Machine Learning techniques for wildfire management

Hamed Izadgoshasb
M.Sc. Engineering for Natural Risk Management

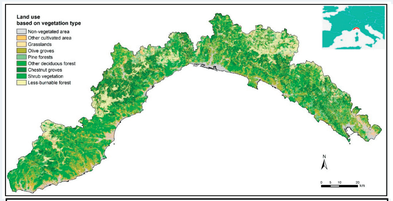
Supervisors:
Prof. Paolo Fiorucci & Dr. Andrea Trucchia

Abstract

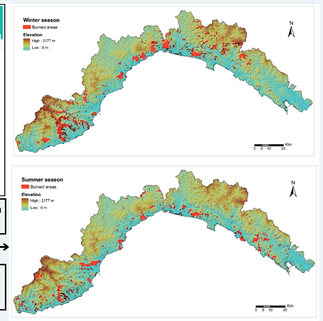
In the Mediterranean region, wildfires are a very important social and environmental problem that influence infrastructure, ecosystems, and human life. Therefore, creating susceptibility maps needs to be defined for managing wildfires. According to the inherent qualities of the territory, the wildfire susceptibility is described as a static probability of experiencing wildfire in a specific location. Moreover, a contemporary tool for addressing fuel management and supporting forest preservation programs is susceptibility mapping. This dissertation discusses a method based on Multilayer Perceptron (MLP), Support Vector Machine (SVM), and Random Forest (RF) that enables the creation of a map of the Liguria region's susceptibility to wildfires. Due to its dense and varied vegetation, which includes the presence of forests on more than 70% of its surface, as well as its ideal climatic circumstances, this area is severely damaged by wildfires. The dataset of the mapped fire perimeters, covering a 21-year period (1997-2017), and several geo-environmental predisposing factors are considered for assessing susceptibility (i.e., derivatives, altitude, land cover, road network, and vegetation type). Additionally, the influence of various local and neighboring vegetation types on the occurrence of wildfires is assessed using a more precise vegetation map as a predictor. Then, wildfire susceptibility maps were developed using the output values from the three machine learning methods that were probabilistically predicted. Random Forest has the best performance amongst three methods with the highest values of AUC (0.944 and 0.953, respectively) and lower values in both seasons (0.329 in summer and 0.335 in winter). Vegetation turned out to be by far the most significant predictor variable, as shown by Random Forest. In order to validate the creation of the summer and winter maps in terms of its capability to detect and identify potential future catastrophic events, another event-based validation is used based on the quantile analysis of susceptibility distribution over a set of burned areas for 2018 to 2021, which had no impact on the ML process.

Study Area

The region of interest for this study is **Liguria** which is located in the northwest of Italy. The following factors significantly affect how wildfires are distributed both spatially and temporally in this area: 1) the climate, which is marked by long dry periods in summer, especially along the coast, but also in winter, caused by dry winds from the north, creating curing conditions for the majority of herbaceous species; 2) the area's topographic assessment, which favours the spread of fire along the forested steep slopes; 3) the heterogeneous vegetation, characterised by a high percentage of forest canopy; and 4) human pressure, in terms of rural exodus, urbanisation, and greening



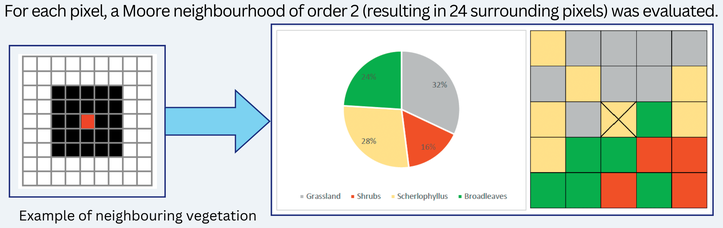
Land use derived from the combination of the initial classifications shown on the regional map of forest types, provided by the Liguria region.



Spatial distribution of burned areas during the winter (on the top) and summer season (on the bottom) in the Liguria region (Italy) during the period 1997–2017.

Neighbouring method

For each pixel, a Moore neighbourhood of order 2 (resulting in 24 surrounding pixels) was evaluated.



Example of neighbouring vegetation

General methodological workflow

I Development of the input data set: Pre-processing of the raster defining the independent variable and the predisposing factors.

II The testing and training subsets are chosen: In order to verify that the testing subset accurately reflected the potential wildfire patterns, three out of the 17 years were chosen at random.

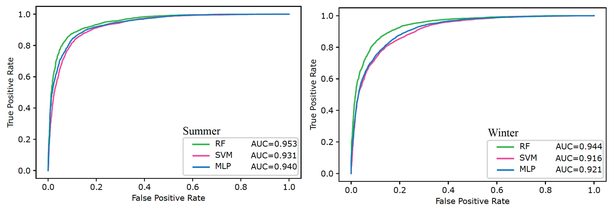
III Dividing the training subset into five folds, each of which was kept out of the validation subset selection process (through spatial-cross validation). The remaining four folds were used to train the model.

IV Application of machine learning Algorithms for the spatially predicting of wildfire susceptibility;
Multi-Layer Perceptron (MLP)
Support Vector Machine (SVM)
and Random Forest (RF)

- Quantile analysis of susceptibility distribution over a set of burned areas to assess the robustness of ML algorithm performance.
- Evaluation of the ranking and marginal impact on the projected outcome to determine the significance of the predictors (with RF).
- Development of wildfire susceptibility maps using the probabilistic outputs from the three models that were built using machine learning.
- Review of the performance metrics for the two seasons and each ML method (The AUC, ROC curve, and the root mean square error (RMSE)).

Results

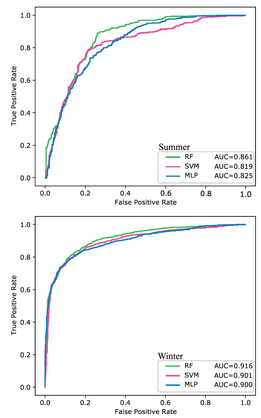
Performances of algorithms

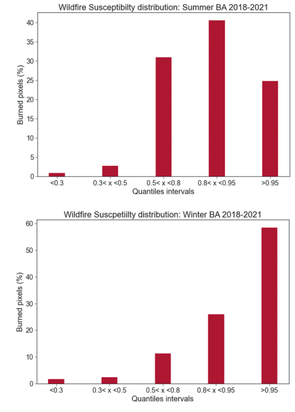


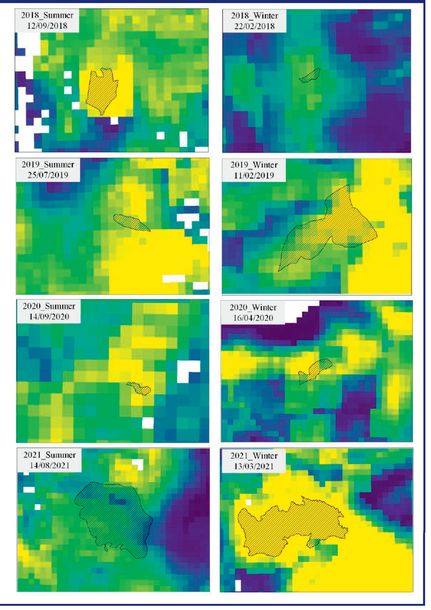
			RMSE	ROC AUC
Winter	SVM	Neighboring Vegetation	0.36	0.916
	MLP	Neighboring Vegetation	0.353	0.921
		Standard Model	0.342	0.939
	RF	Neighboring Vegetation	0.335	0.944
		Without Neighboring Vegetation	0.367	0.906
Summer	SVM	Neighboring Vegetation	0.358	0.931
	MLP	Neighboring Vegetation	0.344	0.94
		Standard Model	0.328	0.952
	RF	Neighboring Vegetation	0.329	0.953
		Without Neighboring Vegetation	0.37	0.911

Event-based validation

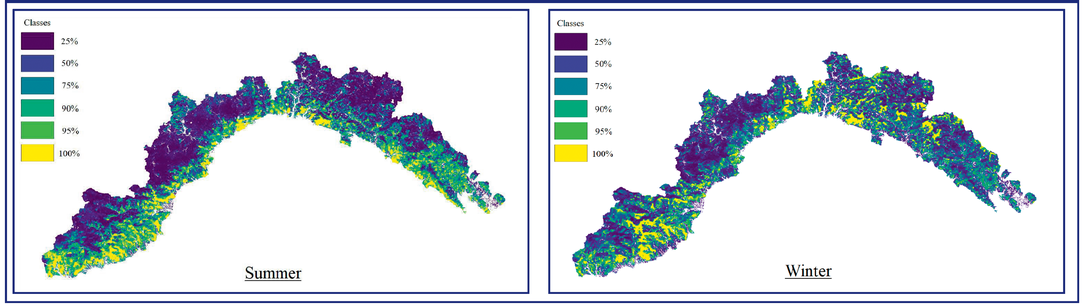
Analysis is done on the distribution of wildfire susceptibility across a number of significant wildfire incidents in the summer and winter in years from 2018 to 2021. Obviously, the more test burned pixels are in the higher 5th percentile of the susceptibility distribution, the more accurate the map is at predicting the locations most likely to experience a fire.



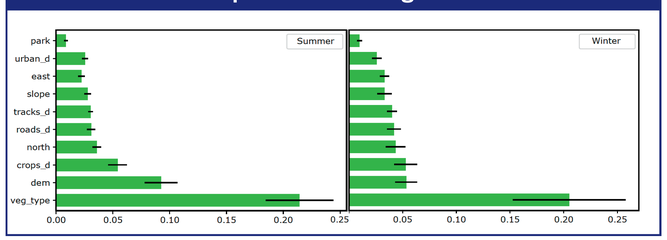




Susceptibility maps from RF algorithm with applying neighbouring vegetation (best performance)



Variable importance ranking for RF model



Conclusion

The results of the study indicated that Random Forest exhibited the best performance in predicting wildfire occurrences, closely followed by Multi-Layer Perceptron and Support Vector Machine. The successful application of machine learning showcased its ability to leverage general predisposing factors and explore interactions, overcoming potential limitations associated with broad classes of land use. The models proved to be robust and accurate in predicting future events, even in scenarios where wildfires were not part of the training process. Decision-makers in wildfire control and long-term land use planning will find the results of this study invaluable.

Overall, this study demonstrates the efficacy of machine learning techniques in understanding and predicting wildfire susceptibility, offering a valuable contribution to the field of environmental risk assessment and management.

Acknowledgement

I would like to sincerely thank my advisors, Professors Paolo Fiorucci and Andrea Trucchia, whose encouragement and genuineness I will never forget. Many thanks to all the members of Wildfire Risk Management and Forest Conservation Department of CIMA Research Foundation for their kind support during my thesis process. My sincere thanks also go to the department of Electrical, Electronics and Telecommunication Engineering and Naval Architecture Department (DITEN) of University of Genoa for organizing this practical and well-organized master course, especially Prof. Gabriele Moser and Prof. Giorgio Boni who inspired me during these two years. I am grateful to mention that part of this work has been published in Geosciences journal of MDPI (doi: 10.3390/geosciences12110424).