Providing Forest Fire Risk Map Using Multivariate Aduptive Regression Spline (Case Studey: Golestan Province)

Ali Shah-Heydaripour[1], Parham Pahlavani[2*], Behnaz Bigdeli[3]

1. GIS M.Sc. Student at School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran
2. Assistant Professor at School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran
3. Assistant Professor at School of Civil Engineering, Shahrood University of Technology, Shahrood, Iran

(Resaved 23 July 2018 – Accept 20 October 2018)

Abstract

Forest areas are among the most important natural and ecological resources on the Earth and are considered as one of the main pillars of sustainable development in any country. Fires ruins almost 5500 hectares of Iran's forests yearly. In this research, firstly, the fire points were identified using the fire data of Forest Organization in combination with MODIS sensor data between 2012 and 2017. Due to the fact that more than 75% of fires were happened in the hot season of the year (June, July, and August), the data of the three months was used for modeling. Then, the effective parameters in fire occurring were evaluated and the dependent parameters were removed. Accordingly, two methods, including multiple linear regression and multivariate adaptive regression spline were studied to predict the fire risk. Some important parameters including the root-mean-square error (RMSE), R^2, the correct estimation percentage of fire and non-fire points, and error distribution were used to evaluate. After modeling, it was found that the multivariate adaptive regression spline has better performance—where its RMSE of test data was 0.1628, its R^2 of test data was 0.893, and its correct estimation percentage of test fire points and test non-fire points was near 94% and 88% respectively, as well as its error distribution was better than the other method. This actually shows that modeling with a local method is very better than modeling with a global method. Therefore, the risk map resulted by multivariate adaptive regression spline has better reliability compared to those of the other method. Finally, the high-risk areas were recognized using the risk map of this method. The traits of these areas were a short distance to residential areas and roads, having rich soil with organic materials, relatively high temperature, and low height.

* Corresponding Author, Email: pahlavani@ut.ac.ir

Introduction

In 2000, a convention was established in the United Nations to improve the quality of human life in which the principles of the Millennium Development Goals were adopted. One of these goals was to ensure the stability of the environment and natural resources. In the contemporary world, the value of forests is about 120 billion dollars and the livelihood of almost 9.1 people is dependent on forest (in)directly.

According to the opinion of global experts including FAO, if the forest cover of a country is less than 25% of that country's area, that country is in critical condition in terms of the human environment. Almost 190000 hectares of Iranian forests have been ruined by fire in a 28-year period. Forest fire not only changes the natural ecosystem and ruins many plant and animal species of a region, but also makes other destructive effects like air pollution, respiratory problems, soil erosion, increased flowing surface waters, increased acidity of soil, decreased fertility, tourism industry losses, manufacturing industry and economy losses, and even climate change.

Immediate and accurate detection of the fire location and the ability to determine the effective parameters on it, as well as the detection of the areas with high-risk of fire is among the main concerns of environmental protection and disaster management. We can prevent the fire by training people, making effective regulations and management policies, and increased monitoring to deal with fire triggers. Moreover, in the case of fire occurrence, we must take necessary actions like deploying fire-fighting equipment near hazardous areas and making easy access to these areas. In fact, nowadays, the increasing importance of protecting the forests and natural resources has led to change the focus from crisis management to risk management.

Methodology

The modeling was not possible without non-fire points. Accordingly, at the beginning, some points are randomly selected in the whole area with a certain distance from the fire points and are identified as non-fire points. To implement the methods in MATLAB programming environment, firstly, the parameters used in the modeling are extracted using the maps of these parameters for fire and non-fire points. These parameters are used as inputs in each of these methods.

Constantly, 70% of the selected data were used as the training data and 30% of them were used as the test data. Initially, the multivariate linear regression and then the multivariate adaptive regression spline were used for modeling. The steps of the research implementation are shown in Figure (1).

After implementation of the modeling, the evaluation parameters of each method were provided to compare. Then, the risk map of the area was provided using trial points and Inverse Distance Weighting (IDW) and by employing 12 lateral points for each method (Figures 2 and 3). The points with a high risk were extracted from the resulted map. Then, the main traits of these points are considered as the traits of high-risk points.

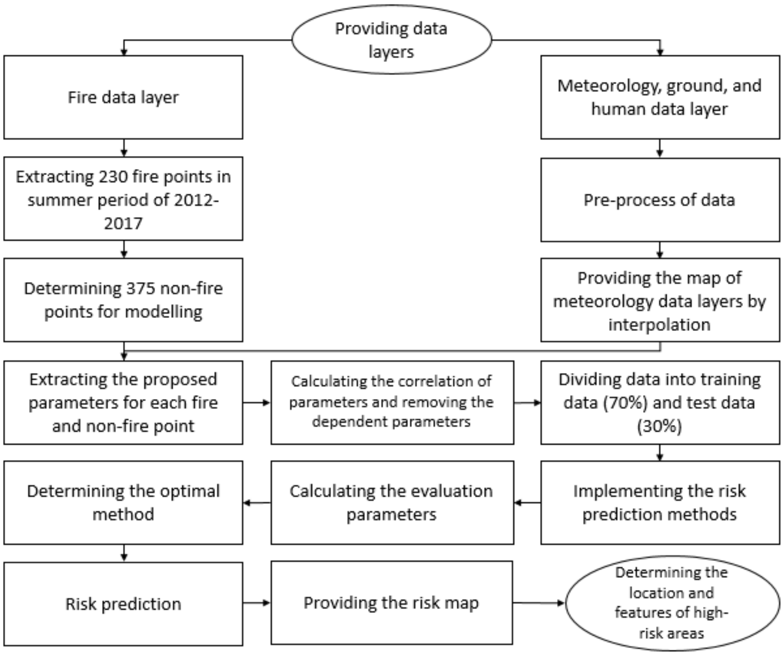

Fig. 1. The steps of the research implementation

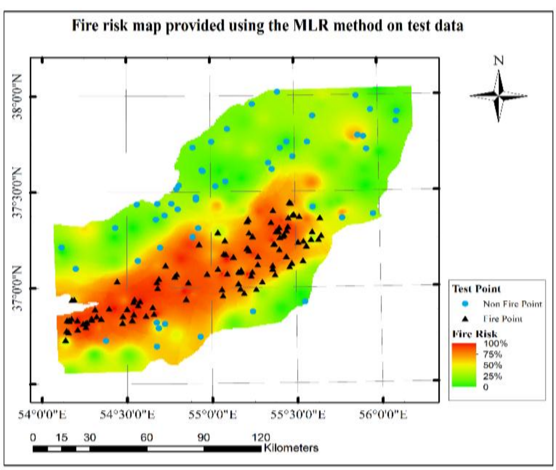

Fig. 2. Fire risk map provided using the MLR method on test data

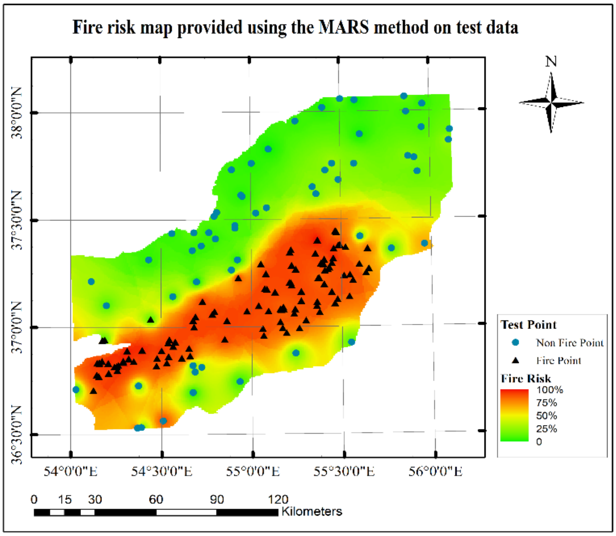

Fig. 3. Fire risk map provided using the MARS method on test data

Discussion and Results

After removing the dependent parameters from the effective parameters on the fire, the optimal effective parameters are presented in Table (1). These parameters are divided into three groups including climate, ground physical, and human parameters.

The modeling of fire risk was done by two methods. In the training and testing data section, the RMSE and R^2 are presented in Table (2) for multivariate adaptive regression spline and multivariate linear regression methods, respectively. The results achieved by the training data section indicate that the training procedure is more accurate (R^2 closer to 1) and with less error (less RMSE) in the multivariate adaptive regression spline than those achieved by the multivariate linear regression method. The appropriate amount of evaluation parameters for test data shows that the model does not experience over-fitting in these methods.

Table 1. Effective parameters on fire occurrence in the case-study area

Climate parameters	Ground's physical parameters	Human parameters
Average temperature (C°)	Soil type	Distance from the residential areas (km)
Rainfall (mm)	Height (m)	Distance from the road (km)
Average wind speed (km/h)	Distance from the river (km)	
	Steep direction	

Table 2. Evaluation parameters of risk modeling methods

	Data/parameter	R^2	**RMSE**	**The correct estimation percentage of nom-fire points**	**The correct estimation percentage of fire points**
MLR	Training data	0.6728	0.2846	42%	74%
	Test data	0.5877	0.3180		
MARS	Training data	0.8932	0.1628	88%	94%
	Test data	0.8211	0.2078		

In the linear regression method, the two parameters of the correct estimation percentage of fire points and non-fire points have a low value, hence, the worst possible scenario has happened and the risk map has the least amount of reliability. In the multivariate adaptive regression spline, the fire and non-fire points are simultaneously estimated with a high accuracy. This makes the risk map provided by the multivariate adaptive regression method becomes to be more reliable.

As seen in the results, the risk map provided by the multivariate adaptive regression spline method has a very higher reliability compared to the risk map provided by multivariate linear regression method. Hence, the risk map resulted by the first method was used to determine the features of the areas with a high risk of fire (Figure 4).

Since the fire risk has a normal distribution, the areas which satisfy Equation (1) are among the 2.5% of the areas that have the most fire risk.

$$\mu + 2\sigma \leq R \tag{1}$$

where μ is the average, σ is the standard deviation, and R is the fire risk. The main features of the mentioned areas can be used as the important tools for decision making. The extraction of high-risk areas is done in ArcGIS environment. Statistical analysis of effective parameters' features in these areas shows some key points. These features include low distance from the residential regions (less than 2 km), low distance from the road (less than 2 km), having mollisol, relatively high average temperature (more than $28\ C^{\circ}$), and low height (less than 50 m).

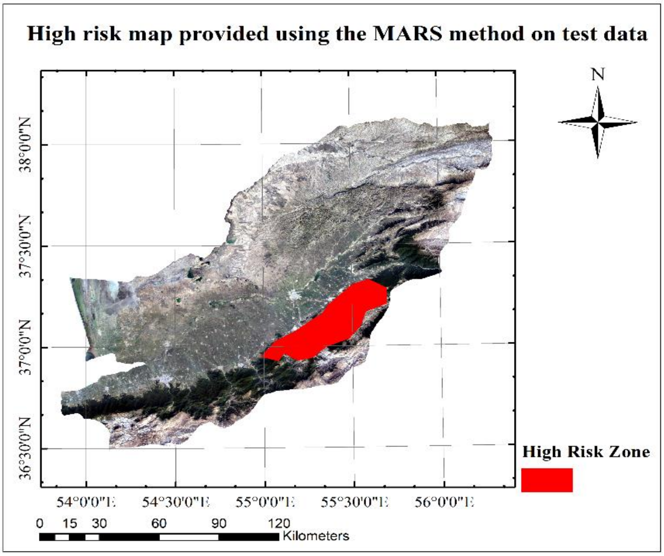

Fig. 4. High risk map provided using the MARS method on test data

Conclusions

This research attempted to identify the optimal method for modeling of fire points risk using climate, ground physical, and human parameters. Therefore, an accurate local method (MARS) was used along with a non-local method (MLR).

In the test data and the training data sections, the MARS method had the lowest RMSE and a value closer to 1. The outputs showed that the MARS method had a more accurate performance in the estimation of the fire and non-fire points compared to the MLR method. This indicated the high reliability of the MARS method. After determining the optimal method for the modeling of the area's fire occurrence, the points of the area with high risk of fire were detected. After doing a statistical analysis it was found that these points have some fundamental features including low distance from the residential regions (less than 2 km), low distance from the road (less than 2 km), having mollisol, relatively high average temperature (more than 28 C°) and low height (less than 50 m).

Keywords: Forest Fire, Multiple Linear Regression, Multivariate Adaptive Regression Spline, Risk Map.

References

[1].López-Mondéjar, R., Brabcová, V., Štursová, M., Davidová, A., Jansa, J., Cajthaml, T., et al. (2018). "Decomposer food web in a deciduous forest shows high share of generalist microorganisms and importance of microbial biomass recycling," The ISME journal, p. 1.

[2].FAO (2010). "Global forest resources assessment," Main report, FAO Forest paper 163.
[3].Molders, N., and Prestemon J.P. (2007). "Comparing production function model for wild fire risk analysis in the wildland-urban interface".Forest policy and economics, 7(5):P. 782-795 [3]
[4].Chuvieco, E. and Congalton, R.G.. (1989). "Application of remote sensing and geographic information systems to forest fire hazard mapping". Remote Sensing of Environment. Vol 29: pp 147–159. [4]
[5].Li, X., Zhao, G., Yu, X., and Yu, Q. (2014). "A comparison of forest fire indices for predicting fire risk in contrasting climates in China," Natural hazards, vol. 70, pp. 1339-1356. [5]
[6].Jafari Goldarag, Y., Mohammadzadeh, A. (2016). "Fire Risk Assessment Using Neural Network and Logistic Regression". Journal of the Indian Society of Remote Sensing. Volume 44, Issue 6, pp 885–894. [6]
[7].Coelho Eugenio, F. , Rosa dos Santos, A. (2016). "Applying GIS to develop a model for forest fire risk: A case study in Espírito Santo", Brazil. Journal of Environmental Management. Vol 173 , pp 65-71.[7]
[8].Z. S. Pourtaghi, H. R. Pourghasemi, and M. Rossi (2015). "Forest fire susceptibility mapping in the Minudasht forests, Golestan province, Iran," Environmental Earth Sciences, vol. 73, pp. 1515-1533. [8]
[9].R. Ajin, A.-M. Loghin, P. Vinod, and M. K. Jacob (2016). "Forest fire risk zone mapping in Chinnar Wildlife Sanctuary, Kerala, India: A study using geospatial tools," Journal of Global Resources, vol. 3, pp. 16-26. [9]
[10]. P. Y. Bernier, S. Gauthier, P.-O. Jean, F. Manka, Y. Boulanger, A. Beaudoin, et al. (2016). "Mapping local effects of forest properties on fire risk across Canada," Forests, vol. 7, p. 157. [10]
[11]. K. Suryabhagavan, M. Alemu, and M. Balakrishnan (2016). "GIS-based multi-criteria decision analysis for forest fire susceptibility mapping: a case study in Harenna forest, southwestern Ethiopia," Tropical Ecology, vol. 57, pp. 33-43. [11]
[12]. M. Luckose, P. Arunkumar, A. Gopi, and J. C. Mathew (2017). "Forest fire hazard zonation mapping of Wayanad district of India using geospatial technology," in Disaster, Risk and Vulnerability Conference 2017, p. 91. [12]
[13]. Rodriguez, T. D. A., Ramirez, M. H. and Tchikoue, J. H., (2008). "Factors affecting the accident rate of forest fires‖", Ciencia Forestal en Mexico, Vol.33, No.104, PP. 38–57. [13]
[14]. Romero-Calcerrada, R., Novillo, C. J., Millington, J. D. A., (2008). "GIS analysis of spatial patterns of human-caused wildfire ignition risk in the SW of Madrid (Central Spain)", Landscape Ecol. Vol.23 PP. 341–354.[14]
[15]. Avila, F. D. A., Pompa-Garcia, M., Antonio-Nemiga, X., (2010). "Driving Factors for Forest Fire Occurrence in Durango State of Mexico: A Geospatial Perspective‖", Chin. Geogra. Sci. Vol.20, No.6, PP. 491–497. [15]
[16]. Raei, A. , Pahlavani, P. , Hasanlou, M. , (2016). "Determining Effective Factors on Forest Fire Using the Compound of Geographically Weighted

Regression and Genetic Algorithm, a Case Study: Golestan, Iran”. Iran. Journal of Geospatial Information Technology. Vol 3, Issue 4, pp 97-120. [16]

[17]. Srivas, T., Artés, T., de Callafon, R. A., and Altintas, I. (2016). "Wildfire Spread Prediction and Assimilation for FARSITE Using Ensemble Kalman Filtering," Procedia Computer Science, vol. 80, pp. 897-908. [17]

[18]. Berger, P. D., Maurer, R. E., and Celli, G. B. (2018). "Multiple Linear Regression," in Experimental Design, ed: Springer, pp. 505-532.

[19]. Friedman, J. H. (1991). "Multivariate adaptive regression splines," The annals of statistics, pp. 1-67. [19]

[20]. Knafl, G. J. and Ding, K. (2016). Adaptive regression for modeling nonlinear relationships: Springer. [20]

[21]. Vidyullatha, P. and Rao, D. R. (2016). "Machine Learning Techniques on Multidimensional Curve Fitting Data Based on R-Square and Chi-Square Methods," International Journal of Electrical and Computer Engineering, vol. 6, p. 974. [21]

[22]. Chai, T. and Draxler, R. R. (2014) "Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature," Geoscientific model development, vol. 7, pp. 1247-1250. [22]