Plot, Residuals v. Leverage and Normal Q-Q) to evaluate the model assumptions and to investigate if there were observations with an undue influence on the analysis. Then also an outlier test was run (Bonferroni Outlier Test, R Core Team 2016) to confirm or not the presence of outlier plots. We also tried to find the best model using the Random Forest algorithm with all variables or with the set of 33 variables previously selected.

When we selected the best models, we extended them to the whole area using the ASCII files as predictive variables and obtaining as a result three maps (one for each canopy variable).

3. RESULTS

3.1. Field work

We measured a total of 969 trees within the 27 sampling plots (Table 2.3). The majority of trees sampled were *Quercus ilex* L.; the other species (which frequencies were less than 10%) were *Arbutus unedo* L., *Phillyrea latifolia* L., *Juniperus oxycedrus* L., *Olea europaea* L. var. *sylvestris*, *Pistacia lentiscus* L. and *Quercus suber* L.

Table 2.3. Summary of field data collected. Stratum refers to the strata in which study area data was divided according to the mean height and the standard deviation of LiDAR pulse height. In the first column of the table, H1S1 means first range of mean height and first range of standard deviation, and so on for the other codes. Diameter at Breast Height (DBH) was measured in all trees of each plot. (*) The values of these parameters were calculated for four trees inside each plot.

	Plot	Number of trees	Mean dbh	Total Height*	Crown Base Height*	Crown Volume*
H1S1	1	14	11.16	4.70	1.36	148.53
	2	9	12.56	3.14	0.84	37.76
	3	20	11.79	5.48	1.48	93.39
	4	31	11.35	4.94	1.34	77.14
	5	20	11.65	4.85	1.38	90.10
	6	21	11.15	4.24	1.29	64.51
H1S2	7	14	14.54	5.73	1.00	288.66
	8	23	12.27	4.40	1.21	148.75
	9	26	13.90	7.23	1.99	260.25
	10	28	13.45	6.65	1.56	125.03
	11	21	11.37	6.10	2.16	62.77
	12	18	11.72	5.79	1.54	122.64
	13	50	11.07	5.70	1.56	124.95
	14	24	12.51	4.99	1.16	90.93
H1S3	15	36	15.60	7.80	2.98	275.28
H2S1	16	75	10.65	7.15	2.39	125.71
	25	80	10.19	5.64	1.44	69.64
H2S2	17	37	13.24	10.15	2.76	686.94
	18	32	16.00	6.93	2.05	493.58
	19	64	12.28	7.85	1.85	180.04
	20	60	12.28	8.75	2.73	256.89
	21	25	12.31	5.71	1.64	87.01
	27	56	12.64	8.05	1.85	493.49
H2S3	22	39	14.28	6.03	2.41	53.75
H3S1	23	62	10.57	8.05	2.37	220.59
H3S2	24	69	10.75	7.53	2.20	95.45
H3S3	26	15	22.18	13.15	2.98	843.13

3.2. Allometric models

As far as tree height and volume are concerned, the best model (i.e. lowest AIC) was that obtained using plot ID in the intercept as random effect. Instead, for CBH the models with random factors did not perform better. The conditional variance explained by the models was the 70, 80, and 33 for tree height, volume and CBH, respectively. Figure 4 shows the fitted relationships for the three variables.

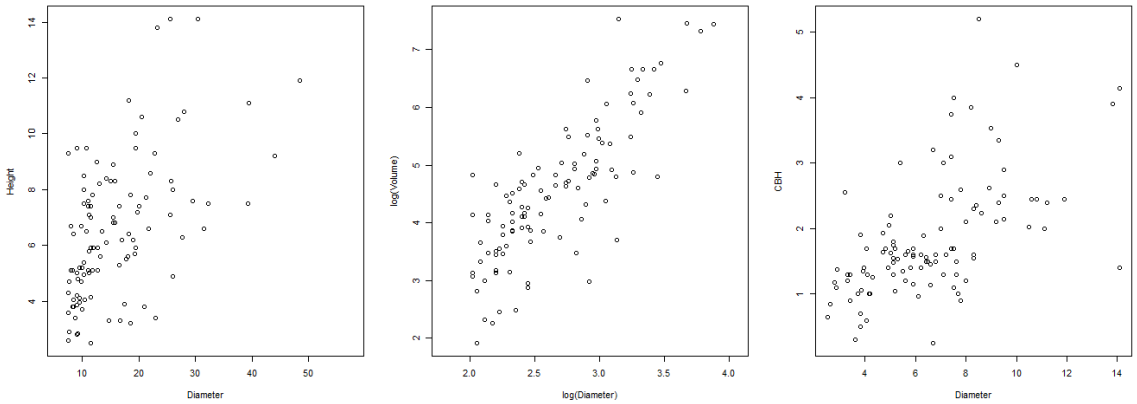

Fig. 2.4. Allometric fitted relationships.

3.3. Models of canopy variables

After the stepwise analysis and trying different variables and combinations of variables, we selected the models by prioritizing the minimum number of variables, the lowest AIC, the highest adjusted correlation coefficient (R^2) and the significance of coefficients.

The selected models and their respective coefficients of determination (R^2) for each canopy fuel variable are summarized in Table 2.4. For all three models, the coefficients of determination are above 0.7.

Scatter plots of field-measured against model values for the three studied variables are shown in Fig. 2.5. In these charts, all plots are near to the 1:1 line, except for plot 26 in SH and Crown Volume.

Table 2.4. Results of the best performing models for the average Stand Height (SH), the average Crown Base Height (CBH) and the average crown Volume (V). Significance codes: (***) = 0; (**) = 0.001; (*) = 0.01; (') = 0.05; ( ) = 0.1. The LiDAR elevations variables are: Elev P70 = 70th percentile; Elev P95 = 95th percentile; Elev stddev = standard deviation; Elev kurtosis = kurtosis; allcover = Percentage of all returns above 2 m.

Selected models	**Adjusted R-squared**
SH = 1.56069** + 0.92950*** · Elev P70	0.7912
CBH = 0.60437*** + 0.15810*** · Elev P70 + 0.21445 · Elev stddev	0.8140
V = 2.55952*** + 0.22290*** · Elev P95 - 0.27668* · Elev kurtosis + 0.0223' · allcover	0.7261

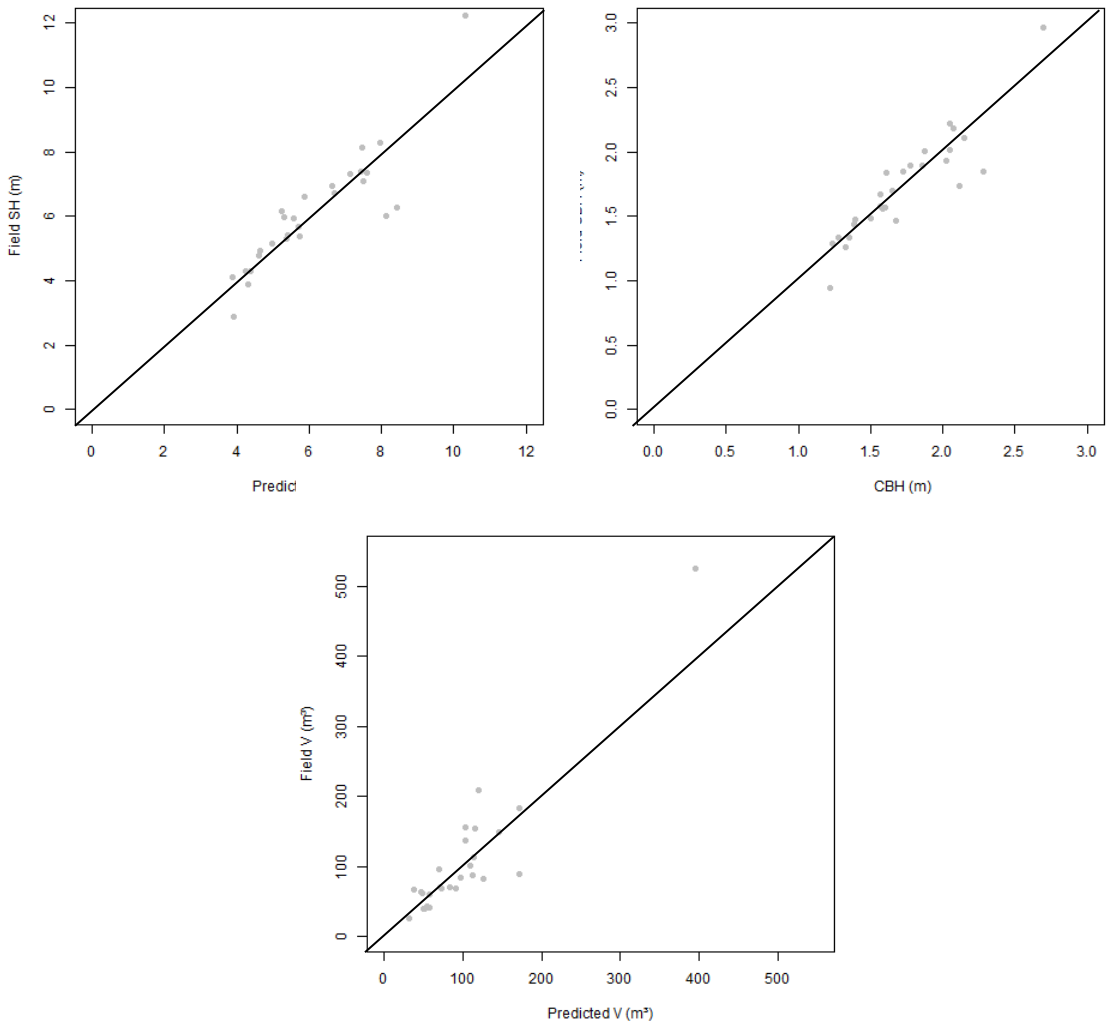

Fig. 2.5. Scatter plot of field-measured values v. model predicted results for the three studied canopy variables: average Stand Height (SH); average Crown Base Height (CBH); and the average crown Volume (V).

Diagnostics graphics showed some plots like 26, 18 and 22 distant from the other ones in terms of influence and residuals (Fig. 2.6). In the residuals v. predicted values plot, especially for the SH, the 26 plot shows a large positive value, while plots 18 and 22 show a large negative value. In addition, both, Regression Influence Plot and Residuals v. Leverage plot, show that plot 26 had a high value of leverage for the three variables. Regarding the Cook's D statistic, again plots 26, 22 and 18 show high values for the three variables, whereas plots 7 and 8 show high variables only for crown volume. Plots 26, 22 and 18 are far away from normal observations in the Normal Q-Q plot for SH. In case of CBH only plots 22 and 18 do not follow the normal observation's line.

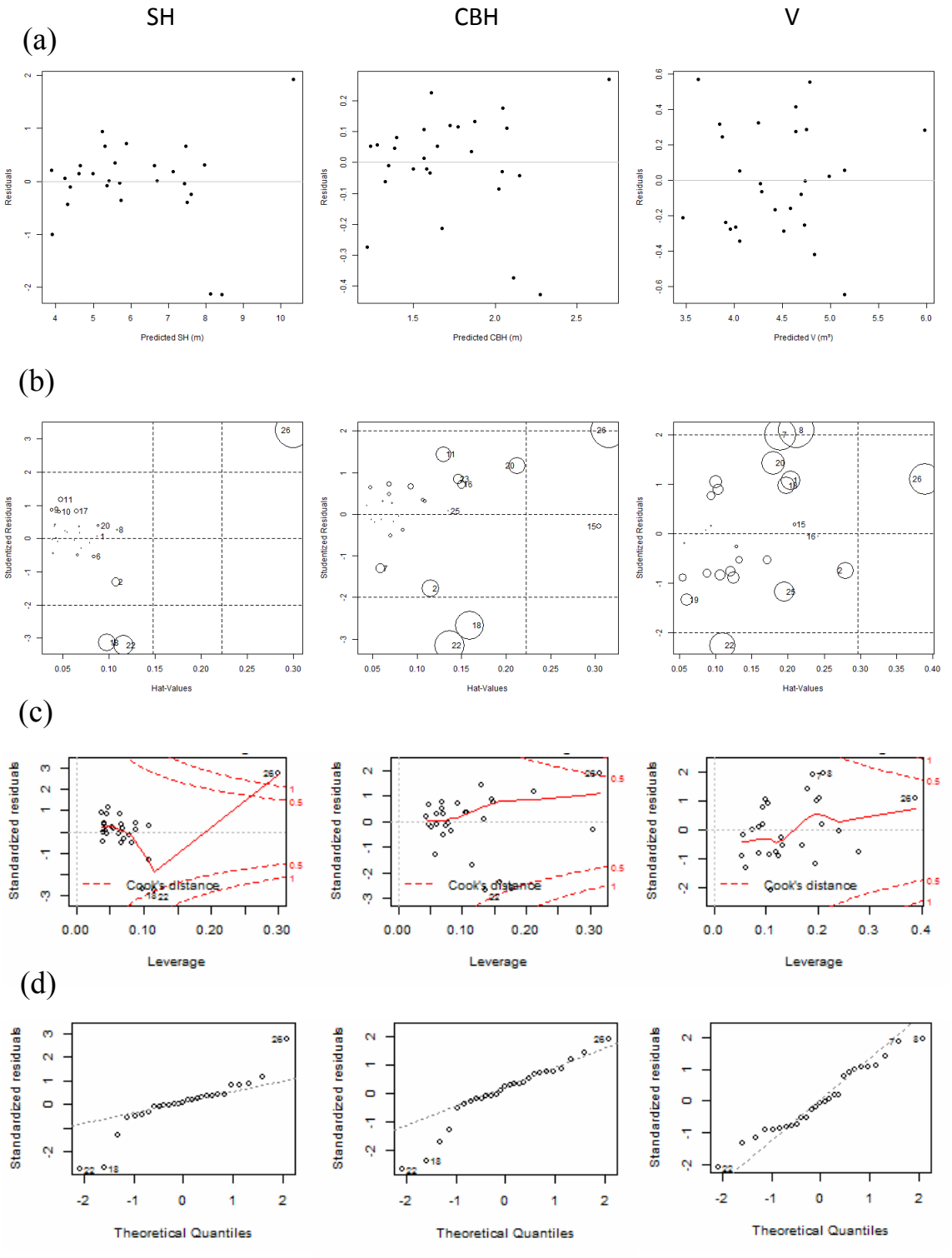

Fig. 2.6. Diagnostics plots for the selected models for the three studied canopy variables: average Stand Height (SH); average Crown Base Height (CBH); and the average crown Volume (V). (a) Residuals v. predicted values; (b) Regression Influence Plot; (c) Residuals v. Leverage; and (d) Normal Q-Q.

We tried also to use the Random Forest algorithm but the resulting percentages of variance explained were always very low, therefore we did not consider these models.

Maps of figures 2.7, 2.8 and 2.9 show the results of extending the fitted linear models to the whole study area. The three maps corresponding to the three studied variables (SH, CBH and crown volume) agreed in the identification of the areas with the highest values. The areas presenting the most pale colors or even white color (corresponding to no data) are also coinciding in the three maps.

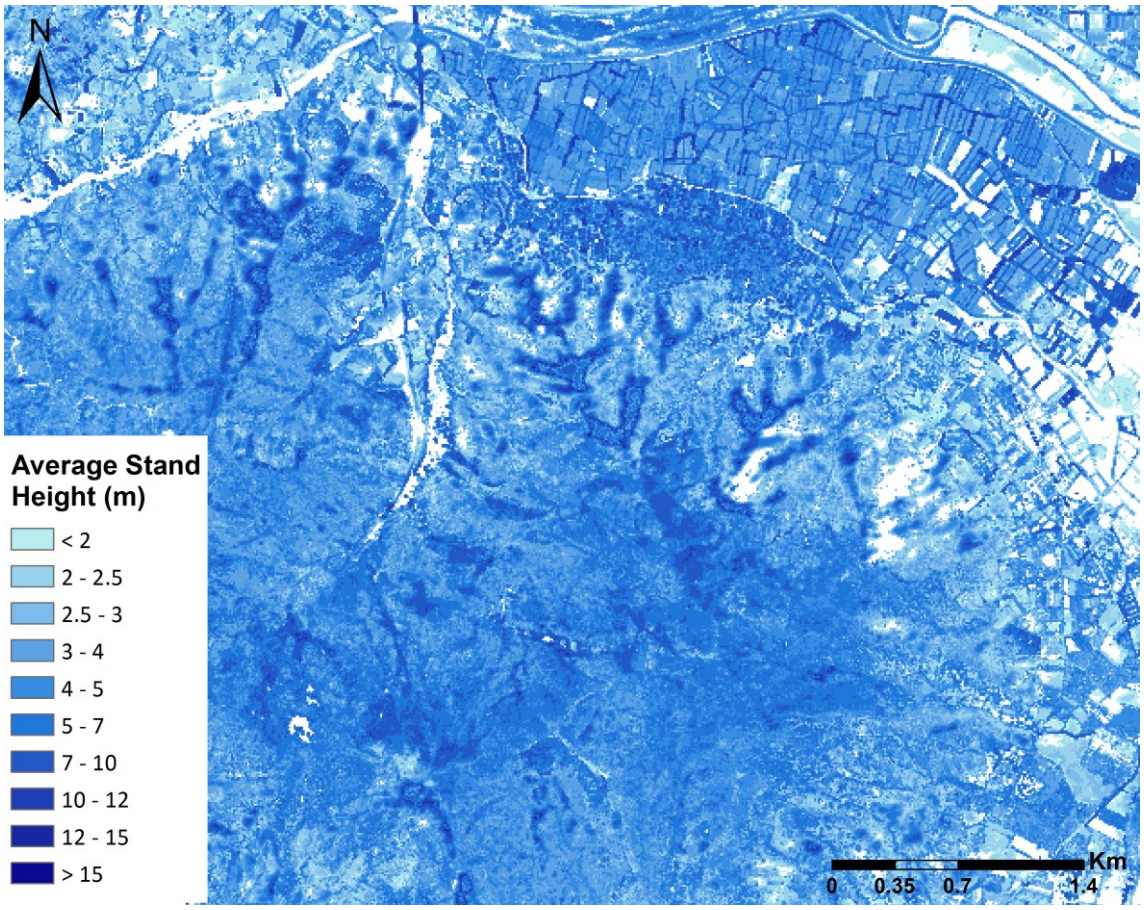

Fig. 2.7. Map of the resulting average Stand Height for the whole study area.

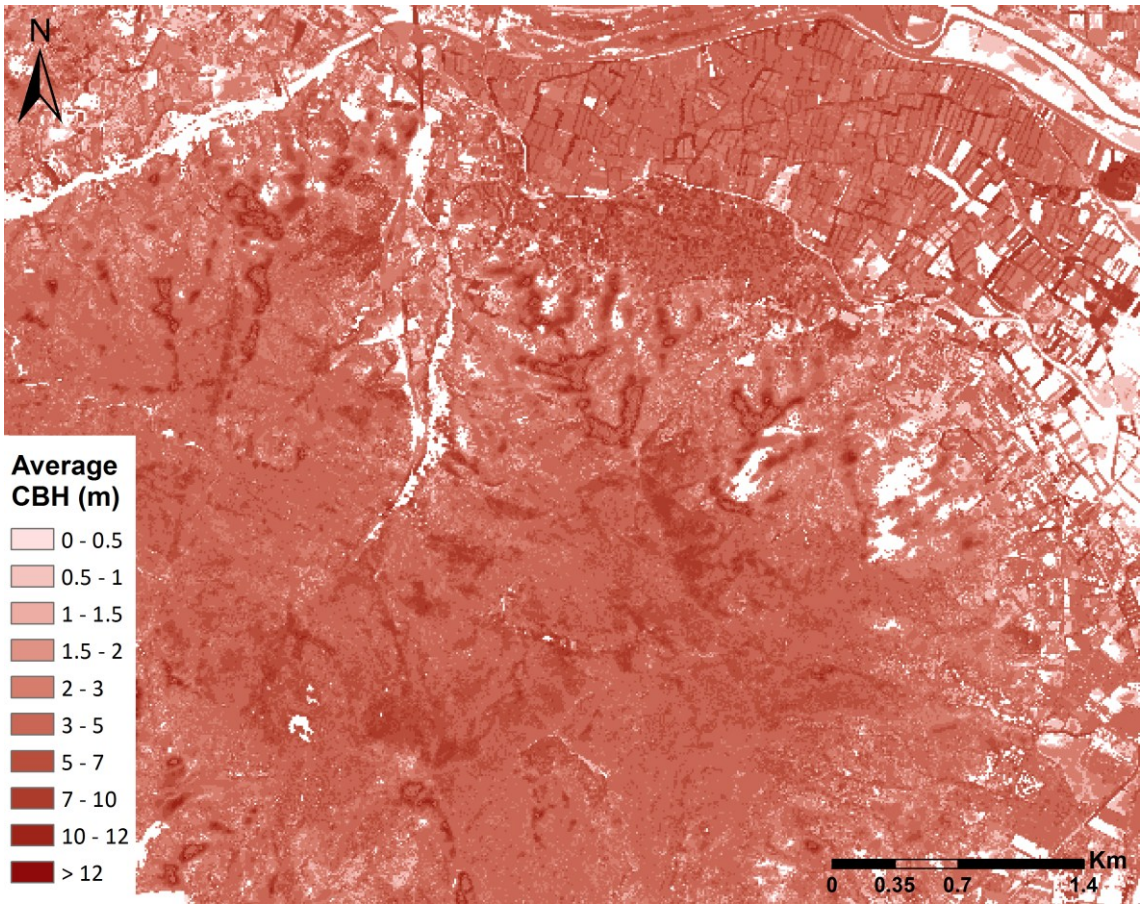

Fig. 2.8. Map of the resulting average Crown Base Height (CBH) for the whole study area.

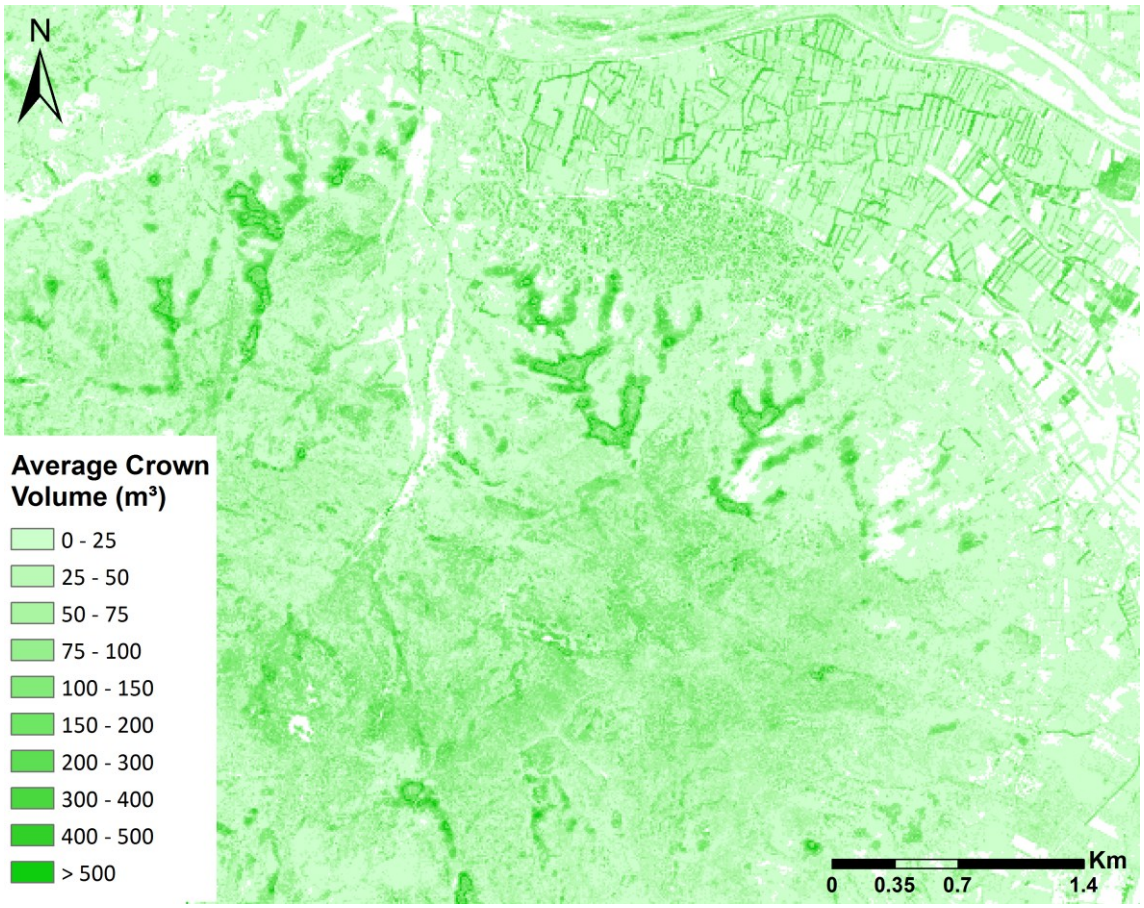

Fig. 2.9. Map of the resulting average Crown Volume for the whole study area.

4. DISCUSSION

According to the R-squared values the three models of canopy variables performed well at predicting the respective canopy variables. In addition, the plots of field data v. predicted data show the goodness of fit obtained with these models (Fig. 2.5).

In some of the diagnostics graphics a few plots appear away from the other ones. However, based on the results of the outlier test, there was no statistically justifiable reason to remove these plots from the data set (Aldas Manzano and Uriel Jiménez 2017). Moreover, we decided to perform a stratified sampling to cover all the variability; therefore, deleting a plot would have resulted in missing some parts of the forest fuels.

Regarding the stand height (SH), previous studies that used low-density LiDAR data found a linear relationship between this variable and LiDAR metrics as in our case (Hall *et al.* 2005; Brubaker *et al.* 2014; González-Ferreiro *et al.* 2014). The coefficients of determination in these studies were lower than our R^2 (0.7912) in of Hall et al. (2005, 0.571); similar in Brubaker et al. (2014, around 0.8); and higher in González-Ferreiro et al. (2014, 0.977).

With respect to CBH, there is not so much agreement among authors about its relationship with LiDAR metrics: Hall et al. (2005) found a logarithmic relationship, whereas González-Ferreiro et al. (2014) and the present study found a linear relationship. For this variable, the coefficient of determination obtained by González-Ferreiro et al. (2014) was higher than ours (R^2 = 0.8140) and those obtained by Hall et al. (2005) which were about 0.8 in both cases.

Our estimation of CBD was indirect; therefore it is difficult to compare our results with other studies carried out using different methodologies. However, we found a logarithmic relationship between LiDAR variables and crown volume as well as Andersen et al. (2005), which reported a logarithmic relationship between LiDAR variables and CBD.

Regarding the resulting maps, the dark areas in the three maps of SH, CBH and crown volume are likely the areas characterized by tree cover. These areas probably mainly correspond to forest areas, but it is also possible to observe the boundaries of the different agricultural areas, which are surrounded by trees. There are also some darker spots inside the village which probably correspond to urban green areas. Instead, the areas showing very pale colors or even white color (no data) most likely correspond to bare ground areas or areas with short vegetation. It is possible to note the white area in the north of the study area, the Flumendosa river, which flows from west to east through the study area. We can also observe the highway as light color area, from north towards south of the study area.

5. CONCLUSIONS

Lidar data are already being used in forest inventories to quantify and map several forest characteristics (Woods *et al.* 2010; Yu *et al.* 2011; Guerra-Hernández et al. 2016a; b, González-Ferreiro *et al.* 2017). Our results showed that this source of data could be also used to characterize the canopy of broadleaf forests.

Even if our accuracy results are lower than those obtained in other works, the results obtained are promising. Moreover, the most of previous works focused on the canopy variables of conifer forests (except for Brubaker et al. 2014), while our study investigated broadleaf forest: the more homogeneous crown shape of conifers could explain the differences in the accuracy values with respect to our work.

The selection of variables is quite time-consuming (explained in section "2.5. Variable selection") and at the end, for the linear models, we only selected five different variables. Since we wanted to develop a methodology which could be extended to more areas, this previous selection phase could be omitted, and we could directly test some commonly used variables (canopy cover, 95^{th} percentile of elevations, standard deviation…) or combinations of these variables. For example, the LiDAR strata variables seem not to be very useful for describing canopy variables, probably because the height of crowns is also variable and the use of only one range of heights (one stratum) is not enough. However, in case of the Random Forest model since all variables could be considered, strata metrics might be included in the analysis. However, in the present work this algorithm did not performed very well for describing canopy variables.

In any case, the most demanding part of this methodology is the field work. Therefore, if the selection of variables optimizes the results, it merits consideration. In this way, it would be possible to take maximum advantage of field-derived data.

The methodology presented in this study provides encouraging results. However, when a new LiDAR dataset will be available for Sardinia, it would be advisable to carry out the field work as soon as possible and this will probably improve the results. Further studies should also consider the possibility of improving the estimations of canopy bulk

density. In this sense, destructive inventories could help since direct measurements should provide better results.

Acknowledgments

LiDAR data were provided by the Autonomous Region of Sardinia. We are also thankful to the Agenzia Fo.Re.S.T.A.S. Especially we would thank Dr. Michele Puxeddu, Walter Bertolutti and the staff of the forest "Baccu Arrodas" for the permission and for the help in the localization of the plots. We would like to acknowledge Dr. Bachisio Arca and Dr. Andrea Ventura and the National Research Council (CNR), Institute of Biometeorology (IBIMET) for helping in the logistics and execution of sampling work, and for providing the materials to carry out the forest inventories. We also thank Liliana del Giudice and José Costa for their invaluable help in the field work.

References

Aldas Manzano J, Uriel Jiménez E (2017) 'Análisis multivariante aplicado con R.' (Alfa Centauro, Ediciones Paraninfo: Madrid)

Andersen H-E, Mcgaughey RJ, Reutebuch SE (2005) Estimating forest canopy fuel parameters using LIDAR data. *Remote Sensing of Environment* **94**, 441–449. doi:10.1016/j.rse.2004.10.013.

Arroyo LA, Pascual C, Manzanera A (2008) Fire models and methods to map fuel types: The role of remote sensing. *Forest Ecology and Management* **256**, 1239–1252. doi:10.1016/j.foreco.2008.06.048.

Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* **67**, 1–48. doi:10.18637/jss.v067.i01.

Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, White JS (2008) Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution* **24**, 127–135. doi:10.1016/j.tree.2008.10.008.

Brubaker KM, Johnson SE, Brinks J, Leites LP, Colleges WS, St P (2014) Estimating Canopy Height of Deciduous Forests at a Regional Scale with Leaf-Off , Low Point Density LiDAR. *Canadian Journal of Remote Sensing* **40**, 123–134. doi:10.1080/07038992.2014.943392.

Burnham KP, Anderson DR (1998) 'Model Selection and Multimodel Inference. A Practical Information-Theoretic Approach.' (Springer: Fort Collins, CO)

Claggett PR, Okay JA, Stehman S V (2010) Monitoring regional riparian forest cover change using stratified sampling and multiresolution imagery. *Journal of the American Water Resources Associaton* **46**, 334–343.

Condés S, Fernandez-Landa A, Rodriguez F (2013) Influencia del inventario de campo en el error de muestreo obtenido en un inventario con tecnología Lidar. In: '6º Congreso Forestal Español. Montes: Servicios y desarrollo rural', Vitoria-Gasteiz. 1–13.

Erdody TL, Moskal LM (2010) Fusion of LiDAR and imagery for estimating forest canopy fuels. *Remote Sensing of Environment* **114**, 725–737. doi:10.1016/j.rse.2009.11.002.

Finney MA (1998) FARSITE : Fire Area Simulator — Model Development and Evaluation. USDA Forest Service, Rocky Mountain Research Station, RMRS-RP-4. (Ogden, UT) http://www.firemodels.org/content/view/52/72/.

Finney MA (2006) An Overview of FlamMap Fire Modeling Capabilities. In: Andrews PL, Butler BW (eds) 'Fuel Management-How to Measure Success: Conference Proceedings. 28-30 March', Portland, OR. 213–220. (USDA Forest Service, Rocky Mountain Research Station, RMRS-P-41)

Frazer GW, Magnussen S, Wulder MA, Niemann KO (2011) Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sensing of Environment* **115**, 636–649. doi:10.1016/j.rse.2010.10.008.

Gobakken T, Næsset E (2009) Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. **39**, 1036–1052. doi:10.1139/X09-025.

González-Ferreiro E, Arellano-Pérez S, Castedo-Dorado F, Hevia A, Vega JA, Vega-Nieva D, Álvarez-Gonzàlez JG, Ruiz-González AD (2017) Modelling the vertical distribution of canopy fuel load using national forest inventory and low-density airbone laser scanning data. *PLoS ONE* **12**, 1–21. doi:e0176114.

González-Ferreiro E, Diéguez-Aranda U, Crecente-Campo F, Barreiro-Fernández L, Miranda D, Castedo-dorado F (2014) Modelling canopy fuel variables for Pinus radiata D. Don in NW Spain with low-density LiDAR data. **23**, 350–362. doi:10.1071/WF13054.

Guerra-Hernández J, Bastos Görgens E, García-Gutiérrez J, Estraviz Rodriguez LC, Tomé M, González-Ferreiro E (2016a) Comparison of ALS based models for estimating aboveground biomass in three types of Mediterranean forest. *European Journal of Remote Sensing* **49**, 185–204. doi:10.5721/EuJRS20164911.

Guerra-Hernández J, Tomé M, González-Ferreiro E (2016b) Using low density LiDAR data to map Mediterranean forest characteristics by means of an area- based approach and height threshold analysis. *Revista de Teledetección* **46**, 103–117. http://dx.doi.org/10.4995/raet.2016.3980.

Hall SA, Burke IC, Box DO, Kaufmann MR, Stoker JM (2005) Estimating stand structure using discrete-return lidar : an example from low density , fire prone ponderosa pine forests. **208**, 189–209. doi:10.1016/j.foreco.2004.12.001.

Hansen EH, Gobakken T, Næsset E (2015) Effects of Pulse Density on Digital Terrain Models and Canopy Metrics Using Airborne Laser Scanning in a Tropical Rainforest. *Remote Sensing* **7**, 8453–8468. doi:10.3390/rs70708453.

Hawbaker T, Keuler NS, Lesak AA, Gobakken T, Contrucci K, Radeloff VC (2009) Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. *Journal of Geophysical Research* **114**, 1–11. doi:10.1029/2008JG000870.

Hermosilla T, Ruiz LA, Kazakova A, Coops NC, Moskal LM (2014) Estimation of forest structure and canopy fuel parameters from small-footprint full-waveform LiDAR data. *International Journal of Wildland Fire* **23**, 224–233. doi:10.1071/WF13086.

Hevia A, Álvarez-González JG, Ruiz-Fernández E, Prendes C, Ruiz-González AD, Majada J, González-Ferreiro E (2016) Modelling canopy fuel and forest stand variables and characterizing the influence of thinning in the stand structure using airborne LiDAR. *Revista de Teledetección* **45**, 41–55. http://dx.doi.org/10.4995/raet.2016.3979%0AModelling.

Jahdi R, Salis M, Darvishsefat AA, Mostafavi MA, Alcasena FJ, Etemad V, Lozano OM, Spano D (2015) Calibration of FARSITE simulator in northern Iranian forests. *Natural Hazards and Earth System Sciences* **15**, 443–459. doi:10.5194/nhess-15-443-2015.

Keane RE, Garner JL, Schmidt KM, Menakis JP, Finney MA (1998) Development of input data layers for the FARSITE fire growth model for the Selway-Bitterroot Wilderness Complex, USA. USDA Forest Service, Rocky Mountain Research Station, RMRS-GTR-3. (Fort Collins, CO)

Lozano OM, Salis M, Ager AA, Arca B, Alcasena FJ, Monteiro AT, Finney MA, Giudice L Del, Scoccimarro E, Spano D (2017) Assessing Climate Change Impacts on Wildfire Exposure in Mediterranean Areas. *Risk Analysis* **37**, 1799–2022. doi:10.1111/risa.12739.

Magnussen S, Boudewyn P (1998) Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research* **28**, 1016–1031.

Mallinis G, Mitsopoulos I, Stournara P, Patias P, Dimitrakopoulos A (2013) Canopy Fuel Load Mapping of Mediterranean Pine Sites Based on Individual Tree-Crown Delineation. *Remote Sensing* **5**, 6461–6480. doi:10.3390/rs5126461.

McGaughey RJ (2014) FUSION/LDV: Software for LIDAR Data Analysis and Visualization. 154.

Nakagawa S, Schielzeth H (2013) A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution* **4**, 133–142. doi:10.1111/j.2041-210x.2012.00261.x.

R Core Team (2016) R: A language and environment for statistical computing. http://www.r-project.org/.

Riaño D, Meier E, Allgöwer B, Chuvieco E, Ustin SL (2003) Modeling airborne laser scanning data for the spatial generation of critical forest parameters in fire behavior modeling. *Remote Sensing of Environment* **86**, 177–186. doi:10.1016/S0034-4257(03)00098-1.

Ruiz LA, Hermosilla T, Mauro F, Godino M (2014) Analysis of the Influence of Plot Size and LiDAR Density on Forest Structure Attribute Estimates. **5**, 936–951. doi:10.3390/f5050936.

Salis M, Del Giudice L, Arca B, Ager AA, Alcasena-urdiroz F, Lozano O, Bacciu V, Spano D (2018) Modeling the effects of different fuel treatment mosaics on wild fi re spread and behavior in a Mediterranean agro-pastoral area. *Journal of Environmental Management* **212**, 490–505. doi:10.1016/j.jenvman.2018.02.020.

Scott JH, Reinhardt ED (2002) Estimating canopy fuels in conifer forests. *Systems for Environmental Management* **62**, 45–50.

Tabacchi G, Di Cosmo L, Gasparini P, Morelli S (2011) 'Stima del volume e della fitomassa delle principali specie forestali italiene, Equazioni di previsione, tavole del volume e tavole della fitomassa arborea epigea.' (Consiglio per la ricerca e la sperimentazione in Agricoltura, Unità di Ricerca per il Monitoraggio e la Pianificazione Forestale: Trento)

Wasser L, Day R, Chasmer L, Taylor A (2013) Influence of Vegetation Structure on Lidar-derived Canopy Height and Fractional Cover in Forested Riparian Buffers During Leaf-Off and Leaf-On Conditions. *PLoS ONE* **8**, 1–13. doi:10.1371/journal.pone.0054776.

Woods M, Pitt D, Penner M, Lim K, Nesbitt D, Treitz P (2010) Operational Implementation of LiDAR Inventories in Boreal Ontario A dvanced F orest R esource I nventory T echnologies. *The Forestry Chronicle* **87**, 512–528.

Yu X, Hyyppä J, Vastaranta M, Holopainen M, Viitala R (2011) Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS Journal of Photogrammetry and Remote Sensing* **66**, 28–37. doi:10.1016/j.isprsjprs.2010.08.003.

Zeide B (1980) Plot Size Optomization. *Forests Science* **26**, 251–257.

Zhao K, Popescu S, Meng X, Pang Y, Agca M (2011) Remote Sensing of Environment Characterizing forest canopy structure with lidar composite metrics and machine learning. *Remote Sensing of Environment* **115**, 1978–1996. doi:10.1016/j.rse.2011.04.001.

Zolkos SG, Goetz SJ, Dubayah R (2013) A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sensing of Environment* **128**, 289–298. doi:10.1016/j.rse.2012.10.017.

Chapter 3: Integrating LiDAR and satellite data with wildfire spread modeling to enhance wildfire exposure analysis in Mediterranean areas

1. INTRODUCTION

Forest managers use wildfire spread models as a tool for assessing fire risk, for fire management decision-making, to manage suppression resources, and to plan fuel reduction treatments across landscapes (Salis *et al.* 2016b; Lozano *et al.* 2017; Sakellariou *et al.* 2017; Salis *et al.* 2018).

Most fire spread models are based on the Rothermel's model, which calculates surface fire rate of spread using as inputs fine scale information on weather, fuels and topography. Two of the most used are FARSITE (Finney 1998), which is based on the Huygen's principle of wave propagation, and FlamMap (Finney 2006) based on the Minimum Travel Time (MTT) algorithm for the fire growth (Finney 2002). In particular, FARSITE is a good simulator for single events and it has been calibrated and validated for different areas including Sardinia (Arca *et al.* 2007; Duguy *et al.* 2007; Salis 2008; Jahdi *et al.* 2015b; Salis *et al.* 2016a).

Fuel maps are one of the most important inputs for wildfire spread simulation. Vegetation layers and databases should quantify fuel information at a high level of detail or resolution for the simulator to work well (Keane *et al.* 2000). For this reason numerous studies have been focused on the development of high accuracy fuel maps (Lasaponara and Lanorte 2007a; b, Mallinis *et al.* 2008, 2014; Otukei and Blaschke 2010). Most of this research is based on remote sensed data which require much less effort than the traditional field methods. However, the elaboration of these accurate layers is always - to a greater or lesser degree – demanding. Therefore, it is important to assess how much could be improved the results of a wildfire spread simulation by using high accurate maps.

Previous studies used Lidar and satellite imagery to build the fuel inputs for wildfire spread models (González-Olabarria *et al.* 2012). Mutlu et al. (2008) modeled fictitious fires with FARSITE to assess the differences in modeling outputs using different fuel model maps. Recently, a study carried out using FlamMap, assesed the impact of error in LiDAR-derived canopy height and canopy base height on modeled wildfire behavior (Kelly *et al.* 2018).

However, to the best of our knowledge, this is the first work that evaluates the improvement of historical wildfire simulations after the use of high accuracy maps. This approach has the advantage that the final area burned is known.

The main objective of this chapter will be to assess differences on the accuracy of FARSITE simulations by using alternatively the fuel model maps obtained from the previous chapters and fuel maps derived from the Land Use Map (LUM) of Sardinia (Autonomous Region of Sardinia 2008) as fuel layer inputs.

2. METHODS

2.1. Study area

The historical fire simulated in this chapter occurred in the area of Muravera, which was described in chapter 2.

2.2. Case study

We selected the only wildfire that affected broadleaf forests inside of the area with available LiDAR data. This wildfire occurred in the study area of Muravera on July 24th 2010 and affected 493 ha (Fig. 3.1).

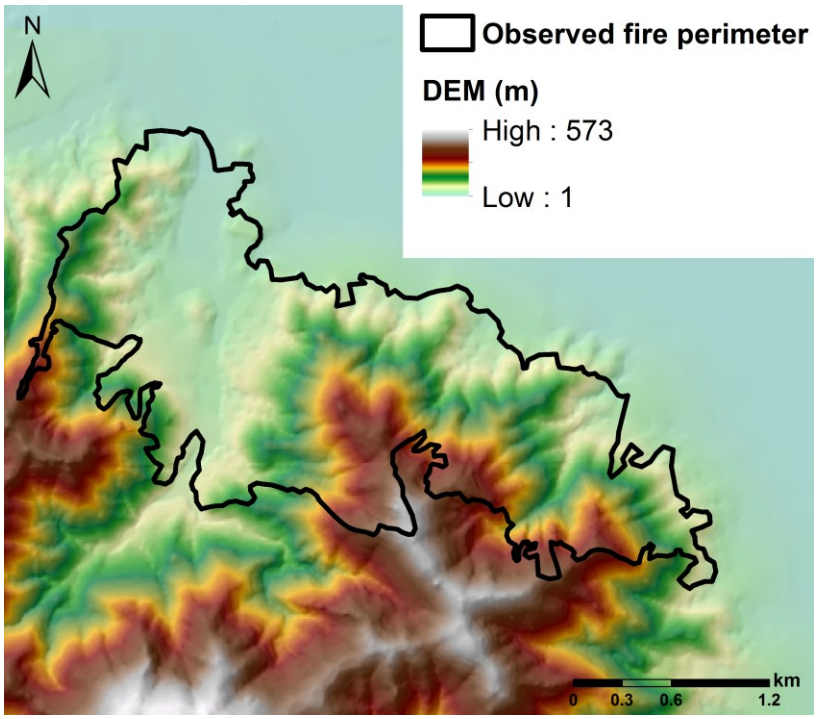

Fig. 3.1. Fire perimeter in the study area of Muravera. DEM = Digital Elevation Model.

Burned area perimeter was downloaded from www.sardegnageoportale.it whereas the ignition point and other information about the fire event were provided by the Forest Service of Sardinia. The fire started at 14:15 pm nearby a highway in a shrubland area mainly covered by *Cistus* ssp. The fire lasted for four hours and showed very high spread rates. Even if the dense smoke complicated the visibility, both spot fires and crown fires were observed and reported in the Sardinia Forest Service documents.

The weather of the day of the fire was characterized by a maximum temperature of 31°C, average relative humidity of 47% and strong northwest winds (24 $km \cdot h^{-1}$ with gusts of 50 $km \cdot h^{-1}$). The fire spread rapidly driven by the mistral wind and the topographic conditions, and was also supported by the dryness of fuels.

2.3. Input data for the simulations

A good approach to test the accuracy of the fuel map will be to use it as input for simulating some fire events in the past from which fire area and behavior are already known. Since other factors (as the other inputs or the model itself) could affect the accuracy of a simulation, the best way to assess the effects of a single input on the

output accuracy is to run the simulation changing only this input (fuel model map and/or canopy layers in our case), while holding constant all the other parameters and inputs. We prepared the simulation inputs as follows:

Topography and fuels

FARSITE required some spatial data, which should be imported as a so-called landscape file. The landscape file is composed of five layers (elevation, slope, aspect, fuel models and canopy cover) and other three optional canopy layers (canopy bulk density, canopy base height, and stand height).

We prepared five different landscape files (at a resolution of 10m); all of them with fixed topography layers. From the Digital Terrain Model (www.sardegnageoportale.it) we calculated, using ArcGIS, the layers of elevation, slope and aspect.

Regarding the fuel layers, each landscape file was different and the combinations of fuel layers and canopy layers of each landscape were as follows:

- LCP-LUM-NO: the fuel model layer was produced from a reclassification of the Land Use Map (LUM, Fig. 3.2) of Sardinia as explained in Chapter 1 (2.3. Fuel types). We built this landscape without canopy layers.

- LCP-LUM-ST: the fuel model layer was the reclassification of the LUM, but in this case, the canopy layers were built using standard values (see Table 3.1) used in other studies carried out in Sardinia in areas covered by similar vegetation types (Salis *et al.* 2016a)

- LCP-CUS-NO: we used the fuel model map (Fig. 3.2) obtained from the classification of selected variables using the random forest algorithm presented in Chapter 1 (which showed the best accuracy coefficients). This landscape file did not include canopy layers.

- LCP-CUS-ST: this landscape was built using our fuel map, but we added the standard canopy values for crown fuels.

- LCP-CUS-CUS: in this landscape file, we combined our fuel model map with the canopy layers resulting from Chapter 2.

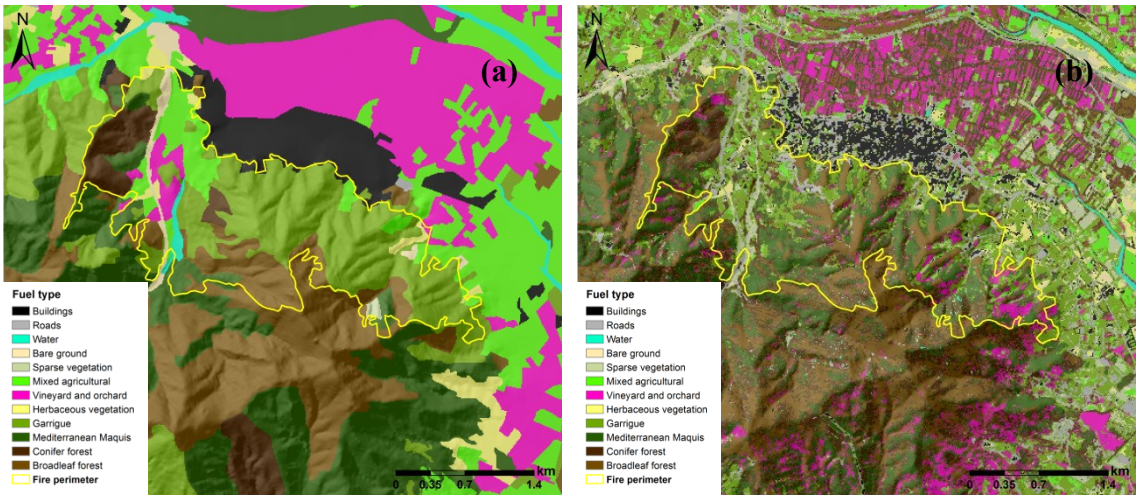

Fig. 3.2. Maps of fuel types from reclassification of the Land Use Map (LUM) of Sardinia (a) and from the classification carried out in chapter 1 combining multispectral data and LiDAR data (b) with the observed perimeter of the fire

The standard (Anderson 1982; Scott and Burgan 2005) or custom fuel models attributed to each fuel type are shown in Table 3.1. The custom fuel models were developed in previous studies carried out for Sardinian fuel models (Arca *et al.* 2007, 2009; Bacciu *et al.* 2009; Salis *et al.* 2013) and were applied to the shrubland vegetation (Mediterranean maquis and garrigue) and pastures.

Table 3.1. Fuel models assigned to each fuel type and the values of canopy variables used for the landscape files including standard canopy layers. [(a)] Scott and Burgan 2005; [(b)] Anderson 1982; and [(c)] Arca et al. 2009.

Fuel types	Class	Reference Fuel Models used	Standard canopy layer values		
			Stand Height (m)	Canopy Base Height (m)	Canopy Bulk Density ($kg \cdot m^{-3}$)
21	Buildings	NB1[(a)]	0	0	0
22	Roads	NB1[(a)]	0	0	0
23	Water	NB8[(a)]	0	0	0
24	Bare ground	GR1[(a)]	0	0	0
25	Sparse vegetation	Mod 1 (reduced load 50%)[(b)]	0	0	0
26	Mixed agricultural	Mod 1[(b)]	0	0	0
27	Vineyard and orchard	Mod 2[(b)]	6	0.8	0.09
28	Herbaceous vegetation	CM 27[(c)]	0	0	0
29	Garrigue	CM 29[(c)]	0	0	0
30	Mediterranean maquis	CM 28[(c)]	7	0.8	0.13
31	Conifer forests	TL6[(a)]	10	2	0.11
32	Broadleaf forests	TL3[(a)]	9	1.8	0.14

Weather data and fuel moisture content

Temperatures and wind data was gathered from San Vito, the nearest available weather station (www.weatherunderground.com), which is located 4 km far from the study area. The values of live and dead fuel moisture content (FMC) were set according to field data collected in other studies carried out in Sardinia, considering the fuel drought conditions of the fire period (Pellizzaro *et al.* 2007, 2009; Arca *et al.* 2009).

2.4. Wildfire simulations

To simulate the wildfire of Muravera we used FARSITE (Finney 1998). FARSITE is a spatially explicit fire growth model that simulates the spread and behavior of fires under heterogeneous conditions (Stratton 2006). This software incorporates models for surface fire (Rothermel 1972), crown fire initiation (Van Wagner 1977, 1993), crown fire spread (Rothermel 1991), spotting (Albini 1979), point-source fire acceleration (Finney 1998), and dead fuel moisture (Nelson 2000).

Different wildfire simulations were performed using alternately as input the five landscape files previously built (see 2.3 Input data for the simulations) and the rest of the inputs and parameters were held fixed. All wildfire spread simulations were performed at 10m resolution with a fire duration of four hours (from 14:15 pm to 18:15 pm). No suppression efforts were considered in fire modeling due to the lack of accurate information.

Regarding the outputs, we exported as shapefiles the perimeters of the burned area at different times. Instead, the fire behavior parameters were exported as raster files (Reaction Intensity, Rate of Spread, Fireline Intensity, Crown Fire Activity, Heat per unit Area, Spread Direction, Time of Arrival, and Flame Length). These fire behavior parameters are explained in the following paragraphs.

The **Reaction Intensity (RCI)** quantifies the rate of released energy per unit area of fire front.

Rate of Spread (ROS), according to the Rothermel (1972) model, can be defined as the ratio between the heat received by unburned fuel and the heat required to ignite unburned fuel. The equation is as follows:

$$ROS = \frac{RCI \times \xi \times (1 + \Phi_W + \Phi_S)}{\rho_b \times \varepsilon \times Q_{ig}}$$

where ROS is the rate of spread, RCI is the reaction intensity, ξ is the propagating flux ratio, Φ_w is the wind factor, Φ_s is the slope factor, ρ_b is the fuel bed bulk density, ε is the effective heating number, and Q_{ig} is the heat of pre-ignition. Under steady-state

conditions, the Rothermel's equation computes the rate of fire spread in the direction of maximum fire spread and assuming that wind and slope are aligned in this direction.

Fireline Intensity (FLI) is defined from fire rate of spread and fuel consumption, according to the Byram's (1959) equation:

$$FLI = H \times wa \times ROS$$

where H is the net low heat of combustion, wa is the fuel consumed in the active flaming front and ROS is the Rate of Spread.

Flame Length (FML) is defined as the distance from the base of the flaming zone to the top of continuous flames. This parameter is used to describe fire intensity and difficulties for the suppression operations.

As **Crown Fire Activity (CFR)** output, FARSITE provides a raster file with three possible values: (1) "Surface" which means no crown fire; (2) "Passive" which means torching crown fires (the rate of spread remains the same as surface fire but the fireline intensity and flame length increase a small amount as a result of the additional fuel consumed by the torching trees); and (3) "Active", which means active crown fires: in this case, the crown fire ROS is much higher than the surface fire ROS. The fireline intensity and flame length also increase significantly for active crown fire.

The **Heat per unit Area (HPA)** is the product of the total heat released by fuel and the net fuel load.

Finally, FARSITE produces other raster outputs with the information on the **Spread Direction (SDR)** for each pixel, and the **Time of Arrival (TOA)** of the simulated fire.

2.5. Statistical analysis

First step for evaluating the simulation performances with different landscape files was the calculation of the error matrix between actual and simulated fire perimeters. Then Sørensen's coefficient (SC; Legendre and Legendre 1998) and Cohen's Kappa coefficient (K; Congalton 1991) were used as measures of the accuracy of the extent of the fire spread (Arca *et al.* 2007; Jahdi *et al.* 2015a; b).

Sørensen's coeffcient (SC) is an indicator of the coincidence between observed and simulated burned areas and is calculated as follows:

$$SC = \frac{2a}{2a+b+c}$$

where *a* is the number of burned cells in both observed and simulated data, *b* is the number of burned cells in the simulation and unburned the observation, and *c* is the number of unburned cells in the simulation and burned in the observation (Arca *et al.* 2007).

Kappa coefficient (K) equation was explained in chapter 1 (2.7. Statistical methods).

Both K and SC coefficient values are close to one when the agreement between simulated and observed fire perimeters is very high.

Regarding the other outputs (not the perimeters), since it is not easy to know exactly how was the behavior of the fire (besides the perimeter that it burned), we only compared the different outputs among themselves. To simplify, we run a zonal statistics using ArcGIS, and we obtained a unique mean value inside the real perimeter.

3. RESULTS

3.1. Perimeters

The simulated perimeters were compared with the observed fire perimeter (Fig. 3.3, Table 3.2). Overall, the landscapes including canopy layers produced more accurate simulations according to both, Søerensen coefficient (0.75 – 0.86) and Cohen's kappa coefficient (0.69 – 0.83). All landscape files including our custom fuel models maps performed better than those based on the reclassification of the LUM. Only the LCP-CUS-NO showed low values for the accuracy coefficients (SC = 0.54 and K = 0.49), even if they were better than the LCP-LUM-NO (SC = 0.47 and K = 0.41).

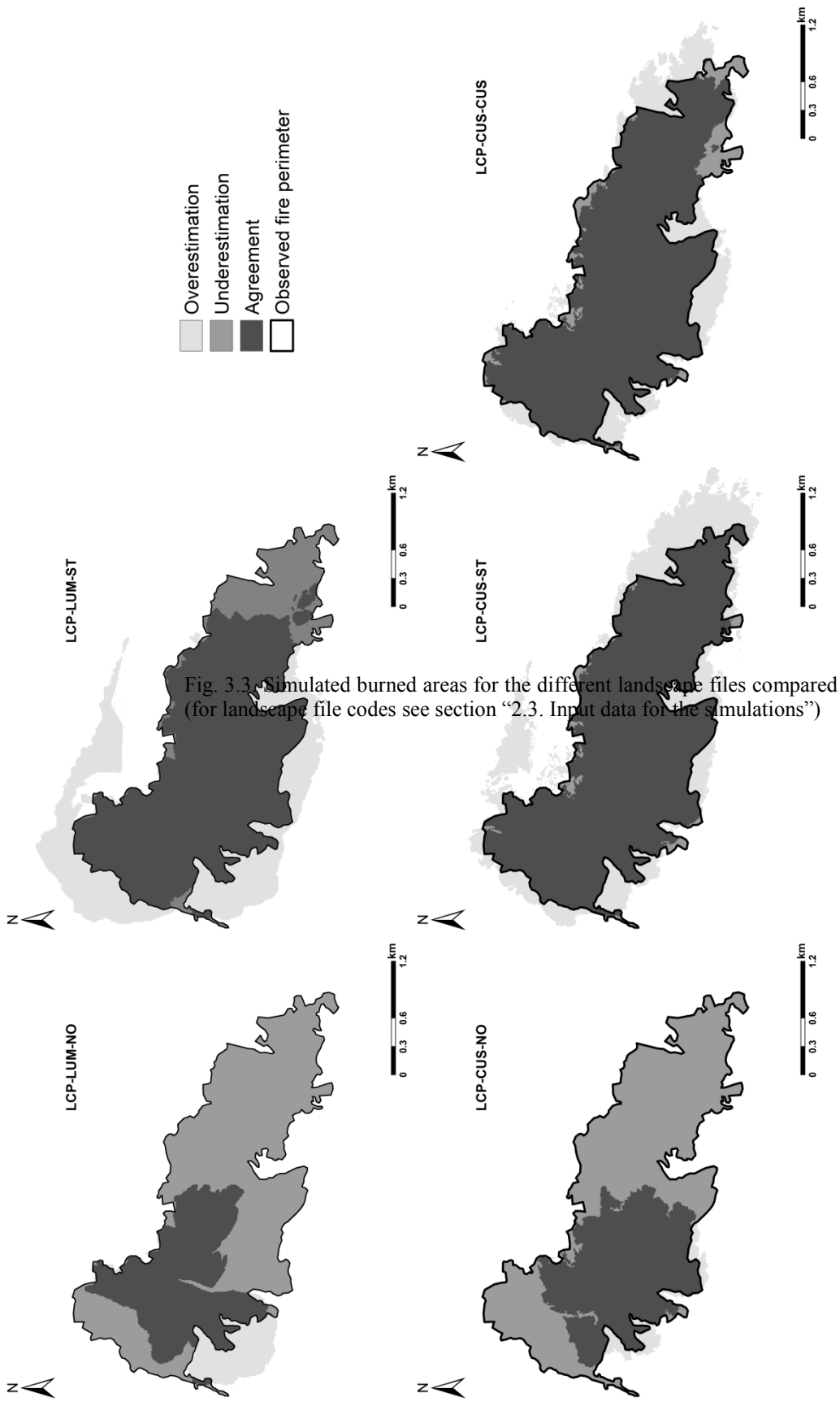

Fig. 3.3. Simulated burned areas for the different landscape files compared whith the actual fire perimeter (for landscape file codes see section "2.3. Input data for the simulations")

When comparing the LCP-LUM-ST with the LCP-CUS-ST, we observed that both accuracy coefficients were higher for the latter. There were also differences between the simulations carried out with LCP-CUS-CUS and LCP-CUS-CUS: our LiDAR-derived canopy layers guaranteed higher accuracy coefficients than those of the outputs from the landscape file including standard canopy layers.

Table 3.2. Results of the statistical evaluation from the comparison between simulated fire perimeter and actual perimeter. LCP = landscape file; ST = standard; CUS = custom (for landscape file codes see section "2.3. Input data for the simulations").

Landscape file	Agreement (ha)	Overestimation (ha)	Underestimation (ha)	Sørensen's coefficient	Cohen's Kappa coefficient
LUM-NO	166.28	46.80	326.67	0.47	0.41
LUM-ST	411.33	199.47	81.62	0.75	0.69
CUS-NO	190.87	19.71	302.08	0.54	0.49
CUS-ST	480.84	187.69	12.11	0.83	0.79
CUS-CUS	457.91	111.18	35.04	0.86	0.83

The outputs of the simulations carried out with the landscape files without canopy layers underestimated much more the area burned than the other ones. The highest overestimations were produced by the landscape files including the standard canopy layers. The landscape file including our custom fuel model map and our custom canopy layers showed the best agreement with respect to the observed perimeter.

3.2. Other outputs

The resulting maps of the raster outputs (Reaction Intensity, Rate of Spread, Fireline Intensity, Crown Fire Activity, Heat per unit Area, Spread Direction, Time of Arrival, and Flame Length) are compiled in Appendix E.

The simulations carried out with the landscape files without the canopy layers produced the lowest mean values for CFR and FML and the highest values of RCI (table 3.3). The LUM landscapes showed fire behavior outputs with the highest means in case of ROS and the lowest values for HPA and FLI.

Table 3.3. Statistics calculates inside the real perimeter of the fire for the outputs obtained from the simulations run with the different landscape file. LCP = landscape file; ST = standard; CUS = custom; CFR = Crown Fire Activity (1 = surface; 2 = passive; 3 = active); FLI = Fireline Intensity; FML = Flame Length; HPA = Heat per unit Area; RCI = Reaction Intensity; ROS = Rate of Spread; SDR = Spread Direction (for landscape file codes see section "2.3. Input data for the simulations").

Landscape file	Output	Minimum	Maximum	Range	Mean	Standard deviation
LUM-NO	CFR (category)	1	3	2	1.08	0.27
	FLI ($kW \cdot m^{-1}$)	0.01	9641.71	9641.70	530.26	730.42
	FML (m)	0.01	12.07	12.06	1.32	1.15
	HPA ($kJ \cdot m^{-2}$)	35.33	30209.08	30173.75	7633.48	7306.18
	RCI ($kW \cdot m^{-2}$)	4.89	1829.93	1825.04	877.78	530.81
	ROS ($m \cdot min^{-1}$)	0.01	35.61	35.60	4.50	3.99
	SDR (Azimuth)	0.00	359.00	359.00	155.34	100.66
LUM-ST	CFR (category)	1	3	2	1.76	0.43
	FLI ($kW \cdot m^{-1}$)	0.09	27255.83	27255.74	1031.98	1141.67
	FML (m)	0.01	24.13	24.12	2.34	1.73
	HPA ($kJ \cdot m^{-2}$)	57.30	116497.13	116439.83	9956.33	8324.49
	RCI ($kW \cdot m^{-2}$)	6.95	1847.49	1840.55	681.90	241.65
	ROS ($m \cdot min^{-1}$)	0.01	36.47	36.46	6.71	5.63
	SDR (Azimuth)	0.00	359.00	359.00	130.38	88.06
CUS-NO	CFR (category)	1	3	2	1.26	0.46
	FLI ($kW \cdot m^{-1}$)	0.72	14746.00	14745.28	1140.57	1546.60
	FML (m)	0.07	16.02	15.95	2.21	2.22
	HPA ($kJ \cdot m^{-2}$)	106.73	31505.61	31398.87	13499.12	10535.05
	RCI ($kW \cdot m^{-2}$)	0.00	1936.15	1936.15	1113.49	532.36
	ROS ($m \cdot min^{-1}$)	0.17	28.59	28.42	4.04	3.48
	SDR (Azimuth)	0.00	359.00	359.00	155.81	94.71
CUS-CUS	CFR (category)	1	3	2	1.47	0.50
	FLI ($kW \cdot m^{-1}$)	0.00	22042.79	22042.79	1310.91	2045.23
	FML (m)	0.00	20.94	20.94	2.51	2.64
	HPA ($kJ \cdot m^{-2}$)	4.95	445585.63	445580.67	15373.11	14417.93
	RCI ($kW \cdot m^{-2}$)	0.00	1915.06	1915.06	686.15	313.73
	ROS ($m \cdot min^{-1}$)	0.01	25.59	25.58	3.89	3.37
	SDR (Azimuth)	0.00	359.00	359.00	152.37	96.43
CUS-CUS	CFR (category)	1	3	2	1.65	0.48
	FLI ($kW \cdot m^{-1}$)	0.05	27925.45	27925.40	1324.19	2083.27
	FML (m)	0.00	24.52	24.52	2.53	2.64
	HPA ($kJ \cdot m^{-2}$)	0.80	633384.25	633383.45	15753.99	14737.69
	RCI ($kW \cdot m^{-2}$)	0.00	1886.17	1886.17	847.66	287.05
	ROS ($m \cdot min^{-1}$)	0.00	24.26	24.26	3.95	3.40
	SDR (Azimuth)	0.00	359.00	359.00	153.82	94.05

4. DISCUSSION

The uncertainty in wildfire spread models is mainly due to the variability of fuels (Lydersen *et al.* 2015). Therefore, when increasing the accuracy of the fuel maps and the canopy layers used as input for fire behavior simulation, an improvement in the outputs is expected. Our work confirmed this fact, and showed that the best results were obtained by the simulations carried out with high accuracy fuel maps and canopy layers.

The accuracy results of the simulations without considering the crown fuels were very weak, which supports the fact that crown fires and spot fires played a relevant role in the fire propagation, as indicated by the Sardinia Forest Service.

Using the fuel map from the reclassification of the LUM of Sardinia and including the standard values for the canopy layers, the results are acceptable. Arca et al. (2007) obtained similar results using as fuel model map a reclassification of the LUM of Sardinia.

When including our custom fuel model maps as input for the simulations, the accuracy of outputs improved considerably. Keeping constant all other inputs (including standard canopy layers), the accuracy coefficients increased from 0.75 to 0.83 for Sørensen coefficient and from 0.69 to 0.79 for kappa coefficient. The inclusion of our custom canopy layers also led to a better performance. However, in this case, the accuracy coefficients increased to a lesser extent (SC from 0.83 to 0.86 and K from 0.79 to 0.83). Probably, this small increase is due to the limited presence of forests in the area burned.

Regarding the overestimation values, they could seem very high. However, it is usual that FARSITE overestimates when suppression activities are not considered in the simulations (Jahdi *et al.* 2015a; b).

We also analyzed other outputs such as: Reaction Intensity, Rate of Spread, Fireline Intensity, Crown Fire Activity, Heat per unit Area, Spread Direction, Time of Arrival, and Flame Length. However, for these outputs it is more difficult to evaluate which simulations were more accurate since the real values were unknown. In any case, variables such as FML and CRW, which are mostly influenced by the canopy characteristics, showed lowest values when using landscape files without canopy layers.

5. CONCLUSIONS

In this study we assessed the effects of the accuracy of fuel models maps on the fire behavior model outputs. In particular, we used FARSITE, which is a tool that could assist forest managers and fire fighters with the mitigation of the effects of wildfire (Finney 1998). Accurate estimation of fire growth area and spread direction is extremely important because it could strongly support the decision-making process.

We demonstrated that increasing the accuracy of fuel model map and canopy layers improves the outputs of the wildfire spread models. The improvement obtained by including more accurate layers in the simulation could be considered moderate. However, especially in areas where houses and vegetation are intermingled, fatalthis effect could definitely be much more relevant.

Acknowledgments

We want to thank Forest Service of Sardinia in particular the Muravera Forest Service Station for providing all the necessary information about the fire under investigation.

References

Albini FA (1979) Spot fire distance from burning trees: a predictive model. USDA Forest Service, Intermountain Forest and Range Experiment Station, General Technical Report INT-GTR-56. (Ogden, UT)

Anderson HE (1982) Aids to determining fuel models for estimating fire behavior. USDA Forest Service, Intermountain Forest and Range Experiment Station, General Technical Report INT-122. (Ogden, UT)

Arca B, Bacciu V, Pellizzaro G, Salis M, Ventura A, Duce P, Brundu G (2009) Fuel model mapping by Ikonos imagery to support spatially explicit fire simulators. In: '7th international workshop on advances in remote sensing and GIS applications in forest fire management towards an operational use of remote sensing in forest fire management', Matera. 4.

Arca B, Duce P, Pellizzaro G, Bacciu V, Salis M, Spano D (2007) Evaluation of Farsite Simulator in Mediterranean maquis. *International Journal of Wildland Fire* **16**, 563–572. doi:10.1071/WF06070.

Autonomous Region of Sardinia (2008) Carta dell'Uso del Suolo in scala 1:25.000. http://www.sardegnageoportale.it/index.php?xsl=1598&s=291548&v=2&c=8831&t=1.

Bacciu V, Salis M, Spano D, Arca B, Pellizzaro G, Duce P (2009) Assessment of smoke emission and carbon estimate from Mediterranean maquis fire events. In: 'Eighth Symposium on Fire and Forest Meteorology.', 6. (American Meteorological Society)

Byram GM (1959) Combustian of Fuels. In: Davis KP (ed): 'For. Fire Control Use'. pp. 61–89. (McGraw-Hill Book Company: New York, Toronto, London)

Congalton RG (1991) A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* **37**, 35–46. doi:10.1016/0034-4257(91)90048-B.

Duguy B, Alloza JA, Röder A, Vallejo VR, Pastor F (2007) Evaluating the Effects of Lanscape-Scale Fuel Treatments on Fire Growth and Behaviour in a Mediterranean Landscape (Eastern Spain). *Gestió Del Risc D'Incendi, Ecologia Del Foc I Restauració De Zones Cremades* 619–632.

Finney MA (1998) FARSITE : Fire Area Simulator — Model Development and Evaluation. USDA Forest Service, Rocky Mountain Research Station, RMRS-RP-4. (Ogden, UT) http://www.firemodels.org/content/view/52/72/.

Finney MA (2002) Fire growth using minimum travel time methods. *Canadian Journal of Forest Research* **32**, 1420–1424. doi:10.1139/x02-068.

Finney MA (2006) An Overview of FlamMap Fire Modeling Capabilities. In: Andrews PL, Butler BW (eds) 'F uel Management-How to Measure Success: Conference Proceedings. 28-30 March', Portland, OR. 213–220. (USDA Forest Service, Rocky Mountain Research Station, RMRS-P-41)

González-Olabarria J-R, Rodríguez F, Fernández-Landa A, Mola-Yudego B (2012) Mapping fire risk in the Model Forest of Urbión (Spain) based on airborne LiDAR measurements. *Forest Ecology and Management* **282**, 149–156. doi:10.1016/j.foreco.2012.06.056.

Jahdi R, Salis M, Darvishsefat AA, Alcasena FJ, Mostafavi MA, Etemad V, Lozano OM, Spano D (2015) Evaluating fire modelling systems in recent wildfires of the Golestan National Park, Iran. *Forestry: An International Journal of Forest Research* **0**, 1–14. doi:10.1093/forestry/cpv045.

Jahdi R, Salis M, Darvishsefat AA, Mostafavi MA, Alcasena FJ, Etemad V, Lozano OM, Spano D (2015) Calibration of FARSITE simulator in northern Iranian forests. *Natural Hazards and Earth System Sciences* **15**, 443–459. doi:10.5194/nhess-15-443-2015.

Keane RE, Mincemoyer SA, Schmidt KM, Long DG, Garner JL (2000) Mapping vegetation and fuels for fire management on the Gila National Forest Complex , New Mexico. Forest Service, Rocky Mountain Research Station, Gen. Tech. Rep. RMRS-GTR-46-CD. (Ogden, UT)

Kelly M, Su Y, Di Tommaso S, Fry DL, Collins BM, Stephens SL, Guo Q (2018) Impact of Error in Lidar-Derived Canopy Height and Canopy Base Height on Modeled Wildfire Behavior in the Sierra Nevada , California , USA. *Remote Sensing* **10**, 1–18. doi:10.3390/rs10010010.

Lasaponara R, Lanorte A (2007a) Remotely sensed characterization of forest fuel types by using satellite ASTER data. *International Journal of Applied Earth Observation and Geoinformation* **9**, 225–234. doi:10.1016/j.jag.2006.08.001.

Lasaponara R, Lanorte A (2007b) On the capability of satellite VHR QuickBird data for fuel type characterization in fragmented landscape. *Ecological Modelling* **204**, 79–84. doi:10.1016/j.ecolmodel.2006.12.022.

Legendre P, Legendre L (1998) 'Numerical Ecology.' (Elsevier: Amsterdam)

Lozano OM, Salis M, Ager AA, Arca B, Alcasena FJ, Monteiro AT, Finney MA, Giudice L Del, Scoccimarro E, Spano D (2017) Assessing Climate Change Impacts on Wildfire Exposure in Mediterranean Areas. *Risk Analysis* **37**, 1799–2022. doi:10.1111/risa.12739.

Lydersen JM, Collins BM, Knapp EE, Roller GB, Stephens S (2015) Relating fuel loads to overstorey structure and composition in a fire-excluded Sierra Nevada mixed conifer forest. *International Journal of Wildland Fire* **24**, 484–494. http://dx.doi.org/10.1071/WF13066.

Mallinis G, Galidaki G, Gitas I (2014) A Comparative Analysis of EO-1 Hyperion, Quickbird and Landsat TM Imagery for Fuel Type Mapping of a Typical Mediterranean Landscape. *Remote Sensing* **6**, 1684–1704. doi:10.3390/rs6021684.

Mallinis G, Mitsopoulos ID, Dimitrakopoulos AP, Gitas IZ, Karteris M (2008) Local-Scale Fuel-Type Mapping and Fire Behavior Prediction by Employing High-Resolution Satellite Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **1**, 230–239. doi:10.1109/JSTARS.2008.2011298.

Mutlu M, Popescu SC, Zhao K (2008) Sensitivity analysis of fire behavior modeling with LIDAR-derived surface fuel maps. *Forest Ecology and Management* **256**, 289–294. doi:10.1016/j.foreco.2008.04.014.

Nelson RM (2000) Prediction of diurnal change in 10-h fuel sticks moisture content. *Canadian Journal of Forest Research* **30**, 1071–1087.

Otukei JR, Blaschke T (2010) Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation* **12**, 27–31. doi:10.1016/j.jag.2009.11.002.

Pellizzaro G, Duce P, Ventura A, Zara P (2007) Seasonal variations of live moisture content and ignitability in shrubs of the mediterranean basin. *International Journal of Wildland Fire* **16**, 633–641. doi:10.1071/WF05088.

Pellizzaro G, Ventura A, Arca B, Arca A, Duce P (2009) Weather seasonality and temporal pattern of live and dead fuel moisture content in Mediterranean shrubland. *Geophysical Research Abstracts* **11**, 12100–12100.

Rothermel RC (1972) A Mathematical model for predicting fire spread in wildland fuels. USDA Forest Service, Intermountain Forest and Range Experiment Station, Research paper INT-115. (Ogden, UT)

Rothermel RC (1991) Predicting Behavior and Size of Crown Fires in the Northern Rocky Mountains. Forest Service, Intermountain Research Station, Research paper INT-438. (Ogden, UT)

Sakellariou S, Tampekis S, Samara F (2017) Review of state-of-the-art decision support systems (DSSs) for prevention and suppression of forest fires. *Journal of Forestry Research* **28**, 1107–1117. doi:10.1007/s11676-017-0452-1.

Salis M (2008) Fire behaviour simulation in Mediterranean maquis using FARSITE (fire area simulator). University of Sassari.

Salis M, Ager AA, Arca B, Finney MA, Bacciu V (2013) Assessing exposure to human and ecological values in Sardinia, Italy. *International Journal of Wildland Fire* **22**, 549–565. doi:10.1071/WF11060.

Salis M, Arca B, Alcasena F, Arianoutsou M, Bacciu V, Duce P, Duguy B, Koutsias N, Mallinis G, Mitsopoulos ID, Moreno JM, Pérez JR, Urbieta IR, Xystrakis F, Zavala G, Spano D (2016) Predicting wildfire spread and behaviour in Mediterranean landscapes. *International Journal of Wildland Fire* **25**, 1015–1032. http://dx.doi.org/10.1071/WF15081.

Salis M, Del Giudice L, Arca B, Ager AA, Alcasena-urdiroz F, Lozano O, Bacciu V, Spano D (2018) Modeling the effects of different fuel treatment mosaics on wildfire spread and behavior in a Mediterranean agro-pastoral area. *Journal of Environmental Management* **212**, 490–505. doi:10.1016/j.jenvman.2018.02.020.

Salis M, Laconi M, Ager AA, Alcasena FJ, Arca B, Lozano O, Fernandes de Oliveira A, Spano D (2016) Forest Ecology and Management Evaluating alternative fuel treatment strategies to reduce wildfire losses in a Mediterranean area. *Forest Ecology and Management* **368**, 207–221. doi:10.1016/j.foreco.2016.03.009.

Scott JH, Burgan RE (2005) Standard Fire Behavior Fuel Models: A Comprehensive Set for Use with Rothermel's Surface Fire Spread Model. USDA Forest Service, Rocky Mountain Research Station, General Technical Report RMRS-GTR-153. (Fort Collins (CO))

Stratton RD (2006) Guidance on Spatial Wildland Fire Analysis: Models, Tools, and Techniques. USDA Forest Service, Rocky Mountain Research Station, RMRS-GTR-183. (Fort Collins, CO)

Van Wagner CE (1977) Conditions for the start and spread of crown fire. *Canadian Journal of Forest Research* **7**, 23–34. doi:10.1139/x77-004.

Van Wagner CE (1993) Prediction of crown fire behavior in two stands of jack pine. *Canadian Journal of Forest Research* **23**, 442–449. https://doi.org/10.1139/x93-062.

Final conclusions

This work suggests that the use of LiDAR and satellite imagery data can contribute to improve estimates of modeled wildfire behavior. The methodology proposed in this work could become a semi-automatic process to assess the potential fire behavior from remote sensed data. However, the procedures for both, creating accurate fuel maps and developing canopy layers from high accurate remote sensed data, are very time-consuming. Therefore it would be advisable to determinate, previously, which level of detail is needed. In that way, for some works it would be enough with the reclassification of the Land Use Map which results are acceptable. For example, for fire spread modeling at no very high resolution, for large areas, this option would be adequate. Even if in some cases more accurate results are needed, the accuracy increases considerably when simply performing the classification. The development of high accurate canopy layers requires (in addition to the data process) field work and therefore it would be necessary only when very high accuracies are expected.

The key aspect of the methodology showed in this work is that it could be reproduced for other areas or when new data would be available (especially LiDAR data which is expensive and therefore it is difficult to have this kind of data at disposal). This is important because it would allow an efficient, operative and low-cost update of maps compared to traditional methods.

Another interesting aspect of this methodology is that the spatial distribution of fuel models and their canopy characteristics is not only useful for modelling wildfire behavior. Forest managers working on prevention might benefit from this information because they could better plan their actions. In addition, even fire fighter could need this information for predicting the fire behavior in real time.

In this work, we developed and validated a methodology to improve the spatial information about fuel models and canopy characteristics. However, there is another aspect of fuel models that is very important for wildfire simulations: the characterization of each fuel model (dead and live fuel load, moisture of extinction, surface area-to-

volume ratio, etc). Some studies have explored the possibility of quantifying the fuel load or biomass from LiDAR data (Guerra-Hernández *et al.* 2016; Chen *et al.* 2017; González-Ferreiro *et al.* 2017). Additionally, other studies investigated the relationship between fuel moisture and the reflectances from different satellite images (Nieto *et al.* 2010; Jurdao *et al.* 2012; Yebra *et al.* 2013). Therefore, maybe in the future, would be possible to gather and create all the necessary information for wildfire modeling from remote sensed data (a combination of LiDAR data and satellite images) following a rather automatic methodology. Thus, future studies can be oriented to improve this topic, particularly for fire-prone areas such as the Mediterranean basin.

At the end, this work, as well as all studies about wildfires, aims to maintain and protect our landscapes and natural heritage, to minimise the damage caused to properties, and especially, to avoid loss of human lives by uncontrolled forest fires. We hope that we have contributed, in some small way, to improve the prevention of these disaters.

REFERENCES

Chen Y, Zhu X, Yebra M, Harris S, Tapper N (2017) Development of a predictive model for estimating forest surface fuel load in Australian eucalypt forests with LiDAR data. *Environmental Modelling and Software* **97**, 61–71. doi:10.1016/j.envsoft.2017.07.007.

González-Ferreiro E, Arellano-Pérez S, Castedo-Dorado F, Hevia A, Vega JA, Vega-Nieva D, Álvarez-Gonzàlez JG, Ruiz-González AD (2017) Modelling the vertical distribution of canopy fuel load using national forest inventory and low-density airbone laser scanning data. *PLoS ONE* **12**, 1–21. doi:e0176114.

Guerra-Hernández J, Bastos Görgens E, García-Gutiérrez J, Estraviz Rodriguez LC, Tomé M, González-Ferreiro E (2016) Comparison of ALS based models for estimating aboveground biomass in three types of Mediterranean forest. *European Journal of Remote Sensing* **49**, 185–204. doi:10.5721/EuJRS20164911.

Jurdao S, Chuvieco E, Arevalillo JM (2012) Modelling fire ignition probability fromsatellite estimates of live fuel moisture content. *Fire Ecology* **8**, 77–97. doi:10.4996/fireecology.0801077.

Nieto H, Aguado I, Chuvieco E, Sandholt I (2010) Dead fuel moisture estimation with MSG-SEVIRI data. Retrieval of meteorological data for the calculation of the equilibrium moisture content. *Agricultural and Forest Meteorology* **150**, 861–870. doi:10.1016/j.agrformet.2010.02.007.

Yebra M, Dennison PE, Chuvieco E, Riaño D, Zylstra P, Hunt ER, Danson FM, Qi Y, Jurdao S (2013) A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products. *Remote Sensing of Environment* **136**, 455–468. doi:10.1016/j.rse.2013.05.029.

Acknowledgements

Firstly, I would like to express my gratitude to my advisor Prof.ssa Donatella Spano for allowing me to experience different working contexts and ways of making research in my stages abroad. Besides, she was ready to be helpful when I needed it.

I would also like to thank my supervisor Dr. Michele Salis for his support and help during my PhD activities. He was also always understanding with regard to my personal circumstances.

My sincere thanks also goes to the Scuola di Dottorato in Scienze Agrarie of the University of Sassari for giving me the opportunity to do this work and also for the possibility of knowing this wonderful island, its really friendly people and the Pecorino Sardo.

I would especially like to thank Prof. Emilio Chuvieco for his patience explaining me everything from the very beginning. He has an extraordinary combination of immense knowledge and a gift for teaching, which I admire. Also his team was always willing to help me at my research, but not only… thank you also for the really funny coffee brakes and lunches.

Prof. Margarida Tomé, Brigite Botequim and Juan Guerra help me also a lot in my research. I want to thank them for their advices which undoubtedly contributed to the accomplishment of this thesis. Also the rest of the ForChange team helped me a lot and I'm very grateful to them for involving me in their synergetic link-making initiatives such as the Friday seminars, the "*Café com Ciência*" and the *Cabaz*.

I would like to acknowledge Dr. Bachisio Arca and Dr. Andrea Ventura and the National Research Council (CNR), Institute of Biometeorology (IBIMET) for helping in the logistics and execution of sampling work, and for providing the materials to carry out the forest inventories. They were also always very encouraging even when technology did not want to help.

The reviewers of the thesis made very useful comments, which contributed to considerably improve the manuscript. I am grateful to them for spending their time reading the initial manuscript and especially for their really prompt responses.

I would like thank Riccardo Nicoló Zallu and Maria Teresa Spano, responsibles for PhD students of the Sassari University, for his help in all the administrative issues. Riccardo Zallu was really patient with me in this last period and he answered each and every one of my thousand questions.

My office mates Antonio Trabbuco, Gavriil Kyriakakis, Antonio Caddeo and Simone Mereu showed always interest in my thesis and want to thank them for their advices, but I'm especially grateful to them for sharing sweets and, in this last period, for pup up with the noise of my breast pump.

I thank the colleagues of the ex-dipNET and CMCC for sharing their knowledge, for the time we spent together attending seminars and congresses, and for keeping me good company at lunch time (especially Liliana and Carla who also shared with me their recipes and a lot more).

Out of the academic sphere, I want to thank all my friends who highly contributed to my emotional stability (and *corpore sano*) in this demanding period. Thank you very much to "Ties Maries", "Diaspora Sassarese", "Españoles en Sassari" and "Brujas".

I must also express my very deep gratitude to my family who has always encouraged me to do everything made me feel good. Especially I would thank my mother for always listening to my problems (even when she didn't understand what I was talking about) and say something which cheers me up. Also my father has been important to me for being always there and for giving me this feeling of calm.

Last but not the least, I am infinitely grateful to my full time partner Chema and my Son Matías. Regarding the thesis, Chema has contributed as labor for the field work, as mentor for some parts, and as spiritually support throughout the whole period (including the phases of "pregnancy + thesis" and "maternity + thesis"). I want to thank also my son because he had to share my attention with the thesis.

Appendices

Appendix A. Maps obtained from the classifications outputs carried out with the different algorithms and in the study area of Siniscola.

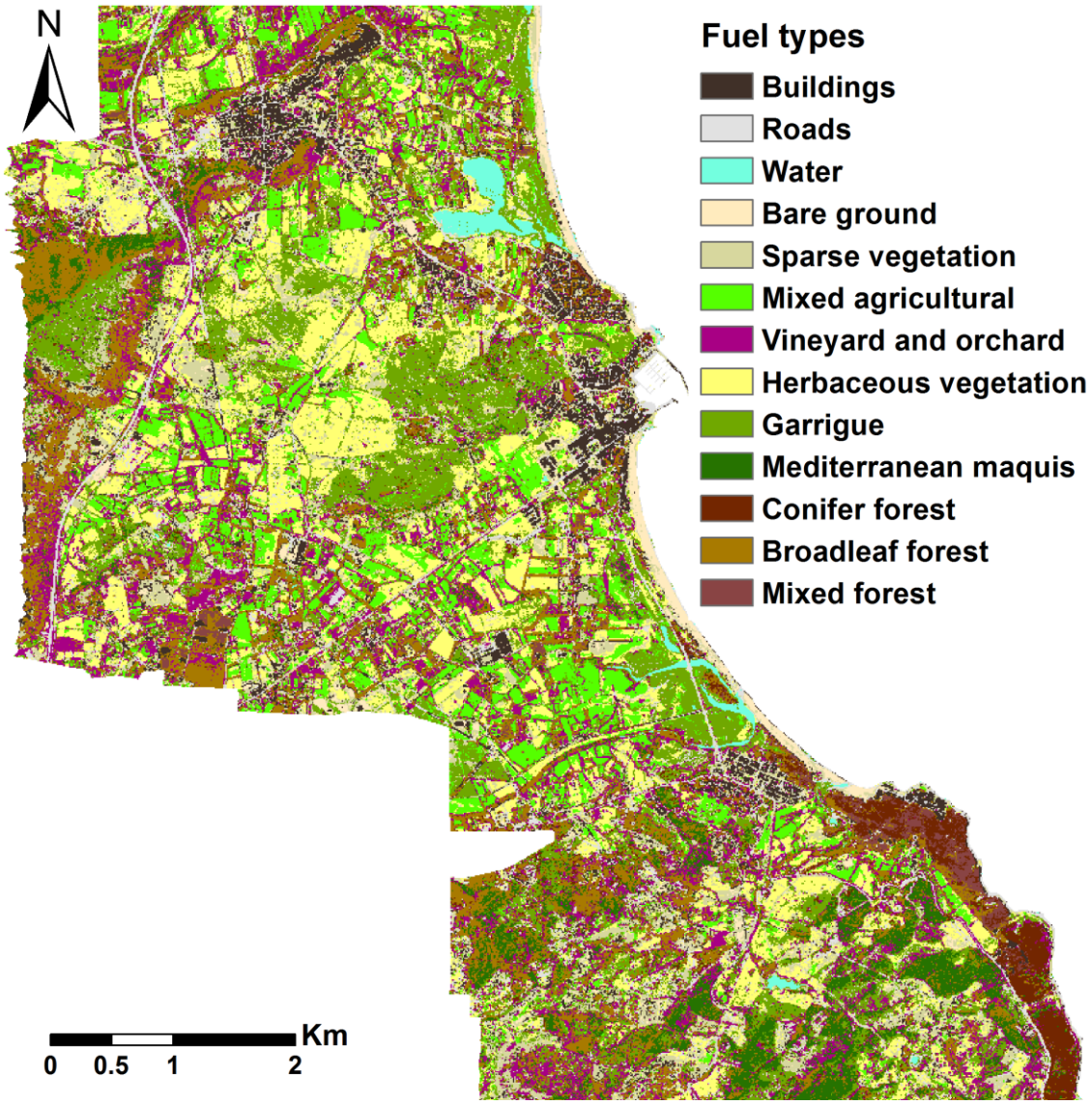

Fig. A.1. Classification output map for the study area of Siniscola using the Maximum Likelihood (ML) algorithm with the selection of variables.

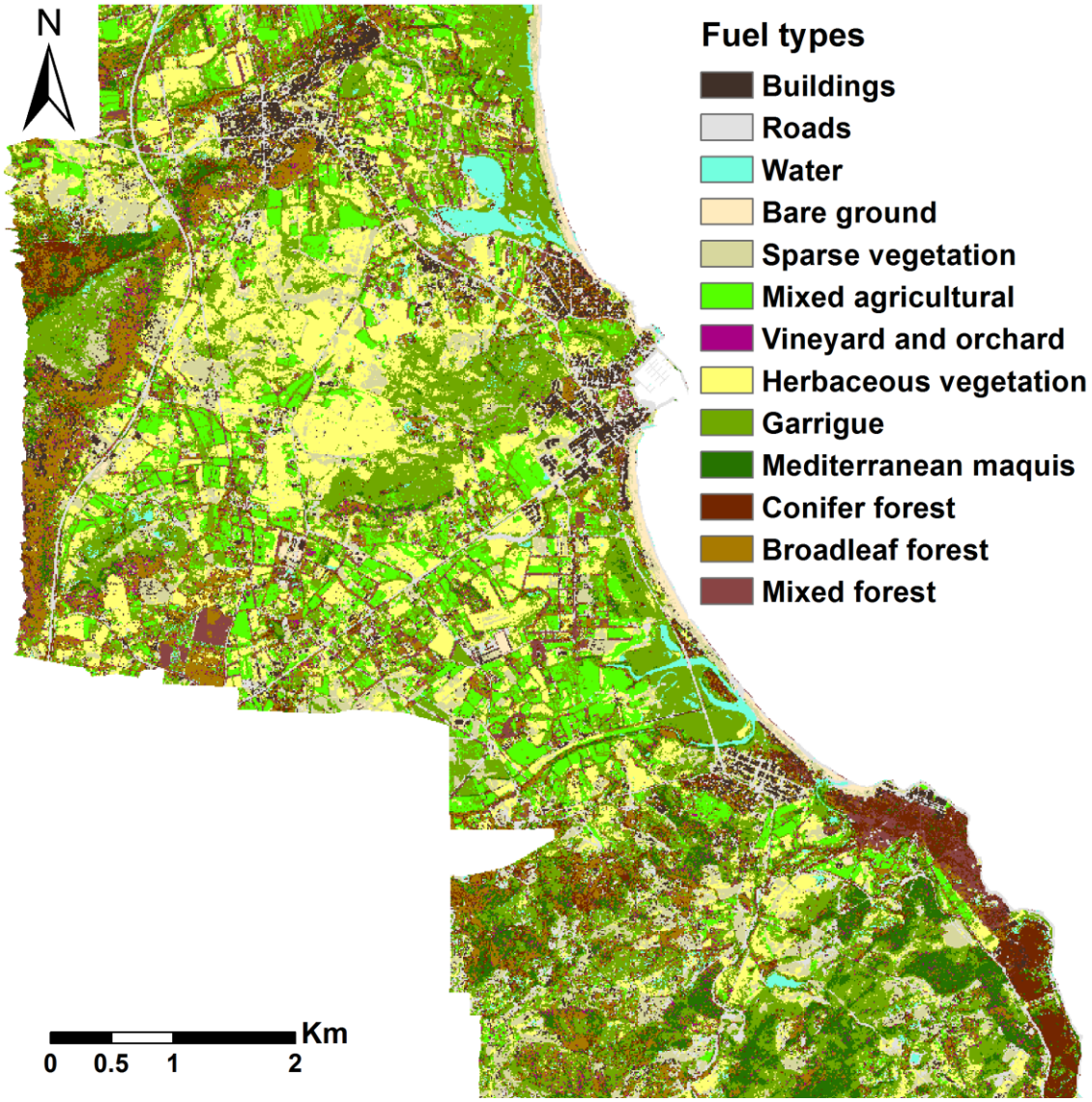

Fig. A.2. Classification output map for the study area of Siniscola using the Neural Networks (NN) algorithm with the selection of variables.

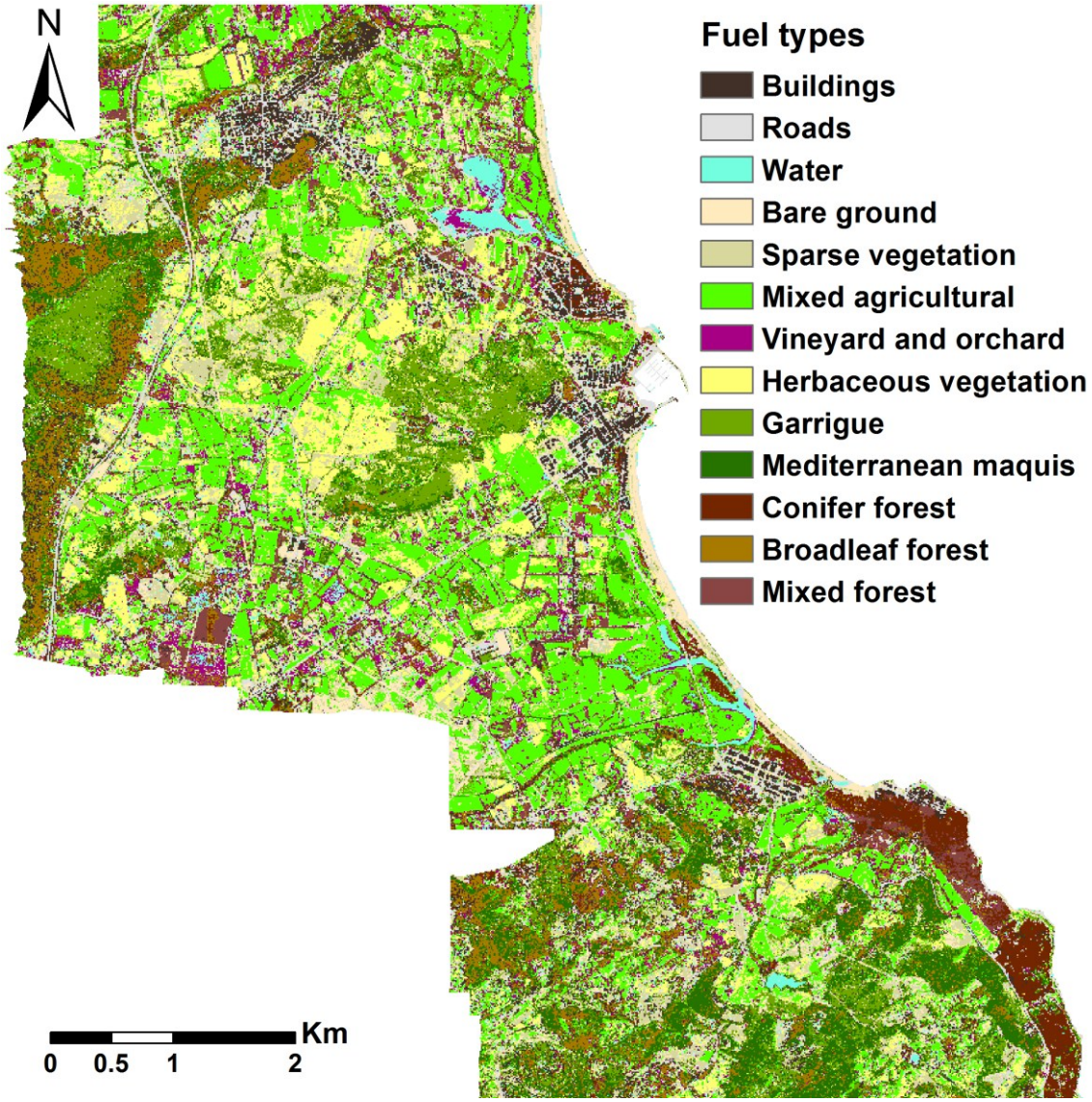

Fig. A.3. Classification output map for the study area of Siniscola using the Neural Networks (NN) algorithm with the all variables.

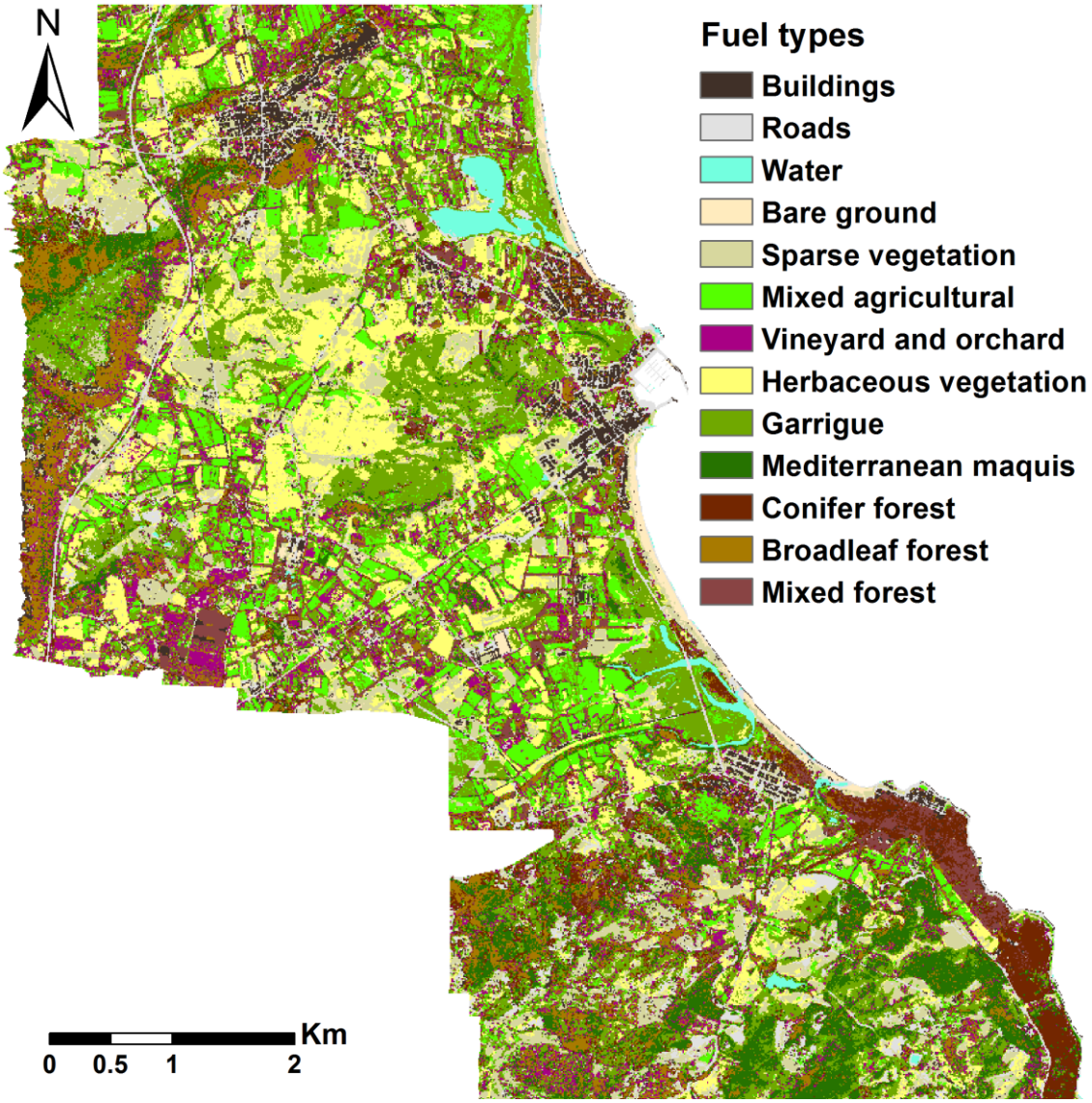

Fig. A.4. Classification output map for the study area of Siniscola using the Support Vector Machine (SVM) algorithm with the selection of variables.

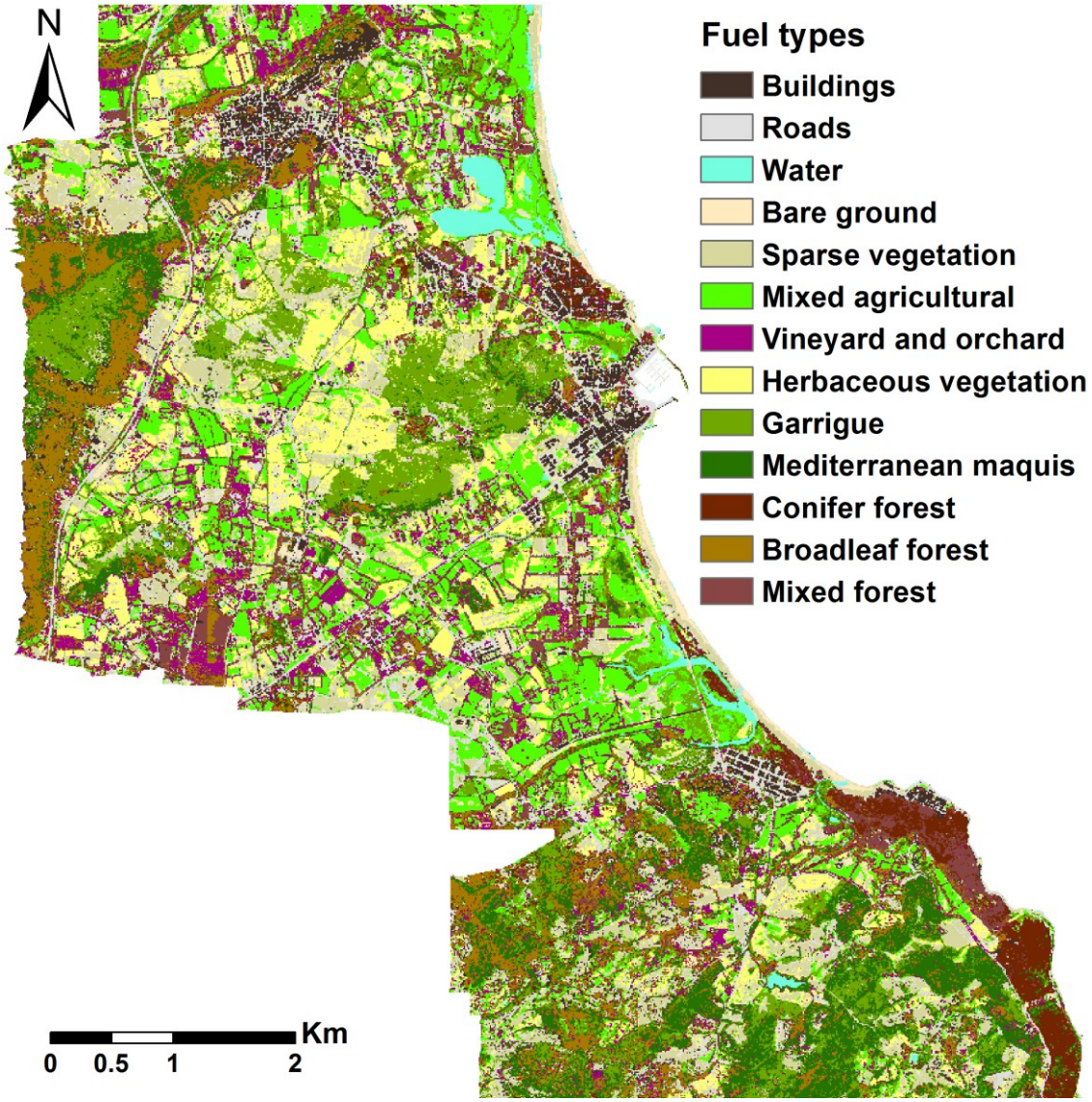

Fig. A.5. Classification output map for the study area of Siniscola using the Random Forest (RF) algorithm with all variables.

Appendix B. Maps obtained from the classifications outputs carried out with the different algorithms and in the study area of Muravera.

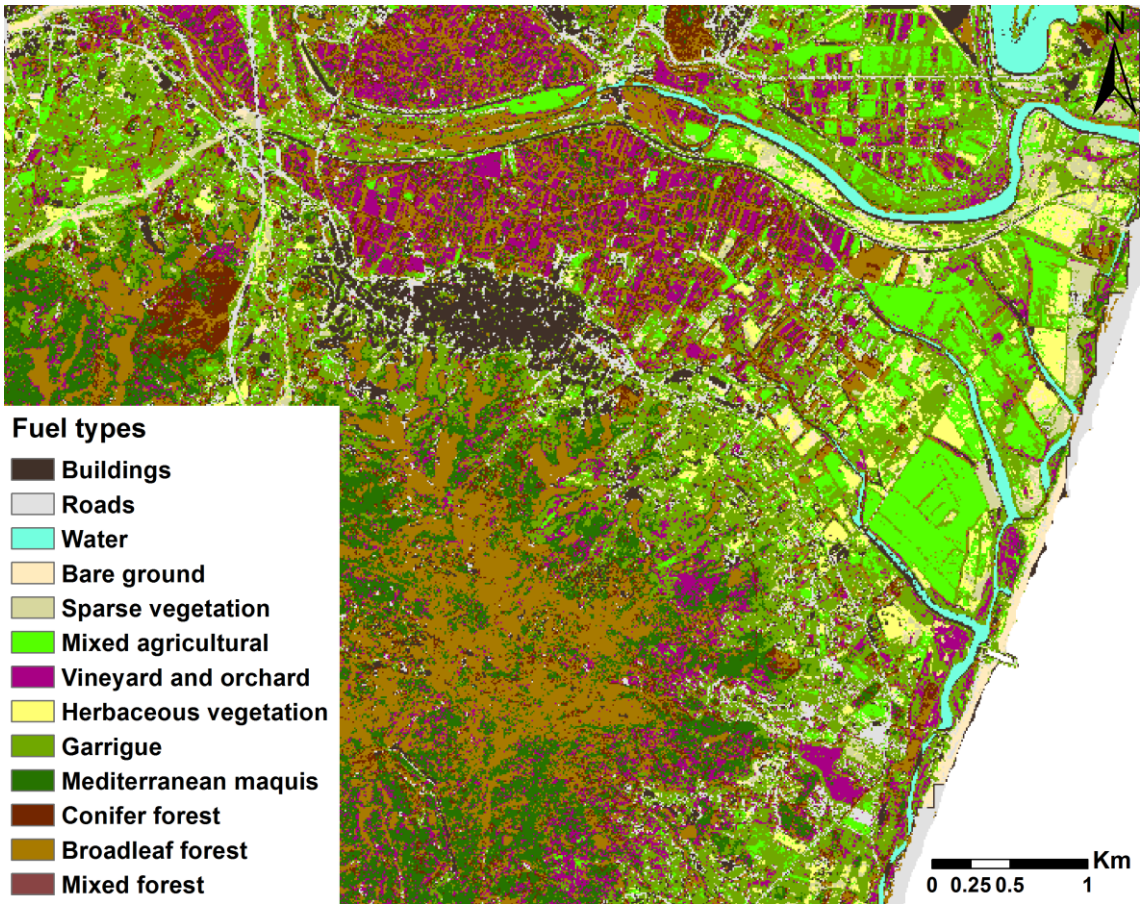

Fig. B.1. Classification output map for the study area of Muravera using the Maximum Likelihood (ML) algorithm with the selection of variables.

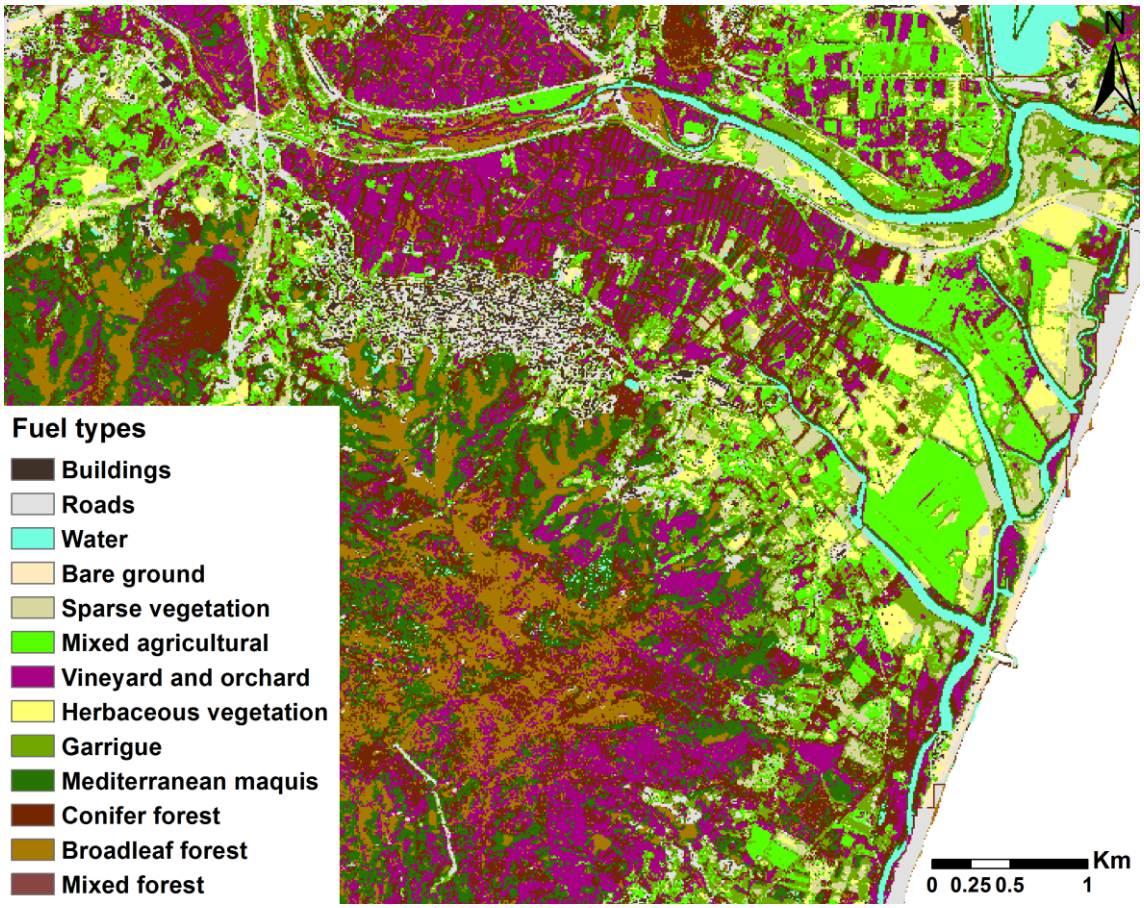

Fig. B.2. Classification output map for the study area of Muravera using the Neural Networks (NN) algorithm with the selection of variables.

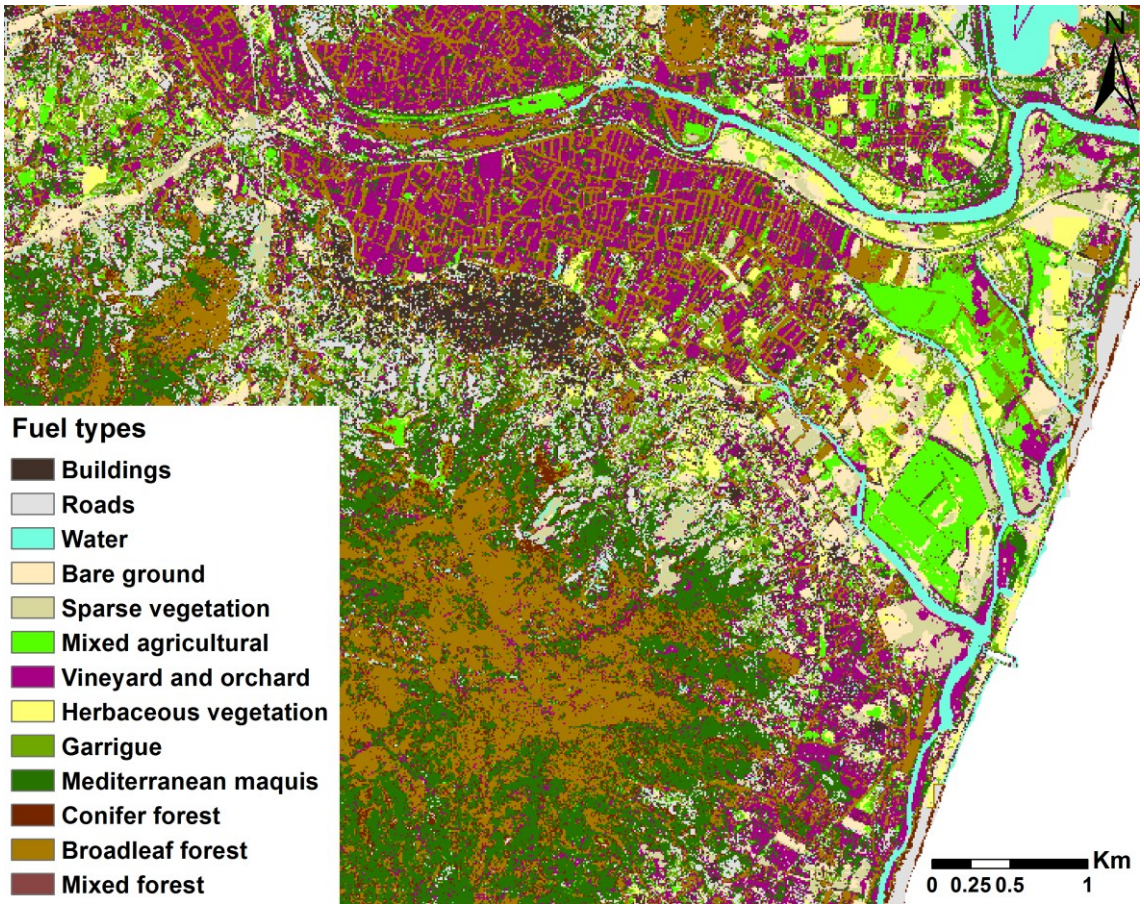

Fig. B.3. Classification output map for the study area of Muravera using the Neural Networks (NN) algorithm with the all variables.

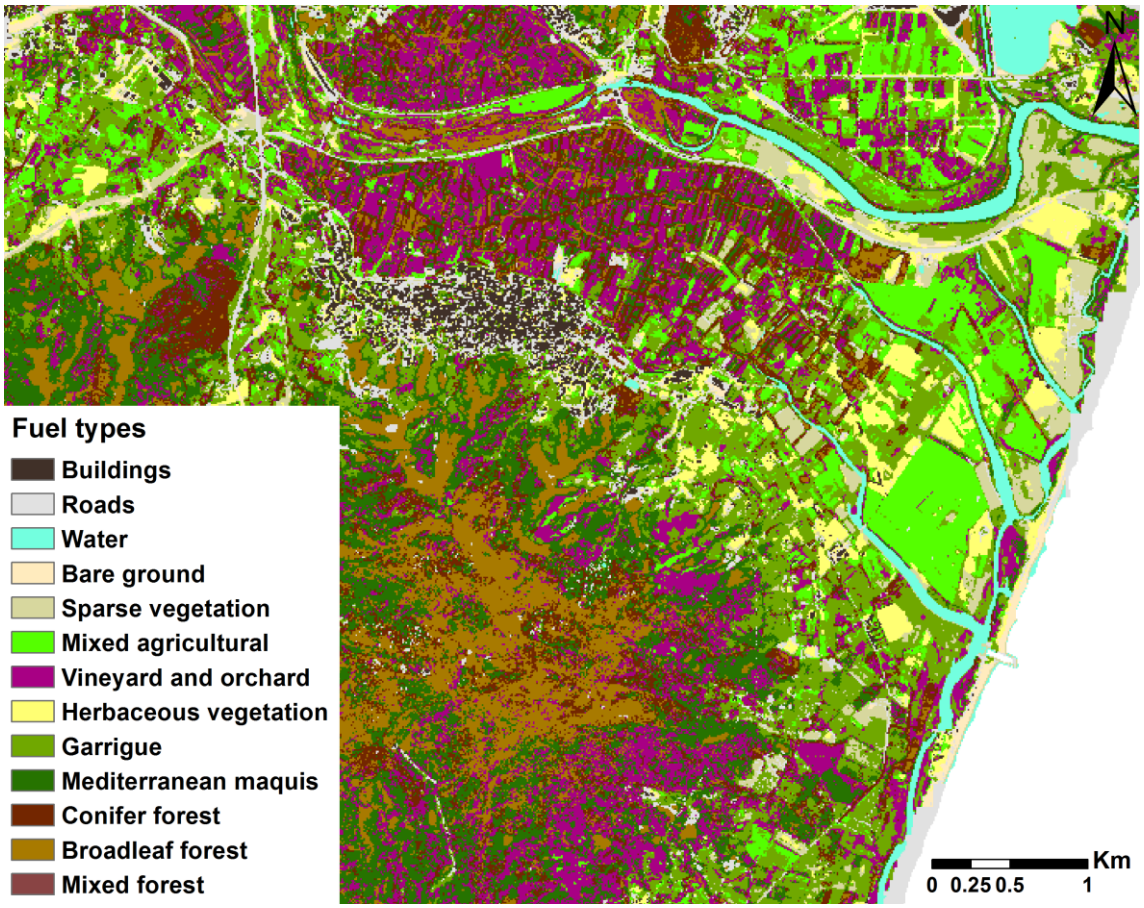

Fig. B.4. Classification output map for the study area of Muravera using the Support Vector Machine (SVM) algorithm with the selection of variables.

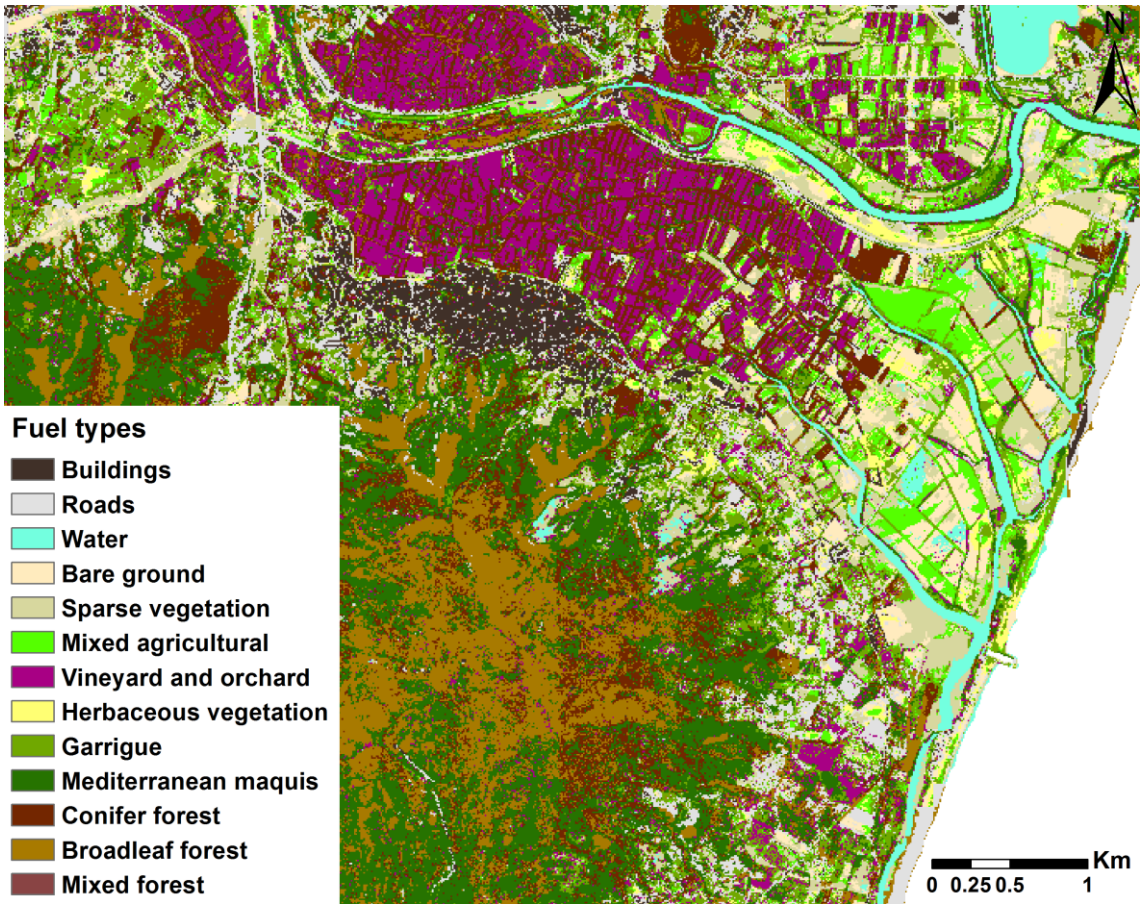

Fig. B.5. Classification output map for the study area of Muravera using the Random Forest (RF) algorithm with all variables.

Appendix C. Confusion matrices from the classifications carried out with the different algorithms for the study area of Siniscola.

Table C.1. Confusion matrix obtained from the classification carried out with the maximum likelihood algorithm (ML) using the subset of variables in the Siniscola study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction													Producer's Accuracy
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**(%)**
Reference data	**1**	79	4	0	1	4	0	0	0	0	0	0	0	0	89.77
	2	5	88	0	8	5	0	0	2	0	0	0	0	0	81.48
	3	0	0	23	0	0	0	0	0	0	0	0	0	0	100.00
	4	0	0	0	46	5	0	0	4	0	0	0	0	0	83.64
	5	2	3	0	16	40	0	0	11	1	0	0	0	0	54.79
	6	0	0	0	0	1	33	0	1	1	0	0	0	0	91.67
	7	0	0	0	0	0	0	14	0	0	0	0	0	0	100.00
	8	0	0	0	0	0	0	0	36	2	0	0	0	0	94.74
	9	0	1	0	0	0	0	0	0	36	0	0	0	0	97.30
	10	0	0	0	0	1	0	4	0	1	39	0	1	0	84.78
	11	0	0	0	0	0	0	0	0	0	0	40	0	2	95.24
	12	0	0	0	0	1	0	0	0	0	0	0	28	4	84.85
	13	0	0	0	0	0	0	0	0	0	0	3	6	29	76.32
User's Accuracy (%)		91.86	91.67	100.00	64.79	70.18	100.00	77.78	66.67	87.80	100.00	93.02	80.00	82.86	

Table C.2. Confusion matrix obtained from the classification carried out with the Neural Networks (NN) using the subset of variables in the Siniscola study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction													Producer's Accuracy
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**(%)**
Reference data	**1**	78	6	0	5	1	0	0	0	0	0	0	0	1	85.71
	2	0	103	1	0	4	0	0	3	1	0	0	0	0	91.96
	3	0	0	25	0	0	0	0	0	0	0	0	0	0	100.00
	4	0	1	0	33	10	0	0	0	0	0	0	0	0	75.00
	5	1	4	0	8	53	0	0	6	0	1	0	0	0	72.60
	6	0	0	0	0	0	28	0	0	5	0	0	0	0	84.85
	7	0	0	0	0	0	0	2	2	5	7	0	6	0	9.09
	8	0	1	0	0	3	0	0	23	1	0	1	0	0	79.31
	9	0	0	1	0	2	0	0	2	33	0	0	0	0	86.84
	10	0	0	0	0	0	0	1	0	3	38	0	0	0	90.48
	11	0	0	0	0	0	0	0	0	0	0	39	2	2	90.70
	12	1	0	0	0	0	0	2	0	0	2	7	23	0	65.71
	13	1	0	0	0	0	0	0	0	0	0	10	6	27	61.36
	User's Accuracy (%)	96.30	89.57	92.59	71.74	72.60	100.00	40.00	63.89	68.75	79.17	68.42	62.16	90.00	

Table C.3. Confusion matrix obtained from the classification carried out with the Neural Networks (NN) using all variables in the Siniscola study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction													Producer's Accuracy
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**(%)**
Reference data	**1**	72	11	1	4	1	0	0	0	0	0	2	0	0	79.12
	2	4	90	1	4	8	0	1	4	0	0	0	0	0	80.36
	3	0	0	23	0	0	0	2	0	0	0	0	0	0	92.00
	4	0	2	0	37	5	0	0	0	0	0	0	0	0	84.09
	5	1	1	0	21	39	6	0	2	3	0	0	0	0	53.42
	6	0	0	0	0	0	32	0	0	1	0	0	0	0	96.97
	7	0	0	3	0	0	0	18	0	0	0	0	1	0	81.82
	8	0	0	0	0	3	9	0	17	0	0	0	0	0	58.62
	9	0	0	0	0	1	2	0	0	32	3	0	0	0	84.21
	10	0	0	0	0	0	0	0	0	4	37	0	1	0	88.10
	11	1	0	0	0	0	0	0	0	0	0	38	1	3	88.37
	12	2	0	0	0	0	0	1	0	0	2	4	26	0	74.29
	13	0	0	0	0	0	0	0	0	0	0	15	0	29	65.91
User's Accuracy (%)		90.00	86.54	82.14	56.06	68.42	65.31	81.82	73.91	80.00	88.10	64.41	89.66	90.63	

Table C.4. Confusion matrix obtained from the classification carried out with the Support Vector Machine (SVM) using the subset of variables in the Siniscola study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction													Producer's Accuracy
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**(%)**
Reference data	**1**	82	4	0	5	0	0	0	0	0	0	0	0	0	90.11
	2	1	103	0	0	4	0	0	4	0	0	0	0	0	91.96
	3	0	0	25	0	0	0	0	0	0	0	0	0	0	100.00
	4	1	2	0	31	10	0	0	0	0	0	0	0	0	70.45
	5	3	6	0	5	54	0	0	5	0	0	0	0	0	73.97
	6	0	0	0	0	0	30	0	0	3	0	0	0	0	90.91
	7	0	0	0	0	0	0	15	0	0	6	0	1	0	68.18
	8	0	3	0	0	4	0	0	22	0	0	0	0	0	75.86
	9	0	0	0	0	0	0	0	0	38	0	0	0	0	100.00
	10	0	0	0	0	0	0	1	0	1	40	0	0	0	95.24
	11	0	0	0	0	0	0	0	0	0	0	41	0	2	95.35
	12	0	0	0	0	0	0	1	0	0	2	3	23	6	65.71
	13	0	0	0	0	0	0	0	0	0	0	4	0	40	90.91
User's Accuracy (%)		94.25	87.29	100.00	75.61	75.00	100.00	88.24	70.97	90.48	83.33	85.42	95.83	83.33	

Table C.5. Confusion matrix obtained from the classification carried out with the Random Forests (RF) using the subset of variables in the Siniscola study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction													Producer's Accuracy
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**(%)**
Reference data	**1**	79	4	0	3	2	0	0	0	0	0	0	0	0	89.77
	2	1	102	0	1	1	0	0	3	0	0	0	0	0	94.44
	3	0	0	23	0	0	0	0	0	0	0	0	0	0	100.00
	4	0	2	0	48	5	0	0	0	0	0	0	0	0	87.27
	5	2	5	0	5	60	0	0	1	0	0	0	0	0	82.19
	6	0	0	0	0	0	34	0	0	2	0	0	0	0	94.44
	7	0	0	0	0	0	0	12	0	0	2	0	0	0	85.71
	8	0	0	0	0	7	0	0	31	0	0	0	0	0	81.58
	9	0	0	1	0	0	0	0	0	36	0	0	0	0	97.30
	10	0	0	0	0	0	0	0	0	1	45	0	0	0	97.83
	11	0	0	0	0	0	0	0	0	0	0	42	0	0	100.00
	12	0	0	0	0	0	0	0	0	0	0	1	30	2	90.91
	13	0	0	0	0	0	0	0	0	0	0	0	2	36	94.74
User's Accuracy (%)		96.34	90.27	95.83	84.21	80.00	100.00	100.00	88.57	92.31	95.74	97.67	93.75	94.74	

Table C.6. Confusion matrix obtained from the classification carried out with the Random Forests (RF) using all variables in the Siniscola study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction													Producer's Accuracy
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**(%)**
Reference data	**1**	78	5	0	2	2	0	0	0	0	1	0	0	0	88.64
	2	3	100	0	1	3	0	0	1	0	0	0	0	0	92.59
	3	0	0	23	0	0	0	0	0	0	0	0	0	0	100.00
	4	0	1	0	41	13	0	0	0	0	0	0	0	0	74.55
	5	2	3	0	5	62	0	0	1	0	0	0	0	0	84.93
	6	0	0	0	0	1	34	0	0	1	0	0	0	0	94.44
	7	0	0	0	0	0	0	14	0	0	0	0	0	0	100.00
	8	0	1	0	0	7	0	0	30	0	0	0	0	0	78.95
	9	0	0	0	0	0	0	0	0	37	0	0	0	0	100.00
	10	0	0	0	0	0	0	1	0	1	44	0	0	0	95.65
	11	0	0	0	0	0	0	0	0	0	0	42	0	0	100.00
	12	0	0	0	0	0	0	0	0	0	0	0	31	2	93.94
	13	0	0	0	0	0	0	0	0	0	0	0	0	38	100.00
User's Accuracy (%)		93.98	90.91	100.00	83.67	70.45	100.00	93.33	93.75	94.87	97.78	100.00	100.00	95.00	

Table C.7. Confusion matrix obtained from the Land Use Map (LUM) of Sardinia in the Siniscola study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction													Producer's Accuracy
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**(%)**
Reference data	**1**	89	0	0	0	0	2	0	0	0	0	0	0	0	94.68
	2	30	16	0	3	0	30	8	0	6	5	0	3	0	15.84
	3	0	0	22	0	0	3	0	0	1	1	0	0	0	81.48
	4	9	0	0	31	0	1	0	0	7	1	0	2	0	60.78
	5	0	0	0	26	1	30	0	2	11	8	0	0	0	1.28
	6	0	0	0	0	0	43	0	1	0	0	0	0	0	97.73
	7	0	0	0	0	0	4	21	0	0	0	0	0	0	84.00
	8	0	0	0	0	0	27	0	4	0	0	0	0	0	12.90
	9	0	0	0	0	0	4	0	0	23	8	0	0	0	65.71
	10	0	0	0	0	0	0	0	0	4	31	0	1	0	86.11
	11	0	0	0	0	0	0	0	0	21	0	21	0	0	50.00
	12	0	0	0	0	0	0	0	0	0	2	0	32	0	94.12
	13	0	0	0	0	0	0	0	0	0	0	33	0	0	0.00
User's Accuracy (%)		69.53	100.00	100.00	51.67	100.00	29.86	72.41	57.14	30.26	55.36	38.89	84.21	0.00	

Appendix D. Confusion matrices from the classifications carried out with the different algorithms for the study area of Muravera.

Table D.1. Confusion matrix obtained from the classification carried out with the maximum likelihood algorithm (ML) using the subset of variables in the Muravera study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction												Producer's Accuracy (%)
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
Reference data	**1**	12	1	0	0	0	0	0	0	1	0	0	0	85.71
	2	1	16	0	0	2	0	0	0	1	0	0	0	80.00
	3	0	0	27	0	0	0	0	0	0	0	0	0	100.00
	4	0	0	0	13	2	0	0	0	0	0	0	0	86.67
	5	0	0	0	0	8	0	0	1	0	0	0	0	88.89
	6	0	0	0	0	0	15	0	0	2	0	0	0	88.24
	7	0	0	0	0	0	0	12	0	2	1	0	1	75.00
	8	0	0	0	1	0	0	0	20	4	0	0	0	80.00
	9	0	0	0	0	1	0	0	0	12	0	0	0	92.31
	10	0	1	0	0	0	0	3	0	0	12	0	3	63.16
	11	0	1	0	0	0	0	0	0	0	2	11	2	68.75
	12	0	0	0	0	0	0	0	0	0	0	1	10	90.91
	User's Accuracy (%)	92.31	84.21	100.00	92.86	61.54	100.00	80.00	95.24	54.55	80.00	91.67	62.50	

Table D.2. Confusion matrix obtained from the classification carried out with the Neural Networks (NN) using the subset of variables in the Muravera study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction												Producer's Accuracy (%)
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
Reference data	**1**	9	4	0	0	0	0	0	0	0	0	0	0	69.23
	2	1	18	0	1	1	2	0	0	1	0	0	0	75.00
	3	0	0	27	0	0	0	0	0	0	0	0	0	100.00
	4	0	2	0	11	0	0	0	0	0	0	0	0	84.62
	5	0	0	0	1	9	0	0	0	3	0	0	0	69.23
	6	0	0	0	0	0	21	0	0	0	0	0	0	100.00
	7	0	0	0	0	0	0	12	0	0	2	0	0	85.71
	8	0	0	0	1	0	0	0	19	1	0	0	0	90.48
	9	0	0	1	0	0	2	2	2	5	2	0	0	35.71
	10	0	0	0	0	0	0	4	0	0	9	0	0	69.23
	11	0	1	0	0	0	0	0	0	0	0	10	1	83.33
	12	0	0	0	0	0	0	1	0	0	0	3	13	76.47
User's Accuracy (%)		90.00	72.00	96.43	78.57	90.00	84.00	63.16	90.48	50.00	69.23	76.92	92.86	

Table D.3. Confusion matrix obtained from the classification carried out with the Neural Networks (NN) using all variables in the Muravera study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction												Producer's Accuracy (%)
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
Reference data	**1**	9	3	0	0	0	0	1	0	0	0	0	0	69.23
	2	5	13	0	2	0	0	1	0	0	1	1	1	54.17
	3	0	0	27	0	0	0	0	0	0	0	0	0	100.00
	4	2	0	0	10	0	0	0	1	0	0	0	0	76.92
	5	0	0	0	2	11	0	0	0	0	0	0	0	84.62
	6	0	0	0	1	0	19	0	1	0	0	0	0	90.48
	7	0	0	0	0	0	0	14	0	0	0	0	0	100.00
	8	0	0	0	0	0	0	0	19	2	0	0	0	90.48
	9	1	0	0	0	0	1	0	2	9	1	0	0	64.29
	10	0	1	0	0	0	0	3	0	0	9	0	0	69.23
	11	0	1	0	0	0	0	0	0	0	0	0	11	0.00
	12	0	0	0	0	0	0	0	0	0	0	1	16	94.12
User's Accuracy (%)		52.94	72.22	100.00	66.67	100.00	95.00	73.68	82.61	81.82	81.82	0.00	57.14	

Table D.4. Confusion matrix obtained from the classification carried out with the Support Vector Machine (SVM) using the subset of variables in the Muravera study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction												Producer's Accuracy (%)
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
Reference data	**1**	10	2	0	0	0	0	0	1	0	0	0	0	76.92
	2	0	19	0	1	2	0	0	0	2	0	0	0	79.17
	3	0	0	27	0	0	0	0	0	0	0	0	0	100.00
	4	0	0	0	13	0	0	0	0	0	0	0	0	100.00
	5	0	0	0	1	10	0	0	0	2	0	0	0	76.92
	6	0	0	0	0	2	19	0	0	0	0	0	0	90.48
	7	0	0	0	0	0	0	9	0	1	3	1	0	64.29
	8	0	0	0	0	0	0	0	20	1	0	0	0	95.24
	9	0	0	0	0	0	0	2	2	9	1	0	0	64.29
	10	0	0	0	0	0	0	1	0	1	11	0	0	84.62
	11	0	1	0	0	0	0	0	0	0	0	10	1	83.33
	12	0	0	0	0	0	0	0	0	0	0	4	13	76.47
User's Accuracy (%)		100.00	86.36	100.00	86.67	71.43	100.00	75.00	86.96	56.25	73.33	66.67	92.86	

Table D.5. Confusion matrix obtained from the classification carried out with the Random Forests (RF) using the subset of variables in the Muravera study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction												Producer's Accuracy (%)
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
Reference data	**1**	13	1	0	0	0	0	0	0	0	0	0	0	92.86
	2	0	16	0	1	0	0	0	0	3	0	0	0	80.00
	3	0	0	27	0	0	0	0	0	0	0	0	0	100.00
	4	0	1	0	12	2	0	0	0	0	0	0	0	80.00
	5	0	0	0	0	9	0	0	0	0	0	0	0	100.00
	6	0	0	0	0	0	15	0	2	0	0	0	0	88.24
	7	0	0	0	0	0	0	14	0	1	1	0	0	87.50
	8	0	0	0	1	0	0	0	24	0	0	0	0	96.00
	9	0	0	0	0	1	0	0	0	11	0	1	0	84.62
	10	0	0	0	0	0	0	0	0	0	19	0	0	100.00
	11	0	0	0	0	0	0	0	0	0	1	15	0	93.75
	12	0	0	0	0	0	0	0	0	0	0	2	9	81.82
	User's Accuracy (%)	100.00	88.89	100.00	85.71	75.00	100.00	100.00	92.31	73.33	90.48	83.33	100.00	

Table D.6. Confusion matrix obtained from the classification carried out with the Random Forests (RF) using all variables in the Muravera study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction												Producer's Accuracy (%)
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
Reference data	**1**	13	1	0	0	0	0	0	0	0	0	0	0	92.86
	2	0	13	0	1	0	0	0	1	5	0	0	0	65.00
	3	0	0	27	0	0	0	0	0	0	0	0	0	100.00
	4	0	1	0	12	2	0	0	0	0	0	0	0	80.00
	5	0	0	0	0	9	0	0	0	0	0	0	0	100.00
	6	0	0	0	0	0	14	0	3	0	0	0	0	82.35
	7	0	1	0	0	0	0	14	0	0	1	0	0	87.50
	8	0	1	0	0	0	1	0	23	0	0	0	0	92.00
	9	0	1	0	0	1	0	0	0	10	0	1	0	76.92
	10	0	0	0	0	0	0	1	0	0	18	0	0	94.74
	11	0	0	0	0	0	0	0	0	0	0	16	0	100.00
	12	0	0	0	0	0	0	0	0	0	0	2	9	81.82
User's Accuracy (%)		100.00	72.22	100.00	92.31	75.00	93.33	93.33	85.19	66.67	94.74	84.21	100.00	

Table D.7. Confusion matrix obtained from the Land Use Map (LUM) of Sardinia in the Muravera study area. 1 = Buildings; 2 = Roads; 3 = Water; 4 = Bare ground; 5 = Sparse vegetation; 6 = Mixed agricultural; 7 = Vineyard and orchard; 8 = Herbaceous vegetation; 9 = Garrigue; 10 = Mediterranean maquis; 11 = Conifer forests; 12 = Broadleaf forests; 13 = Mixed forests.

		Prediction												Producer's Accuracy (%)
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
Reference data	**1**	11	0	0	0	0	0	0	0	0	0	0	0	100.00
	2	0	0	1	9	0	7	4	0	0	0	0	0	0.00
	3	0	0	26	0	0	0	0	0	0	0	0	0	100.00
	4	0	0	4	5	0	1	0	0	1	1	0	0	41.67
	5	0	0	6	0	0	1	0	0	1	0	0	0	0.00
	6	0	3	0	0	0	9	1	2	0	0	0	0	60.00
	7	0	0	0	0	0	0	14	0	0	0	0	0	100.00
	8	0	0	0	0	0	12	2	11	0	0	0	0	44.00
	9	0	0	0	0	0	10	0	0	4	0	0	0	28.57
	10	0	0	0	0	0	0	0	0	6	8	2	1	47.06
	11	0	0	0	0	0	0	0	0	0	0	16	0	100.00
	12	0	0	0	0	0	0	0	0	2	2	1	18	78.26
User's Accuracy (%)		100.00	0.00	70.27	35.71	0.00	22.50	66.67	84.62	28.57	72.73	84.21	94.74	

Appendix E. Outputs from the simulations carried out using as input the different landscape files.

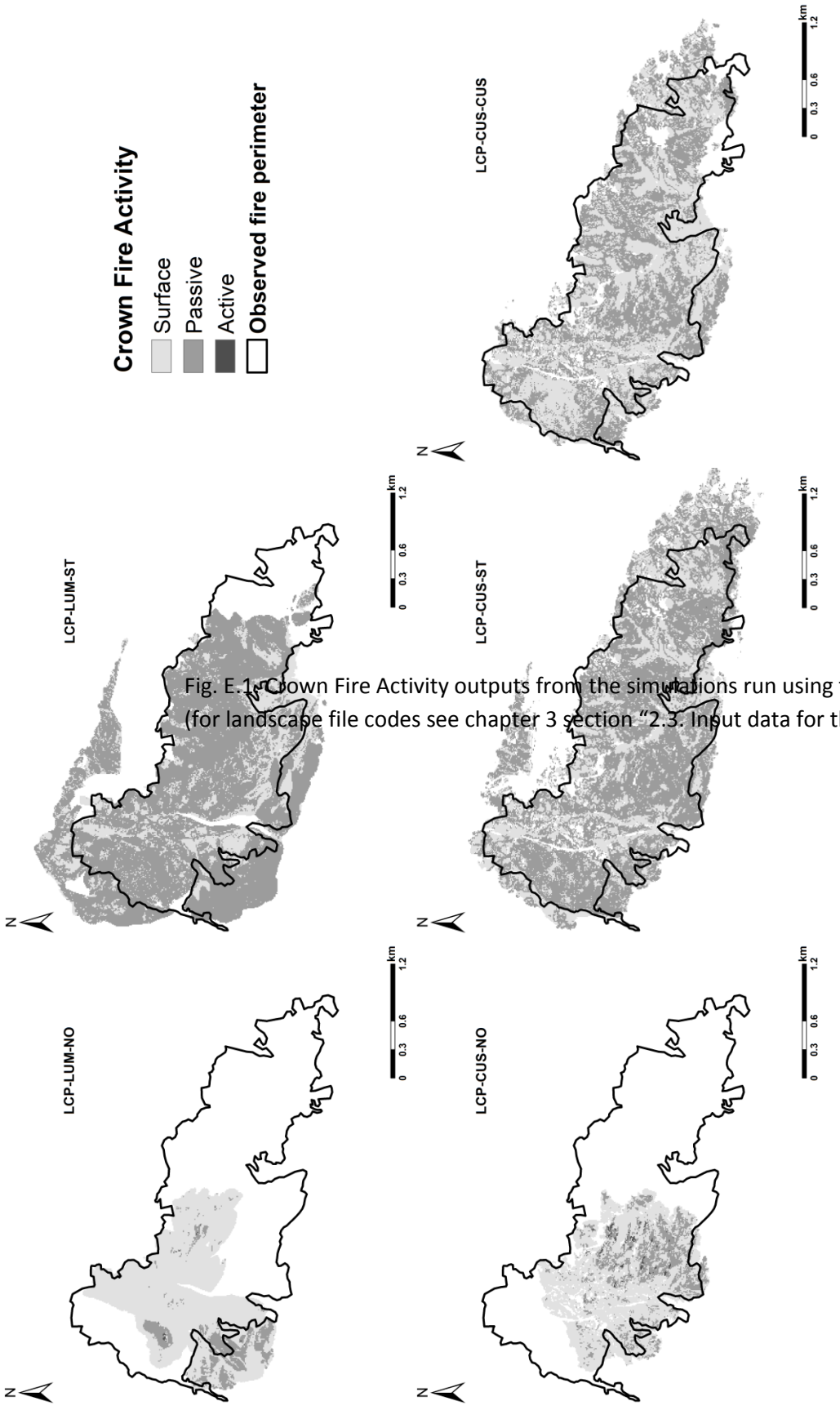

Fig. E.1. Crown Fire Activity outputs from the simulations run using the different landscape LCP (for landscape file codes see chapter 3 section "2.3. Input data for the simulations")

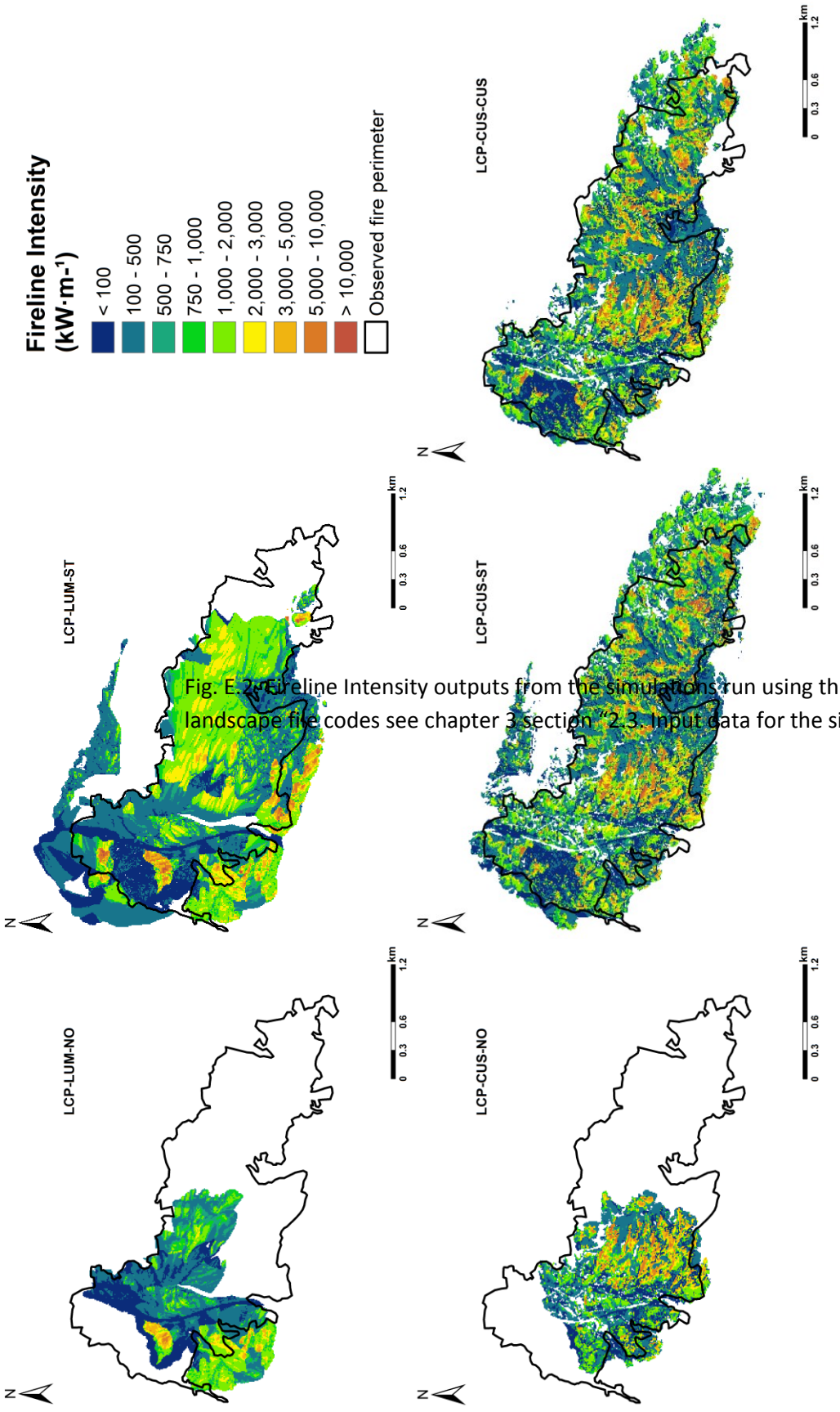

Fig. E.2. Fireline Intensity outputs from the simulations run using the different landscape LCP = l
landscape file codes see chapter 3 section "2.3. Input data for the simulations")

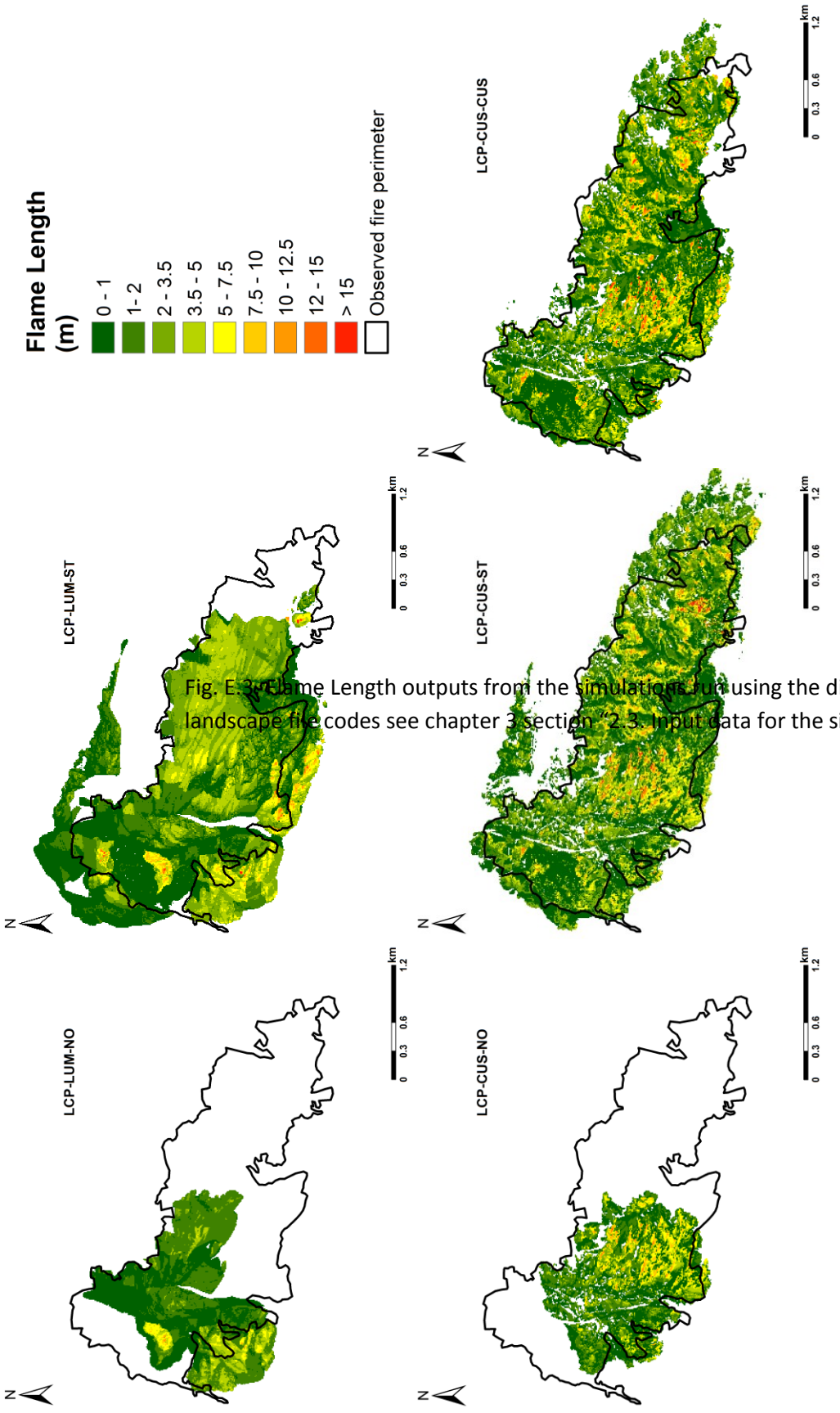

Fig. E.3: Flame Length outputs from the simulations run using the different landscape LCP = land
landscape file codes see chapter 3 section "2.3. Input data for the simulations")

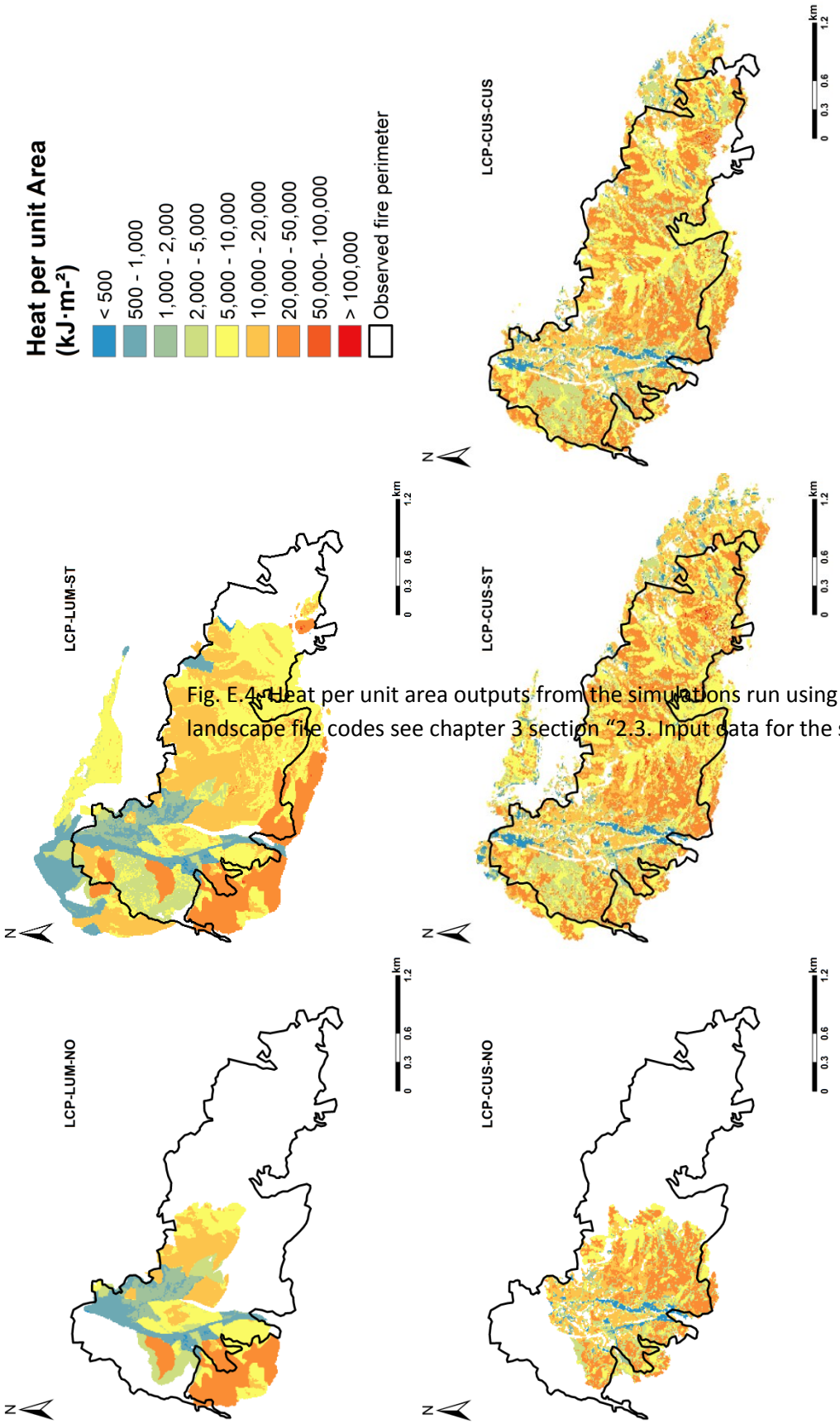

Fig. E.4 Heat per unit area outputs from the simulations run using the different landscape LCP = landscape file codes see chapter 3 section "2.3. Input data for the simulations")

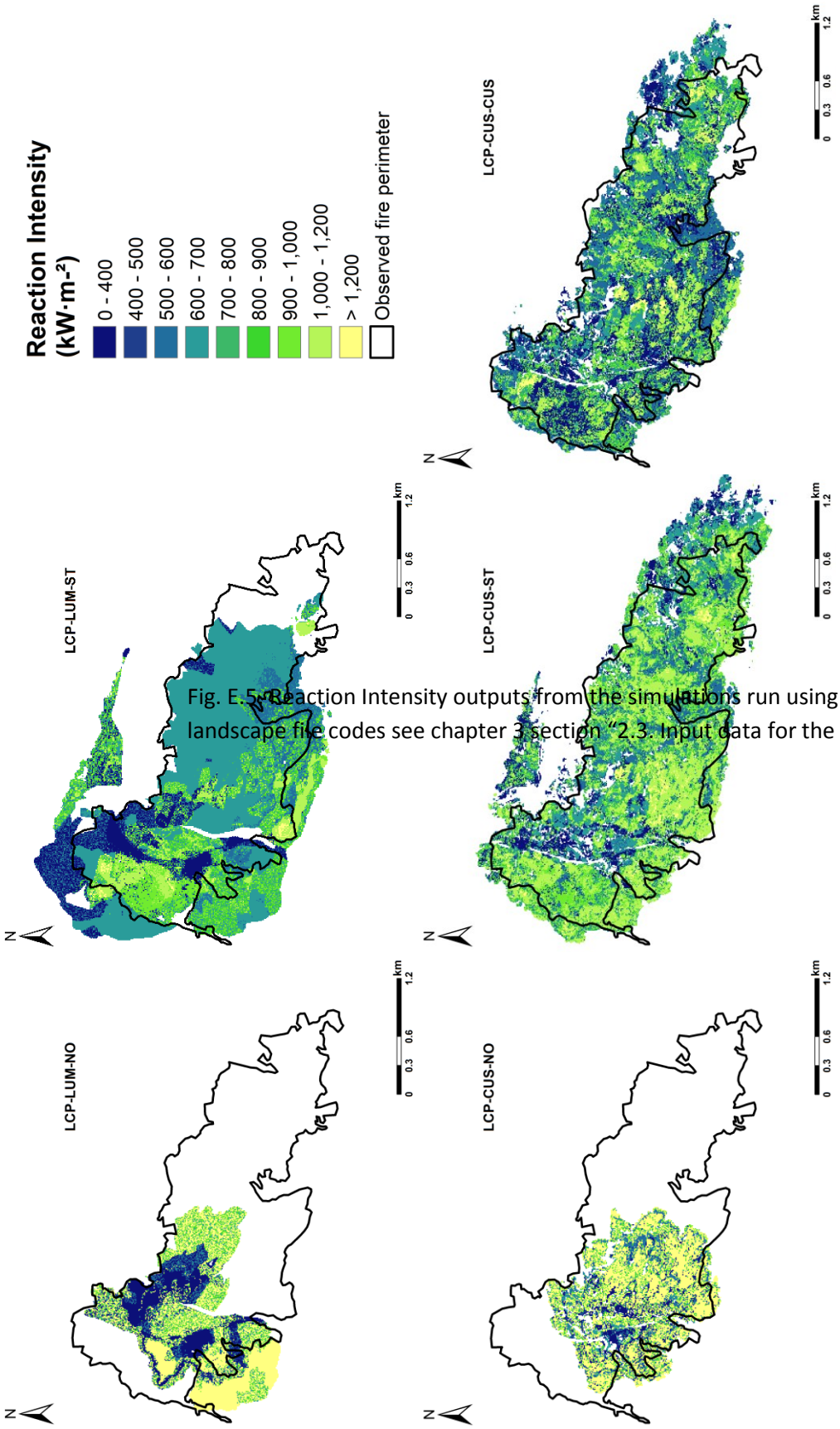

Fig. E.5. Reaction Intensity outputs from the simulations run using the different landscape LCP = landscape file codes see chapter 3 section "2.3. Input data for the simulations")

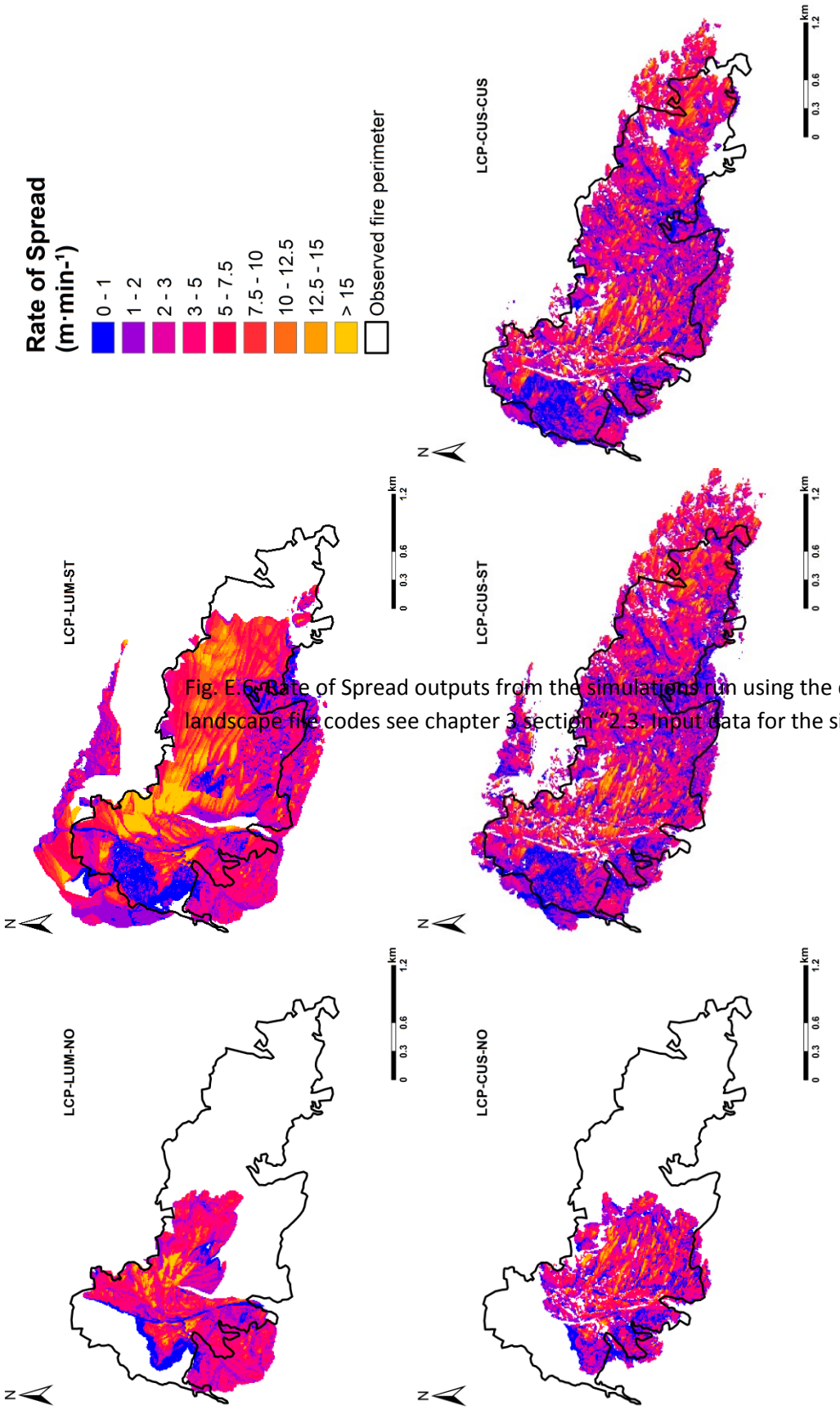

Fig. E.6. Rate of Spread outputs from the simulations run using the different landscape LCP = lan
landscape file codes see chapter 3 section "2.3. Input data for the simulations")

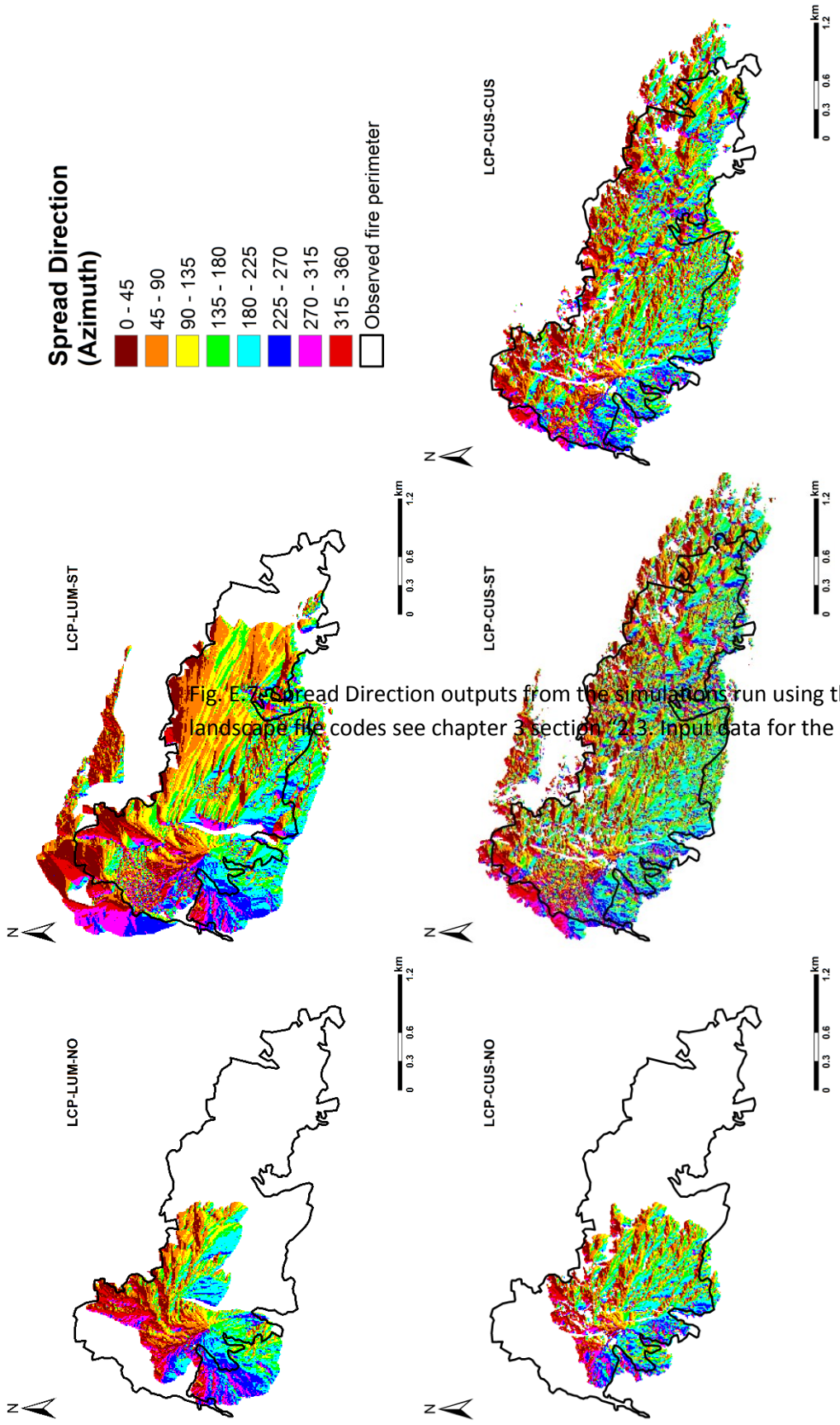

Fig. E 7 - Spread Direction outputs from the simulations run using the different landscape LCP = l
landscape file codes see chapter 3 section "2.3. Input data for the simulations")

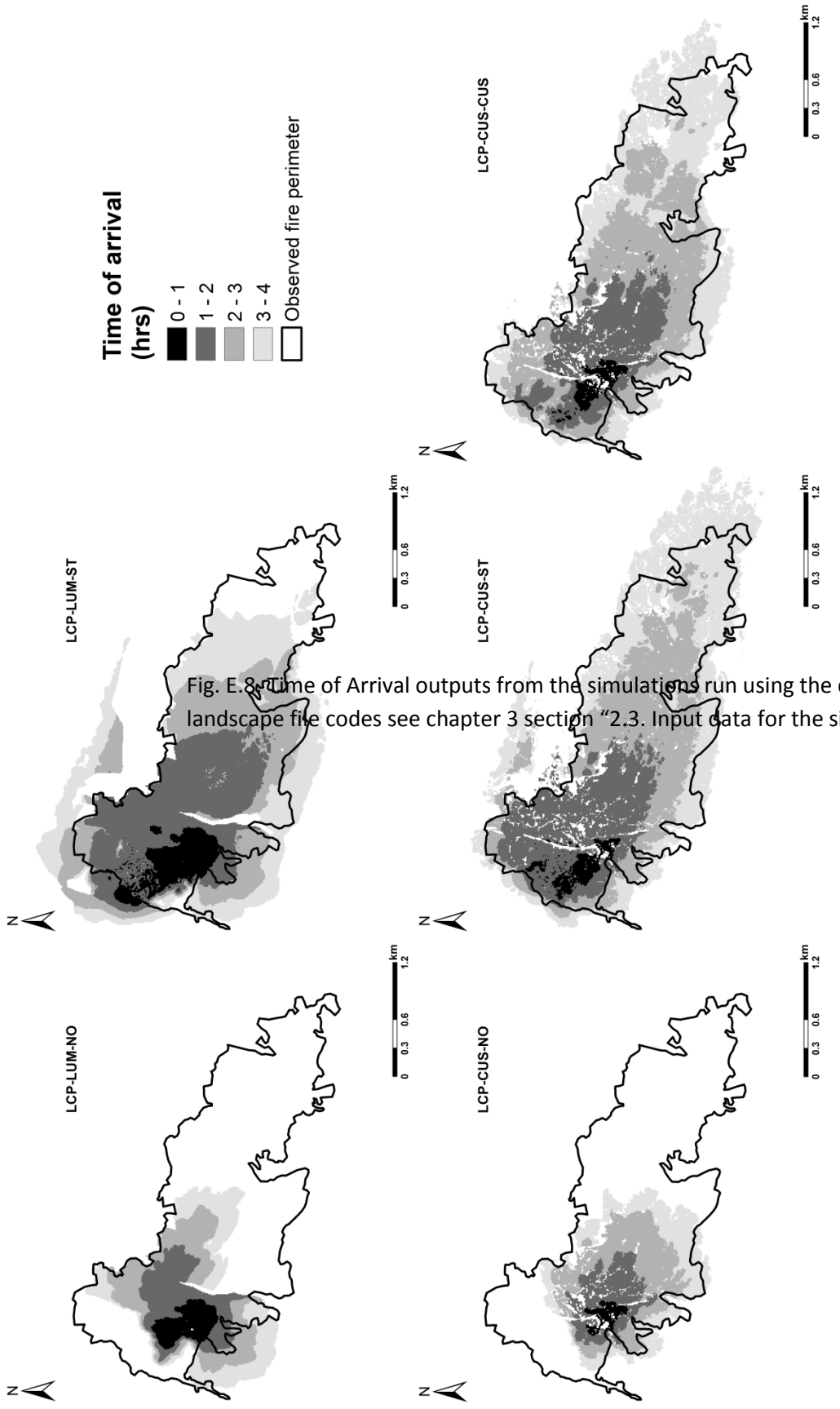

Fig. E.8. Time of Arrival outputs from the simulations run using the different landscape LCP = lan
landscape file codes see chapter 3 section "2.3. Input data for the simulations")