MODELLING FOREST FIRE OCCURENCE IN LEBANON USING SOCIO-ECONOMIC AND BIOPHYSICAL VARIABLES IN OBJECT-BASED IMAGE ANALYSIS

G. Mitri [a]*, E. Antoun [b], S. Saba [a], D. McWethy [c]

[a] Institute of the Environment, University of Balamand, Kelhat, El Koura, Lebanon - Email: george.mitri@balamand.edu.lb
[b] Directorate of Rural Development and Natural Resources, Ministry of Agriculture, Lebanon
[c] Department of Earth Sciences, Montana State University, Montana, USA

KEY WORDS: Mediterranean, fire, fire risk, wildland-urban interface, fire occurrence, socio-economic, biophysical, object-based

ABSTRACT:

Like many other countries in the Mediterranean, the occurrence and spread of forest fires in Lebanon are related to human activities. More specifically, landcover and land use changes (e.g., conversions of lands, abandonment of land and accumulation of fuel) driven by socio-economic changes occurring in the country have increased the probability of occurrence and spread especially in the Wildland-Urban Interface. The aim of this work was to model the influence of both socio-economic and biophysical variables on fire occurrence in Lebanon. The specific objectives were to 1) analyze socio-economic and biophysical drivers of forest fires, and 2) use object-based image analysis to derive a spatially explicit probability of fire occurrence across the country. Forward stepwise binary logistic regression analysis of 24 socio-economic and biophysical variables was used to predict wildfire occurrence. Spearman correlation analysis was conducted in order to eliminate multi-collinearity between selected variables. Eighty percent of the total number of administrative units was randomly selected for use in the development of the modelling, while the remaining 20% of units were used for testing and validating the final model. Object-based image analysis was used to map the spatial distribution of fire occurrence by modelling socio-economic and biophysical drivers including land-cover and land-use changes. The final map showed 5 different fire danger classes ranging from very low to very high. The quality of the classification results was evaluated and under- and overestimations errors of fire occurrence were mapped. The accuracy of the fire occurrence mapping model was approximately 85% when tested on the validation data set. The probabilistic spatial output of the fire threat model was considered satisfactory given the challenges of using multi-source data in an object-based image analysis approach. Results suggest increasing the resolution of socio-economic data would improve modelling accuracy of fire occurrence in Lebanon.

1. INTRODUCTION

Many Mediterranean ecosystems are experiencing increases in fire frequency, intensity and severity (Vilar et al., 2014) resulting in negative ecological, social and economic consequences (Westerling et al., 2006; FAO, 2007). The increase in forest fires in the Mediterranean are mostly related to human activities as more than 90% of fires of known cause are related to deliberate or accidental human ignitions (Ganteaume et al., 2013).

Population expansion into the Wildland-Urban Interface (WUI) is increasing the risk of natural fire occurrence and spread (Koulelis et al., 2010). More specifically, socio-economic and demographic changes (i.e., rural exodus, abandonment of arable land, and increase in ex-urban populations) are leading to accumulation in forest fuels and the close proximity of communities to these fuels (Pausas et al., 2009; FAO, 2007).

In Lebanon, negligence was reported as the main cause of fires (i.e., 14% of reported fire causes), followed by human activities on natural landscapes (i.e., 9% of reported fire causes) although the causes of ~67% of forest fires are not identified and reported (MOE/UOB, 2016). Landfill and agricultural practices represented 6% and 4%, respectively, of the total fire causes.

In this context, it is important to identify the factors which contribute to wildfire occurrence and integrate these factors into a mathematical expression in order to evaluate and quantify fire threat level (Sebastian-Lopez et al., 2008). More specifically, knowledge of the impact of socio-economic factors and their contribution to wildfire occurrence and proximity to populations is essential for improved evaluation and management of wildfire risk. In recent years, many studies (De Torres Curth et al., 2012; Zumbrunnen et al., 2012; Martinez et al., 2009) have shown that socio-economic variables and human factors have a considerable impact on wildfire regime.

In order to accurately predict fire risk, new analytical methods need to be employed to evaluate multi-scale biophysical and socio-economic data (Mitri et al., 2014a). One powerful analysis approach that is useful for analysing multi-dimensional datasets is GEographic Object-Based Image Analysis (GEOBIA). This approach addresses the development of automated methods of generating meaningful image-objects and assessing their characteristics (Blaschke, 2010).

The aim of this work was to model the influence of both socio-economic and biophysical factors on fire occurrence in Lebanon. The specific objectives were to 1) analyze socio-economic and biophysical drivers of forest fires, and 2) use

---

* Corresponding author

object-based image analysis to derive a spatially explicit probability of fire occurrence across the country.

2. STUDY AREA AND DATASET DESCRIPTION

2.1 Study area

The study area covers the country of Lebanon (with a total surface of 10,452 km^2) on the Eastern Mediterranean. It is divided into four distinct physiographic regions: the coastal plain, the Lebanon mountain range, the Beqaa Valley, and the Anti-Lebanon mountains (Mitri et al., 2015). Forests and other wooded land cover 24.5% of the Lebanese territory (MOE/UNDP/ECODIT, 2011).

The climate in Lebanon is characterized by dry summers extending from June to November (Salloum and Mitri, 2013) with average daytime temperatures above 30°C, and little rain (i.e., around 90% of the total annual precipitation falling between November and March).

Mitri et al. (2014b) found that approximately 89% of the WUI in Lebanon occurs in areas classified between moderate to very high risk of fire spread and 37% of the roads occur in areas classified between moderate to very high risk of fire spread.

Dataset description

The administrative map of Lebanon used for this study comprises all administrative units at the level of cadastral unit representing the approximate extent of a municipality/village. Biophysical data included slope, aspect, and vegetation fuel type (Mitri et al., 2015). The Prometheus fuel type classification system (Lasaponara et al., 2006) was used for characterizing fuel type (Table 1).

Fuel type	% coverage	Description
1	Ground fuels (cover > 50%)	Grass
2	Surface fuels (shrub cover > 60%; tree cover < 50%)	Grassland, shrubland (smaller than 0.3-0.6 m and with a high percentage of grassland), and clear-cuts, where slash was not removed.
3	Medium-height shrubs (shrub cover > 60%; tree cover < 50%)	Shrubs between 0.6 and 2.0 m
4	Tall shrubs (shrub cover > 60%; tree cover <50%)	High shrubs (between 2.0 and 4.0 m) and young trees resulting from natural regeneration or forestation.
5	Tree stands (>4 m) with a clean ground surface (shrub cover < 30%)	The ground fuel was removed either by prescribed burning or by mechanical means. This situation may also occur in closed canopies in which the lack of sunlight inhibits the growth of surface vegetation.
6	Tree stands (>4m) with medium surface fuels (shrub cover > 30%)	The base of the canopies is well above the surface fuel layer (>0.5 m). The fuel consists essentially of small shrubs, grass, litter, and duff.
7	Tree stands (> 4m) with heavy surface fuels (shrub cover >30%)	Stands with a very dense surface fuel layer and with a very small vertical gap to the canopy base (<0.5 m).

Table 1. The Prometheus fuel type classes

Spatial distribution of fruit trees, mixed culture, field crop, olive trees, vineyard, and greenhouses were extracted from the landcover/land use map of Lebanon produced in 1998. Conversion of lands (i.e. cropland, forest, and grassland to settlement) between 1998 and 2005 were also acquired (MOE/UNDP/GEF, 2015). Animal production was obtained from the homogeneous agricultural zones data of the Ministry of Agriculture produced in 1998.

In addition, population density as acquired from the Central Administration of Statistics (CAS), distance to big cities (i.e., 1,000 to 10,000 inhabitants), distance to bigger cities (i.e., above 10,000 inhabitants), the agricultural-urban interface, and roads in agricultural lands were also acquired (Mitri et al., 2014b). The annual extent of burned areas from 1998 throughout 2012 was collected from MOE/UNDP/GEF (2015).

The list of datasets employed in this study, and their corresponding unit and sources were listed in Table 2.

Description	Unit	Source/reference
Population density	Inhabitant/km^2	Central Administration of Statistics
Distance to big cities (1,000 to 10,000)	Km	Mitri et al, 2014b
Distance to big cities (> 10,000)	Km	Mitri et al, 2014b
Roads Density	Km/Km2	Mitri et al, 2014b
Agricultural Urban Interface Length	km	Mitri et al, 2014b
Roads in agricultural land	Km	Mitri et al, 2014b
Fruit trees	ha	1998 landuse/landcover map- MOE/UNDP/GEF, 2015
Mixed culture	ha	1998 landuse/landcover map- MOE/UNDP/GEF, 2015
Field Crop	ha	1998 landuse/landcover map- MOE/UNDP/GEF, 2015
Olive trees	ha	1998 landuse/landcover map- MOE/UNDP/GEF, 2015
Vineyard	ha	1998 landuse/landcover map- MOE/UNDP/GEF, 2015
Greenhouses	ha	1998 landuse/landcover map- MOE/UNDP/GEF, 2015
Animal production	Big Cattle Unit	Ministry of Agriculture
Conversion of cropland to settlement	ha	MOE/UNDP/GEF, 2015
Conversion of	ha	MOE/UNDP/GEF, 2015

forest to settlement		
Conversion of grassland to settlement	ha	MOE/UNDP/GEF, 2015
Slope	%	Digital Elevation Model
Aspect	degree	Digital Elevation Model
Fuel types (a total of 6 variables)	% coverage	Mitri et al. 2015

Table 2. Datasets employed as independent variables

3. METHODOLOGY

3.1 Logistic regression analysis

A number of studies used logistic regression (Verdú et al., 2012; Catry et al., 2010; Vilar et al., 2010; Andrews et al., 2003) to model fire occurrence. This approach best describes binary variables (i.e., presence or absence of burned areas) and is considered a more flexible and statistically robust approach to use than multiple regression (Cohen et al., 2013). In this study, Forward Stepwise binary logistic regression was used in order to identify a group of predictors affecting wildfire occurrence in Lebanon.

The presence/absence of burned areas in each cadastral unit was considered the dependent variable and was coded in a binary format (i.e., 0 for no burned areas and 1 for burned areas). The analysis was carried out on randomly selected cadastral units representing 80% of the total number of units. The remaining 20% of cadastral units were employed for validation of the final model. Before running the regression analysis, Spearman correlation analysis was applied (Artusi et al., 2002) to eliminate highly collinear variables (>0.60). Accordingly, the variable "distance to big cities (1,000 to 10,000 inhabitants" was removed. The final analysis was based on the following regression function:

$$f(z)= 1/(1+e\text{-}z) \qquad \text{Equation 1}$$

Where f is the probability of occurrence of an event and z is the linear function of the independent variables:

$$z= b_0+b_1x_1+b_2x_2+\dots+b_nx_n \qquad \text{Equation 2}$$

Stepwise forward regression was used to identify the contribution of a group of predictors on wildfire occurrence. Eleven distinct regression models were generated and a final best fit model was selected (Martinez et al., 2008). More specifically, the introduction of variables in the model was based on Wald significance (i.e., $p<0.05$) and their removal from the equation was applied when $p > 0.1$. The goodness of fit was estimated by the Nagelkerke's R2 test and the likelihood ratio -2LL was used to test the predictive capability of the model.

In order to predict dichotomous fire occurrence (yes = 1, no = 0) cadastral units with a value exceeding 0.5 was considered equal to 1 (i.e., high risk of fire occurrence). Otherwise, cadastral units were considered low risk of fire occurrence (Hosmer et al., 2013).

3.2 GEOBIA

In previous study we used GEOBIA to map vegetation fuel type (Mitri et al., 2015) and conversion of lands (MOE/UNDP/GEF, 2015). In this work, GEOBIA was further developed to integrate additional socio-economic and biophysical datasets using the software ecognition.

Image segmentation was applied using the cadastral units thematic map. As a result, each of the generated polygon/segment corresponds to a cadastral unit. A customized feature was built to integrate the previously final best fit model in the mapping analysis. The final output of the logistic regression model for each image object was classified into one of the five classes of fire threat taking into account the equally divided ranges.

3.3 Validation of the results

The binary classification of image object (i.e., cadastral units) was compared with known fire occurrence values of the original dataset in order to assess over- or under-estimation of fire occurrence. Underestimations include those cadastral units where the model predicted low fire risk of occurrence and at least one known fire occurred, while overestimation include cadastral units where the model predicted high fire occurrence and no known fires occurred.

4. RESULTS AND DISCUSSION

The final logistic regression model included a total of 10 significant variables ($p<0.05$) for predicting fire occurrence:

$Z = - 2.078 + 0.045x_1 + 0.181x_2 + 0.347x_3 + 0.040x_4 + 0.014x_5 + 0.094x_6 + 0.03x_7 + 0.042x_8 + 0.202x_9 + 0.004x_{10}$

Where
x_1: slope – x_2: area of grassland conversion to settlement (1998-2005) – x_3: area of forest conversion to settlement (1998-2005) – x_4: area of fuel type 6 and 7 – x_5: area of fuel type 4 – x_6: area of fuel type 3– x_7: area of fuel type 2– x_8: area of fuel type 1 – x_9: length of Agricultural-Urban Interface – x_{10}: area of fruit trees cultivation.

The goodness of fit was estimated by the Nagelkerke's R2 test (i.e., 0.469) and the likelihood ratio -2LL (i.e., 943.815) to test the predictive capability of the model.

A 2x2 classification table (Table 3) of observed and predicted fire occurrence was developed.

Observed		Predicted					
		Model building data			Validation data		
		Binary		%	Binary		%
		Low	High		Low	High	
Binary	0	190	128	59.7	39	21	65
	1	74	872	92.2	19	200	91.3
Overall %				84.0			85.7

Table 3. Comparison of predicted versus observed fire occurrence

Validation results showed that the model correctly predicted an average of 85.7% fire occurrences or non-occurrences. More

specifically, 91.3% of fire occurrences were correctly predicted, while 65% of non-fire occurrences were correctly predicted.
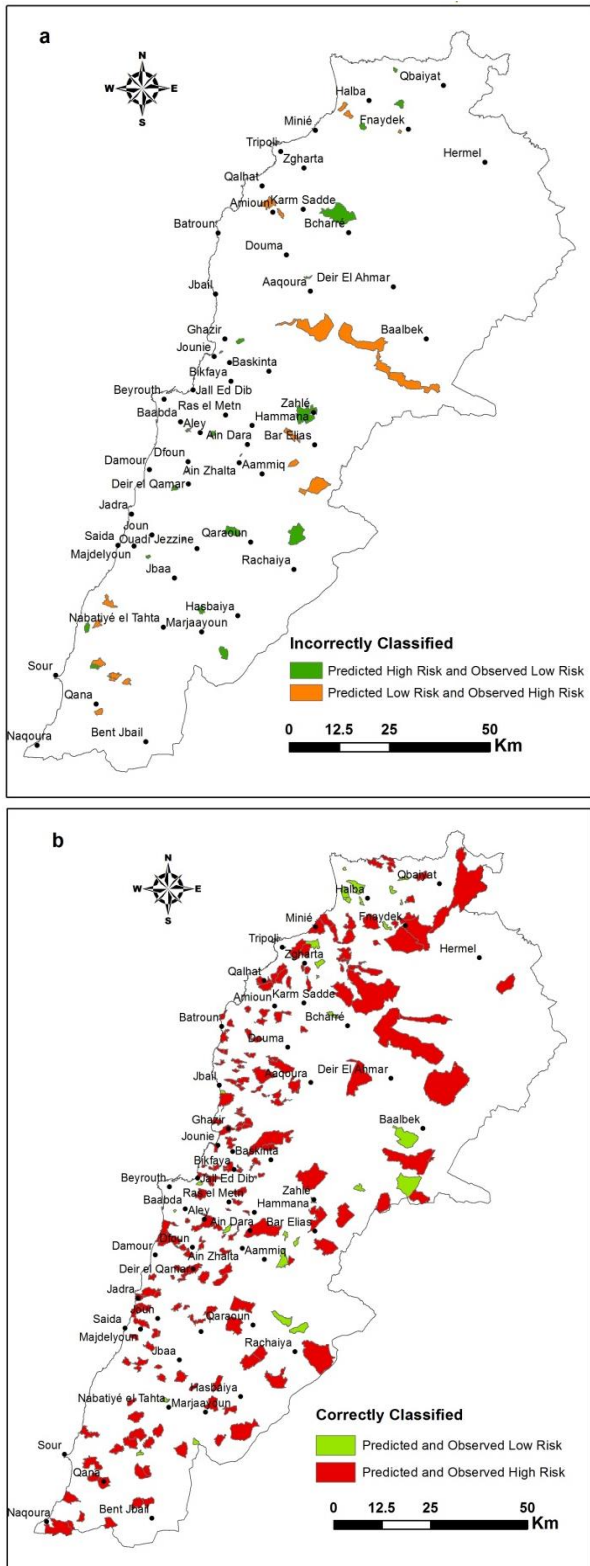

Figure 1. Cadastral units in which predicted fire occurrence did not match observed data (upper) and in which matched observed data (lower)

Figure 2 displays the spatial distribution of the model's final outcome at the level of the cadastral units. The range of estimated fire occurrence probabilities was divided into 5 equal-size classes. The majority of the cadastral units were characterized with a very high risk of fire occurrence. Cadastral units with very low to moderate risk of fire occurrence were mainly located in relatively highly urbanized or agricultural areas such as in the coastal zone and the central part of the Bekaa valley (i.e., inland).
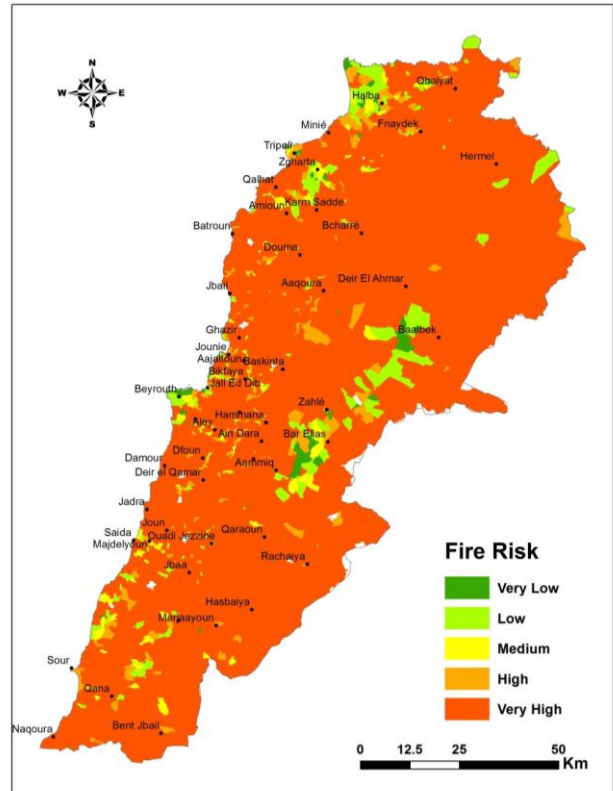

Figure 2. Spatial distribution of the logistic regression model outcome

The influence of the significant socio-economic variables in the model was identified through a sensitivity analysis. All the 10 variables were ranked according to a final score which was calculated by adding the ranks of the five evaluation criteria for significance (Wald test, standard coefficient (B), -2LL, Exp (B) and regression coefficient). The level of importance of the variables in the model increases with rank value (Martinez-Fernandez et al., 2013). Results showed that conversion of forest to settlement was considered as the most significant socio-economic variable increasing fire occurrence followed by the length (km) of Agricultural-Urban Interface, conversion of grassland to settlement, and cultivation of fruit trees.

The final logistic regression model was evaluated in GEOBIA for analysis of multi-source socio-economic and biophysical spatial data. This allowed the assessment of fire occurrence at the level of individual polygons (i.e., cadastral units). GEOBIA was previously conducted to map landcover/landuse changes and vegetation fuel type with the use of satellite imagery (Mitri et al., 2015). It was used here to further integrate socio-economic data to predict fire risk at the cadastral level. This approach showed that GEOBIA can serve as a tool for integrated analysis and assessment of fire occurrence for multiple purposes and scales.

5. CONCLUSIONS

In this study, forward stepwise binary logistic regression analysis was conducted to assess fire occurrence using 24 socio economic and biophysical variables. A total of 10 socio-economic and biophysical variables were finally used in the best performing predictive model. The resulting final logistic regression model was evaluated in GEOBIA to derive a spatially explicit probability of fire occurrence across the country.

Model validation showed an average accuracy of 85.7% in predicting fire occurrence. Socio-economic variables including conversion of forest to settlement, length of Agricultural-Urban Interface, conversion of grassland to settlement, and the cultivation of fruit trees were found to significantly contribute to increased probability of fire occurrence. The primary biophysical variables contributing to an increased probability of fire occurrence included fuel type and slope conditions.

Future work focusing on integrating higher-resolution socio-economic datasets at the cadastral unit will help improve the accuracy of spatially-explicit mapping of fire occurrence across the Lebanon. GEOBIA serves as an essential tool for integrating multiple high-level datasets to predict the spatial probability of fire occurrence.

ACKNOWLEDGEMENTS

REFERENCES

Andrews, P. L., Loftsgaarden, D. O., and Bradshaw, L. S. 2003. Evaluation of fire danger rating indexes using logistic regression and percentile analysis. *International Journal of Wildland Fire*, 12(2), 213-226.

Artusi, R., Verderio, P., & Marubini, E. 2002. Bravais-Pearson and Spearman correlation coefficients: meaning, test of hypothesis and confidence interval. *International Journal of Biological Marker*s, 17(2), 148-151.

Blaschke, T., 2010. Object based image analysis for remote sensing. *International Journal of Photogrammetry and Remote Sensing*, 65, pp. 2-16.

Catry, F. X., Rego, F. C., Bação, F. L., & Moreira, F. 2010. Modeling and mapping wildfire ignition risk in Portugal. *International Journal of Wildland Fire*, 18(8), 921-931.

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. 2013. Applied multiple regression/correlation analysis for the behavioral sciences. Routledge.

De Torres Curth, M., Biscayart, C., Ghermandi, L., and Pfister, G. 2012. Wildland-Urban Interface fires and socioeconomic conditions: a case study of a Northwestern Patagonia City. *Environmental Management* 49(4): 876-891.

FAO 2007. Fire management global assessment. A thematic study prepared in the framework of the Global Forest Resources Assessment 2005. FAO Forestry Paper 151. Food and Agriculture Organization for the United Nations: Rome

Available at http://www.fao.org/forestry/fra2005/en [Verified on 24-5-2016].

Ganteaume, A., Camia, A., Jappiot, M., San-Miguel-Ayanz, J., Long-Fournel, M., Lampin, C. 2013. A Review of the Main Driving Factors of Forest Fire Ignition Over Europe. *Environmental Management*, 51(3), 651-662.

Hosmer Jr, D. W., Lemeshow, S., and Sturdivant, R. X. 2013. *Applied logistic regression*. John Wiley & Sons.

Koulelis, P., and Mitsopoulos, I. 2010. A Study on the Socioeconomic Factors Influencing Wildfire Occurrence in Mediterranean Basin Countries. In. Proceedings of the Third International Symposium on Fire Economics, Planning, and Policy: Common Approaches and Problems (Vol. 227, p. 314). DIANE Publishing.

Lasaponara, R., A. Lanorte, and S. Pignatti, 2006. Characterization and Mapping of Fuel Types for the Mediterra-nean Ecosystems of Pollino National Park in Southern Italy by Using Hyperspectral MIVIS Data. *Earth Interactions*, 10(13): 1-11.

Martínez-Fernández, J., Chuvieco, E., &Koutsias, N. 2013.Modelling long-term fire occurrence factors in Spain by accounting for local variations with geographically weighted regression. *Natural Hazards and Earth System Science*, 13(2), 311-327.

Martínez, J., Vega-Garcia, C., and Chuvieco, E. 2009. Human-caused wildfire risk rating for prevention planning in Spain. *Journal of Environmental Management*, 90(2), 1241-1252.

Martínez, J., Chuvieco, E., & Martín, P. 2008. Estimation of risk factors of human ignition of fires in Spain by means of logistic regression. In Proceedings of the Second International Symposium on Fire Economics, Planning, and Policy: A Global View, Gen. Tech. Rep. PSW-GTR-208, pp. 265-278.

Mitri, G., Jazi, M., McWethy, D. 2015. Assessment of wildfire risk in Lebanon using Geographic Object-based image analysis. *Photogrammetric Engineering & Remote Sensing*. Vol. 81, No. 6, June 2015, pp. 499–506.

Mitri, G., Nader, M., Van der Molen, I. and Lovett, J. 2014a. Evaluating exposure to land degradation in association with repetitive armed conflicts in North-Lebanon using multi-temporal satellite data. *Environmental Monitoring and Assessment*. Volume 186, Issue 11, Page 7655-7672. http://dx.doi.org/10.1007/s10661-014-3957-5.

Mitri, G., Jazi, M., Antoun, E., and McWethy, D. 2014b. *Managing wildfire risk in Lebanon*. University of Balamand. Kelhat, El Koura, Lebanon.

MoE/UNDP/GEF (2015). National Greenhouse Gas Inventory Report and Mitigation Analysis for the Land Use, Land-Use Change and Forestry Sector in Lebanon. Beirut, Lebanon.

MOE/UNDP/ECODIT, 2011. State and Trends of the Lebanese Environment 2010, Ministry of Environment, Beirut, Lebanon.

MOE/UOB. 2016. State of Lebanon's wildfire in 2015. A joint report of the Ministry of Environment and the University of Balamand, Beirut, Lebanon.

Pausas, J. G., Llovet, J., Rodrigo, A., & Vallejo, R. (2009). Are wildfires a disaster in the Mediterranean basin? A review. *International Journal of Wildland Fire*, 17(6), 713-723.

Salloum, L. and Mitri, G. 2014. Assessing the temporal pattern of fire activity and weather variability in Lebanon. International Journal of Wildland Fire. Volume 23(4): 503-509. http://dx.doi.org/10.1071/WF12101.

Sebastián-López, A., Salvador-Civil, R., Gonzalo-Jiménez, J., & San Miguel-Ayanz, J. 2008. Integration of socio-economic and environmental variables for modelling long-term fire danger in Southern Europe. *European Journal of Forest Research*, 127(2), 149-163.

Verdú, F., Salas, J., & Vega-García, C. 2012. A multivariate analysis of biophysical factors and forest fires in Spain, 1991–2005. *International Journal of Wildland Fire*, 21(5), 498-509.

Vilar, L. Camia, A., San-Miguel-Ayanz, J. 2014. Modelling socio-economic drivers of forest fires in the Mediterranean Europe. Advances in forest fire research (Viegas, D.). Pp 1874-1882.

Vilar, L., Nieto, H., & Martín, M. P. 2010. Integration of lightning-and human-caused wildfire occurrence models. *Human and Ecological Risk Assessment: an International Journal*, 16(2), 340-364.

Westerling A., Hidalgo H., Cayan D., Swetnam T. 2006. Warming and earlier spring increase western US forest wildfire activity. *Science* 313, 940-943.

Zumbrunnen, T., Menéndez, P., Bugmann, H., Conedera, M., Gimmi, U., and Bürgi, M. 2012. Human impacts on fire occurrence: a case study of hundred years of forest fires in a dry alpine valley in Switzerland. *Regional Environmental Change*, 12(4), 935-949.