Geophysical Research Abstracts
Vol. 15, EGU2013-3238, 2013
EGU General Assembly 2013

Forest Fires in a Random Forest

Michael Leuenberger, Mikhaïl Kanevski, and Carmen D. Vega Orozco
CRET, FSGE, University of Lausanne, Switzerland (michael.leuenberger@unil.ch)

Forest fires in Canton Ticino (Switzerland) are very complex phenomena. Meteorological data can explain some occurrences of fires in time, but not necessarily in space. Using anthropogenic and geographical feature data with the random forest algorithm, this study tries to highlight factors that most influence the fire-ignition and to identify areas under risk.

The fundamental scientific problem considered in the present research deals with an application of random forest algorithms for the analysis and modeling of forest fires patterns in a high dimensional input feature space. This study is focused on the 2,224 anthropogenic forest fires among the 2,401 forest fire ignition points that have occurred in Canton Ticino from 1969 to 2008. Provided by the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), the database characterizes each fire by their location (x,y coordinates of the ignition point), start date, duration, burned area, and other information such as ignition cause and topographic features such as slope, aspect, altitude, etc. In addition, the database VECTOR25 from SwissTopo was used to extract information of the distances between fire ignition points and anthropogenic structures like buildings, road network, rail network, etc.

Developed by L. Breiman and A. Cutler, the Random Forests (RF) algorithm provides an ensemble of classification and regression trees. By a pseudo-random variable selection for each split node, this method grows a variety of decision trees that do not return the same results, and thus by a committee system, returns a value that has a better accuracy than other machine learning methods. This algorithm incorporates directly measurement of importance variable which is used to display factors affecting forest fires. Dealing with this parameter, several models can be fit, and thus, a prediction can be made throughout the validity domain of Canton Ticino.

Comprehensive RF analysis was carried out in order to 1) understand the importance of environmental features, 2) to assess the predictability of forest fires using environmental variables, and 3) to compare RF with other machine learning algorithms for this particular case study.

Risk maps were plotted by estimating the burned area taking into account the environmental variables and according to the most relevant selected variables.

Key words: Random Forest, classification and regression trees, forest fires, risk maps.

Acknowledgements
This work was partly supported by the SNFS Project No. 200021-140658, “Analysis and Modelling of Space-Time Patterns in Complex Regions”.

References
- Breiman L., 2001, Random Forest, Machine Learning, vol. 45-1: 5-32.
- Oliveira S., Oehler F., San-Miguel-Ayanz J., Camia A., Pereira J.M.C., 2012, Modeling spatial patterns of fire occurrence in Mediterranean Europe, Forest Ecology and Management, vol. 275: 117-129.