Nat Hazards
DOI 10.1007/s11069-008-9326-3

ORIGINAL PAPER

Identifying wildland fire ignition factors through sensitivity analysis of a neural network

Christos Vasilakos · Kostas Kalabokidis · John Hatzopoulos · Ioannis Matsinos

Received: 14 February 2008 / Accepted: 17 November 2008

Abstract Artificial neural networks (ANNs) show a significant ability to discover patterns in data that are too obscure to go through standard statistical methods. Data of natural phenomena usually exhibit significantly unpredictable non-linearity, but the robust behavior of a neural network makes it perfectly adaptable to environmental models such as a wildland fire danger rating system. These systems have been adopted by many developed countries that have invested in wildland fire prevention, and thus civil protection agencies are able to identify areas with high probabilities of fire ignition and resort to necessary actions. Since one of the drawbacks of ANNs is the interpretation of the final model in terms of the importance of variables, this article presents the results of sensitivity analysis performed in a back-propagation neural network (BPN) to distinguish the influence of each variable in a fire ignition risk scheme developed for Lesvos Island in Greece. Four different methods were utilized to evaluate the three fire danger indices developed within the above scheme; three of the methods are based on network's weights after the training procedure (i.e., the percentage of influence—PI, the weight product—WP, and the partial derivatives—PD methods), and one is based on the logistic regression (LR) model between BPN inputs and observed outputs. Results showed that the occurrence of rainfall, the 10-h fuel moisture content, and the month of the year parameter are the most significant variables of the Fire Weather, Fire Hazard, and Fire Risk Indices, respectively. Relative humidity, elevation, and day of the week have a small contribution to fire ignitions in the study area. The PD method showed the best performance in ranking variables' importance, while performance of the rest of the methods was influenced by the number of input parameters and the magnitude of their importance. The results can be used by local forest managers and other decision makers dealing with wildland fires to take the appropriate preventive measures by emphasizing on the important factors of fire occurrence.

C. Vasilakos (✉) · J. Hatzopoulos · I. Matsinos
Department of Environmental Studies, University of the Aegean, University Hill, 81100 Mytilene, Greece
e-mail: chvas@env.aegean.gr

K. Kalabokidis
Department of Geography, University of the Aegean, University Hill, 81100 Mytilene, Greece

Published online: 05 December 2008

Keywords Greece · Neural networks · Partial derivatives · Risk assessment · Sensitivity analysis · Wildfire occurrence

1 Introduction

Forest fires are complicated events that take place as a result not only of natural processes but also due to human factors. Wildland fire danger estimation systems have been developed by many countries meeting fire problems to manage these complicated phenomena (Viegas et al. 1999). The knowledge acquired through these systems enables civil protection agencies to define and identify high-level risk areas and plan the necessary preventive and control actions (Deeming et al. 1977; Hoffmann et al. 1999; Taylor and Alexander 2006; Van Wagner 1987). The majority of these systems adopt different approaches with regard to their spatial and temporal resolution (see, e.g., Yuan 1997), as well as the correlation of their input parameters (Viegas et al. 1999). Spatial data that are taken into consideration in identifying wildland fire pattern occurrence are often collected, managed, analyzed, and presented with Geographic Information Systems—GIS (e.g., Chou 1992a; Chuvieco and Congalton 1989; Preisler et al. 2004).

Wildland fire danger evaluation is an integration of weather, topography, vegetative fuel, and socioeconomic input variables to produce numeric indices of fire potential outputs (Andrews et al. 2003; Pyne et al. 1996). Various quantitative methods have been explored for the correlation of the input variables in fire danger assessment; most of these methods include the input variable's importance as a direct or indirect output. For example, some fire studies are based on the selective weighting that relies on expert opinion regarding input's importance (see Salas and Chuvieco 1994), and other studies are based on multi-criteria analysis where experts proceed to a pair-wise comparison among the inputs using a relative importance scale (Alcázar et al. 1998). Alternatively, to subjective weighting of inputs, statistical methods were widely used for obtaining the expression of fire danger calculation. More specifically, linear and logistic regression techniques were proposed, so the coefficients of the models reflect the influence of inputs on fire danger (Chou 1992b; Chou et al. 1993; Kalabokidis et al. 2007; Vasconcelos et al. 2001).

Artificial neural networks (ANNs) have been also used in fire ignition risk estimation (Chuvieco et al. 1999b; Vasconcelos et al. 2001; Vasilakos et al. 2007; Vega-García et al. 1996). The ANNs are particularly useful for pattern recognition and for modeling complex problems for which the explicit form of the relationships among certain variables is not known (Fausett 1994). The development of ANNs began over 50 years ago as a result of scientists' attempt to better comprehend the human brain and simulate some of its abilities. Traditional applications of ANNs include classification, noise reduction, and prediction (Masters 1993). An ANN is characterized by its architecture, its training algorithm, and its activation function. In general, an ANN consists of processing elements called neurons or nodes organized in layers. Every node is connected to other nodes to form a network, and every connection link has an associated weight. These weights represent the information that is stored in an ANN during the training process. The nodes receive inputs from other nodes or model's input values. Then, the output of each node is computed as a function—so-called activation function—of the weighted sum of its inputs according to Eq. 1:

$$y_i = f\left(\sum_j w_{ij} y_j\right) \tag{1}$$

where y_i is the output value of the node, w_{ji} is the connection weight from node j to node i, and y_j is the output value of node j. During training, the ANN initiates the learning process through the random values of its weights and by using a set of input and output values. The computed output is then compared with the actual output value and the weights are corrected so as to minimize the error function. The same process is repeated many times so that the error is gradually diminished until it becomes small and tolerable. To ensure that the trained network will approximate target values that are not included in the training data set, a validation dataset is used which includes cases that are not used in training. At the same time, at least one validation dataset is used and the training is stopped when the error starts to increase in the validation dataset, despite the fact that the error could still be decreasing in the training set. This is an indication that the network has a good generalization and an over-fitting to the training dataset has been avoided (Bishop 1995).

One of the disadvantages of the ANNs is the interpretation of the output model to find out the most important variables that affect the model's output. Contrary to the statistical methods, e.g., regression, where the influence of each independent variable is expressed by the estimated coefficients of the model, ANN's output model cannot be explained directly. Hence, ANNs have been characterized as a "black box" (Andersson et al. 2000; Howes and Crook 1999; Olden and Jackson 2002; Papadokonstantakis et al. 2006). Previous research on wildfire risk estimation with ANNs is limited to predictability results of the output model or to utilizing methods other than the methods presented and compared here. An operational fire ignition risk system might predict where and when a fire may start, but why a fire starts at a specific place and time is still not predictable, especially in a complex human–natural environment. For example, negligence or arson causative factors require potential human presence; and factors of human presence include urban areas, road network, and agricultural fields among others. The way that these factors influence wildfire ignition at a specific place is barely answered by prediction models.

The aim of this research was the evaluation of the significance of input variables in a Fire Ignition Index (FII) subsystem by comparing four different methods, as part of a fire danger rating system analyzed by Vasilakos et al. (2007). The outputs of these methodologies could be useful in quantitative fire danger rating of the study area, because forest fire management may be focused on preventive measures aiming at decreasing fire ignitions, such as proper legislation regarding property land development and law enforcement, personnel and volunteers' training, technological engineering, as well as staffing for fire control.

Several methods have been applied to examine the input's influence in a trained ANN. Howes and Crook (1999) proposed, among others, the "general influence" method of each input to be used as a parameter based on the analysis of network's weights, similarly to a method introduced by Yoon et al. (1994). Both the above methods are closely related to Garson's algorithm (1991) and the "weight product" (WP) method by Tchaban et al. (1998) that are used in the present research. The "general influence" method was also used in a comparison review by Papadokonstantakis et al. (2006), along with the method proposed by Nord and Jacobsson (1998)—which sequentially zeroes the weights of each input variable to measure the impact on the network response—and the automatic relevance determination method that is based on a probabilistic interpretation of the network training procedure. Montano and Palmer (2003) proposed numeric sensitivity analysis, based on the calculation of the slopes between inputs and outputs; Özesmi and Özesmi (1999) followed a different approach than numerical, based on the visual interpretation of the connection weights among neurons. Our research also uses the method of partial derivatives (PD) of neural network's response with respect to each input. This method has shown an adequate performance in real case studies by Dimopoulos et al. (1995) and Ibarra et al. (2003) and in comparisons with other methods. More specifically, PD were applied on simulated data by Gevrey et al. (2003), Montano and Palmer (2003), and Olden et al. (2004). Chuvieco et al. (1999b) in their ANN approach of fire ignitions replaced the original values of each input variable by random values, after the network training. It was assumed that an increase in the root mean square error, produced by the changed values of each variable, reflected its high significance.

2 Data and methodology

2.1 Study area

The island of Lesvos is located in the northeastern Aegean Sea of Greece and covers an area of 1,636 km^2 with a variety of geological formations, climatic conditions, and vegetation types (Fig. 1). The climate is Mediterranean, with warm and dry summers and mild and moderately rainy winters. Annual precipitation average is 670 mm; the average annual air temperature is 18°C with high oscillations between maximum and minimum daily temperatures. The terrain is rather hilly and rough, with a highest peak of 960 m a.s.l. Slopes greater than 20% are dominant covering almost two-thirds of the island. The soils of Lesvos are widely cultivated, mainly with rain-fed crops such as cereals, vines, and olives. Due to low productivity, many sites were abandoned 40–50 years ago; after abandonment, these areas were moderately grazed and the shrub regeneration has been occasionally cleared by illegal burning. The vegetation of these areas, defined on the basis of the dominant species, includes *phrygana* or *garrigue*-type shrubs in grasslands, evergreen-

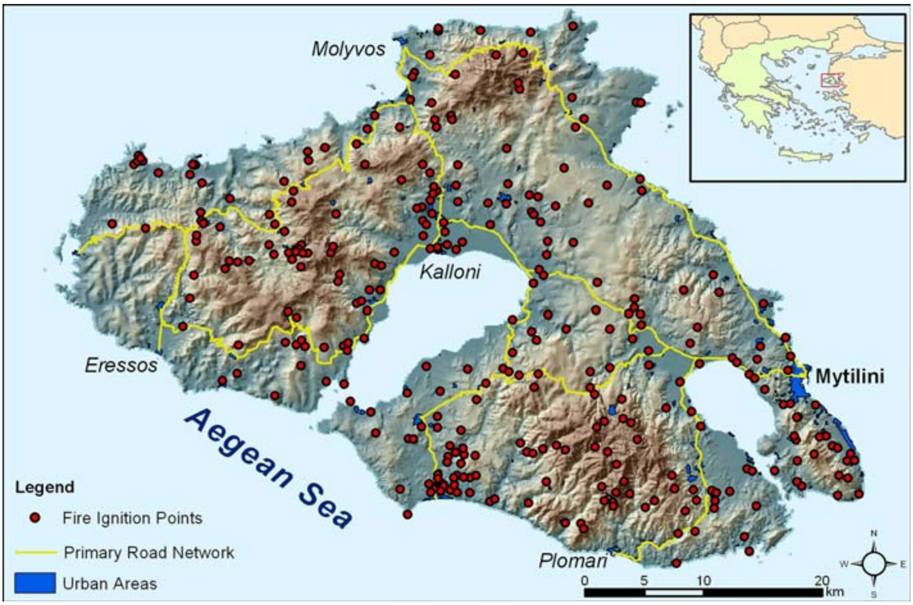

Fig. 1 The study area of Lesvos Island, Greece, with wildfire occurrence (1970–2001), road network, and urban areas

sclerophylous or *maquis*-type shrubs, pine forests, deciduous oaks, olive groves, and other agricultural lands.

More than 420 fires, mostly human caused, occurred in the period 1970–2001 resulting in approximately 80 km^2 of burned area in Lesvos Island (Fig. 1). Even though the number of fires has been increased in recent years, the burned area has declined due to the active fire suppression undertaken by the Greek Fire Brigade using effective tactics and sometimes heavy equipment. During the study period, 62% of wildfires of known causes (approximately 55% of all fires) were due to human negligence; 16% arson, 10% lightning caused, 6.5% military activities, 2.8% garbage disposal related, and 1.7% from electric power lines.

2.2 Neural networks and ignition variables

In our previously published research (Vasilakos et al. 2007), three different neural networks were developed and trained to calculate three intermediate outcomes of FII: i.e., the Fire Weather Index (FWI), the Fire Hazard Index (FHI), and the Fire Risk Index (FRI). Weather is a critical element of the fire environment and at times has more influence than topography and vegetation—once a threshold level of fuel presence and continuity is reached. The weather component is included in most of the operational Fire Danger Rating Systems, e.g., the well-known FWI of the Canadian Forest Fire Weather Index System (Van Wagner 1987). In our research, the FWI is different than the Canadian one as it examines the wind, humidity, temperature, and rainfall variables that are generally considered to determine fire danger potential (Schroeder and Buck 1970), and how these parameters influence the fire ignition potential.

Fire hazards result from the interactions of biophysical factors such as vegetation and topography (see Chen et al. 2003); these parameters are included in the FHI. Wildfire ignition potential is strongly correlated with the quantity, size, density, quality, continuity (all of them standardized in fuel types or models), and moisture content (i.e., the amount of water in a fuel particle) of vegetation that determine the availability of fuel for combustion (Andrews et al. 2003). Topography modifies the general climate over the landscape and thereby affects fuel availability.

The FRI refers to the probability that a wildland fire will start as determined by the presence and activities of causative agents. Human agents are of great importance in fire risk assessment, especially in the Mediterranean countries where they are primary causes of forest fires either by accident or arson (Henderson et al. 2005). Spatial analysis of human risk is quite complex, due to the difficulty of spatially correlating socioeconomic activities in the biophysical and anthropogenic environment (Chen et al. 2003; Kalabokidis et al. 2002a). The primary method used to delineate human risk has been the correlation of the spatial distribution of fire ignition to the proximity of human activities (e.g., Chuvieco and Congalton 1989; Chuvieco and Salas 1996; Kalabokidis et al. 2007).

The functional approach to calculate the three indices of wildfire occurrence was performed through the use of a multi-layer perceptron (MLP) that had been trained with the most widely used method for training, i.e., the back-propagation algorithm (Rumelhart and McClelland 1986). The input parameters of the FII such as meteorological conditions, distances from human activities, vegetation, and topography have been chosen to be easily defined, thus enabling the system for immediate operational application on a local level. It should be noticed also that the parameters have been chosen to reflect the wildland fire ignition pattern based on the fire ignition causes of our study area. Table 1 analytically presents the input variables for each index on fire weather, fire hazard, and fire risk

Table 1 Description of input variables for each index

Index	Input variable	Type	Description
FWI	Air temperature	C	°C
	Wind speed	C	m/s
	Relative humidity	C	%
	Rain	B	Rainfall in the last 24 h
FHI	Fuel models	C	Flammability index
	10-h Fuel moisture content	C	%
	Elevation	C	m
	Aspect	C	°
FRI	Distance to primary road network	C	m
	Distance to secondary road network	C	m
	Distance to power lines	C	m
	Distance to urban areas	C	m
	Distance to landfills	C	m
	Distance to recreational areas	C	m
	Distance to agricultural land	C	m
	Month	C	% of total fire ignitions
	Day of the week	B	Weekend or weekday

FWI Fire Weather Index, *FHI* Fire Hazard Index, *FRI* Fire Risk Index, *C* continuous, *B* binary, *D* discrete

assessment as defined in the fire danger evaluation scheme. The dependant variable for all three models was binary expressed by the presence (1) or absence (0) of fire ignition.

Training and validation samples were created from the total fire history database, to be used in the neural networks for each index. Due to the absence of daily meteorological data, with the exception of some of doubtful credibility, fires that occurred during the 1997–2001 period (May–September) had to be used for the training of the FWI and FHI. For the training of FRI, all the fires having occurred from May to September (1970–2001) were used. For the simulation of the system's operation and its better validation, fires that occurred in 2003 were used. A final validation under operational conditions was conducted in 2004.

To evaluate the performance of the ANNs during training, the mean square error (MSE) function was used:

$$\mathrm{MSE} = \frac{1}{n}\sum_{k}(t_k - d_k)^2 \tag{2}$$

where t_k is the desired outcome, d_k is the actual outcome in the output layer, and n is the total number of the training sample. The logistic function that follows was used as an activation function, which is necessary for the implementation of non-linearity in the network:

$$f(z) = \frac{1}{1 + \mathrm{e}^{-z}} \tag{3}$$

This function approaches 1, for large positive values of z, and 0 for large negative values of z, and is appropriate for the occurrence or non-occurrence of fires since the dependant variable has a binary value of 0 or 1 (Jordan 1995; Sarle 1997). Moreover, the use of this continuous, differentiable, and monotonically non-decreasing function as the activation function allows for the interpretation of the result as a probability (Bishop 1995; Hampshire and Pearlmutter 1990). The logistic function was also used in the output neurons to avoid effects from noisy data that do not conform to the identity function or any other linear function (Masters 1993). Multiple tests were performed to select the final structure of the neural networks for each index. For all three networks, one hidden layer was selected with 6, 4, and 8 hidden nodes for FWI, FHI, and FRI, respectively, while a second hidden layer induced the problem of local minima. In addition, the MSE of the training and the validation datasets was monitored during training; the process was continued until the MSE of the training or the validation dataset started to increase. A learning rate $r = 0.1$ was chosen, while the output neuron was considered activated in case of an output value above 0.5. According to the results, the FWI function was more easily approached, whereas the FRI had better classification percentages regarding the evaluation year of 2003.

2.3 Methods of ranking each variable's importance

Four methods were followed to examine the contribution of each variable at the final models: Garson's (1991) algorithm and Tchaban et al.'s (1998) "weight product" were based on connection weights after the training procedure; logistic regression was used for ranking based on the calculated coefficients; and a ranking method based on PD of the output was applied.

2.3.1 Garson's algorithm

Garson (1991) proposed a method for measuring the relative importance of input variables in a neural network based on connection weights of an already successful trained network, while modifications of this method were presented later (Goh 1995; Howes and Crook 1999; Yoon et al. 1994). Garson's algorithm uses the absolute values of the connection weights; therefore, it is unknown whether an input variable has a positive or negative impact on the output. It is based on Eq. 4 that calculates the percentage of influence Q_{ik} (hereinafter PI) of the input variable x_i on the output y_k, assuming a network where w is the connection weight between N input neurons and L hidden neurons, and u is the connection between the hidden neurons and the output. According to the equation, the sum of percentages of inputs equals 100%:

$$Q_{ik} = \frac{\sum_{j=1}^{L}\left(\frac{w_{ij}}{\sum_{r=1}^{N} w_{rj}} u_{jk}\right)}{\sum_{i=1}^{N}\left(\sum_{j=1}^{L}\left(\frac{w_{ij}}{\sum_{r=1}^{N} w_{rj}}\right)\right)} \tag{4}$$

2.3.2 Tchaban's "weight product"

Tchaban et al.'s (1998) "weight product" is a more simplified method compared to Garson's, and it is also based on weight magnitudes. It depends on the ratio of input variable x_i and the calculated output y_k to estimate the influence WP_{ik} by the following equation:

$$\mathrm{WP}_{ik} = \frac{x_i}{y_k}\sum_{j=1}^{L} w_{ij}u_{jk} \tag{5}$$

According to the above equation, every input pattern returns a different "weight product"; thus, the arithmetic mean and the standard deviation are calculated for all the training dataset to evaluate the influence of each input.

2.3.3 Logistic regression

The logistic regression model (LR) is considered as the most suitable method for the probability estimation of an event to occur (e.g., fires or vegetation types) in case that the dependant or response variable is expressed in a binary way (Hosmer and Lemeshow 1989; Kalabokidis and Koutsias 2000; Vasconcelos et al. 2001). The logistic regression model can be expressed with Eq. 6 (Mendenhall and Sincich 1996):

$$f(z) = \frac{1}{1 + \mathrm{e}^{-(b_0 + b_1x_1 + b_2x_2 + \cdots + b_ix_i)}} \tag{6}$$

where x_i are the qualitative or quantitative independent variables and b_i are the estimated coefficients.

2.3.4 Partial derivatives

The importance of ANN inputs can be calculated through the PD of the network's output with respect to each input, and the method is considered quite efficient in measuring the influence of variables within a neural network. By calculating the PD, the Jacobian matrix can be evaluated to provide the sensitivity of the output as a result of small changes in the inputs (Bishop 1995). In the present case study, the PD show how much the value of each fire index is changing when the input variable (e.g., one of the weather variables) changes. This method has been extensively described by Dimopoulos et al. (1995) and has been used widely (Dedecker et al. 2005; Dimopoulos et al. 1999; Gevrey et al. 2003; Montano and Palmer 2003; Park et al. 2007). PD depend on the input x_i, the output y_k, the connection weights between inputs and hidden modes w_{ij}, the connection weights between hidden and output nodes u_{jk}, and the activation function. The sensitivity S_{ik} of the output y_k with respect to input x_i can be expressed through Eq. 7:

$$S_{ik} = \frac{\partial y_k}{\partial x_i} = f'(\mathrm{net}_k)\sum_{j=1}^{L} u_{jk}f'\left(\mathrm{net}_j\right)w_{ij} \tag{7}$$

where $f'(\mathrm{net}_j)$ and $f'(\mathrm{net}_k)$ are the derivatives of the activation function of hidden nodes and output nodes, respectively. Under the assumption that the trained network will not be linear, every input pattern will result in a different slope for each input vector. In order to rank the relative contribution of each input to the output, the sum of the square partial derivatives (SSD) per input variable was calculated. Therefore, every input variable results in one SSD value, while high value of SSD means significant influence. The plots of the PD of the neural network response with respect to each input variable provide a useful representation of the influence of each input parameter, throughout its domain, to the output.

3 Results and discussion

The scope of this article is to focus on presenting the results from sensitivity analysis to understand each variable's influence on the fire weather, hazard, and risk indices. The results from the four methods are summarized in tabular form showing the magnitude of influence for each input and in a scatter-plot format of the PD. The latter show the kind of influence of each input on each fire potential index, i.e., positive values of PD means positive influence and vice versa, while the magnitude of values indicates the rate of importance. Thereby, systematic assessment of hazards and risks aids in composing quantitative indices of wildland fire behavior and effects with spatial layers of meteorological, vegetative, topographic, and socioeconomic information that could eventually develop geographical fire danger indices (Kalabokidis et al. 2002a).

3.1 Fire Weather Index (FWI)

Table 2 shows the results of the four methods applied to examine how meteorological conditions affect fire ignition probability. Except LR, all other methods show the occurrence of rainfall in the previous 24 h as the most significant parameter. More specifically, the PI method estimates the rainfall's influence to be 35.9%, in agreement with the rest of the methods that classify this parameter as by far the most important of all variables involved. Temperature is ranked second among all methods, while wind speed and relative humidity follow according to the PI, WP, and PD methods. According to the signs of the mean in WP and of the coefficient in LR methods, temperature and relative humidity appear to have a positive influence on wildfire ignition while wind speed and rain have a negative one. This association suggests that increased summer temperatures and low amounts of rainfall account for high fire potential. The signs of relative humidity and wind speed are in contrast with the general perception of how these variables affect fire ignitions, although wind is considered to have a dominant influence on fire (e.g., Finklin and Fischer 1990). This is explained by the very low added effect of these variables compared to temperature and rain (i.e., low mean and coefficient values of relative humidity and wind speed versus high mean and coefficient values of temperature and rainfall in Table 2), as well as by studying more thoroughly Fig. 2 where PD of the FWI neural network response with respect to each input variable are presented.

The study of Fig. 2 shows how meteorological conditions affect fire ignition probability throughout their input domain. The PD values of FWI with respect to temperature are all

Table 2 Results of the four methods applied for the sensitivity analysis of the input variable in Fire Weather Index

Variable	PI (%)	WP		LR		PD (SSD)
		Mean	SD	*B*	Sig.	
Temperature	28.7	3.893	1.710	5.963	0.002	11.246
Wind speed	20.9	−0.438	0.294	−17.100	0.004	1.078
Relative humidity	14.5	0.277	0.186	0.919	0.559	0.008
Rain	35.9	−6.002	9.707	−2.528	0.004	20.657
Total	100.0					

PI percentage of influence, *WP* weight product, *LR* logistic regression model, *PD* partial derivatives

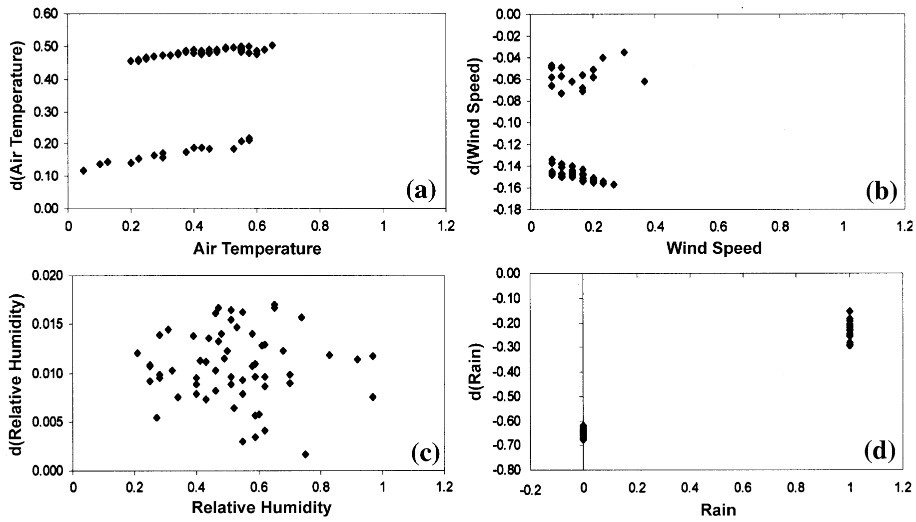

Fig. 2 Partial derivatives of the FWI neural network response with respect to each input variable

positive (Fig. 2a), meaning that an increase in the temperature leads to an increase in fire ignition potential. The higher values of temperature result in a larger positive change of fire potential. The PD values of FWI with respect to rain are all negative (Fig. 2d), meaning that when rain occurs in the last 24 h then fire danger decreases. This association confirms the insight that increased summer temperatures and low amounts of rainfall account for high fire potential. Diurnal variations in temperature and precipitation accordingly have significant influence on fuel moisture content and therefore on fire potential (e.g., Andrews et al. 2003).

The PD values of FWI with respect to relative humidity are all positive (Fig. 2c), while the derivative values of FWI with respect to wind speed are all negative (Fig. 2b). Even if this effect seems to be opposite to reality, it practically shows that relative humidity and wind speed have minimal influence on fire ignitions for our study area. This can also be seen by studying more thoroughly Fig. 2b and c. There is not a precise direction of plots and the maximum values are quite low leading to the conclusion that fire ignition is either not directly affected by these variables in our study area or influenced in combination with other variables. This is also verified by a previous study (Kalabokidis et al. 2002b) where in a similar ecological and socio-economical study area, summer temperature and rain were significant in fire ignitions while relative humidity had no influence on wildfire occurrence mainly due to the narrow range of summer relative humidity (40–60%) in Greece. Wind has very high variability in speed and direction, with topography, vegetation, and local heating and cooling having influences that result in micro-climates with localized conditions (Trewartha and Horn 1980).

3.2 Fire Hazard Index (FHI)

Table 3 shows how vegetation (described in terms of both fuel models and fuel moisture) and topography (elevation above sea level and terrain aspect) influence wildland fire ignition. All methods result in the large influence of 10-h fuel moisture content (defined by Deeming et al. 1977) as the first in ranking. The PI calculates the 10-h fuel moisture's

Table 3 Results of the four methods applied for the sensitivity analysis of the input variable in Fire Hazard Index

Variable	PI (%)	WP		LR		PD (SSD)
		Mean	SD	*B*	Sig.	
Fuel models	12.2	422.999	2267.412	3.694	0.041	760.877
10-h Fuel moisture	60.3	−16246.136	81357.857	−6.532	0.035	29578.630
Elevation	10.6	−1551.479	7968.604	−1.061	0.478	288.322
Aspect	16.9	−1042.417	5666.958	0.712	0.275	299.648
Total	100.0					

PI percentage of influence, *WP* weight product, *LR* logistic regression model, *PD* partial derivatives

influence equal to 60.3%, while the mean of WP and the coefficient of LR both show very high influence. With the exception of the first ranking variable, there is no total agreement with the rest of the variables regarding their ranking, apart from PD that rank the fuel models second, the aspect third, and the elevation fourth, according to the respective magnitude of their statistics. Regarding the type of influence, the signs of the means in WP and of the coefficient of LR show that fuel models (i.e., flammability) have positive influence on wildfire ignition; the 10-h fuel moisture, elevation, and aspect have all negative effect. The only exception is the aspect that has positive influence according to the LR, suggesting increases in fire hazard as the sunlight moves across different aspects during the day.

Figure 3 shows the PD of the FHI neural network response with respect to each input variable. The PD values of FHI with respect to the flammability of fuel models are almost all positive, meaning that an increase in the flammability index naturally leads to the increase in fire potential ignition (Fig. 3a). The PD values of FHI with respect to 10-h fuel moisture are all negative for small values and near zero for higher values

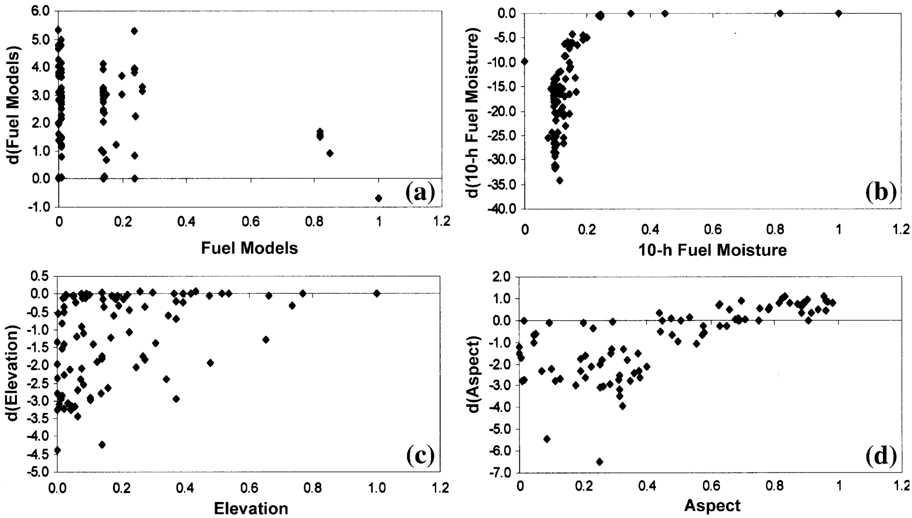

Fig. 3 Partial derivatives of the FHI neural network response with respect to each input variable

(Fig. 3b), meaning that an increase in 10-h fuel moisture leads to a decrease in fire potential ignition, while for high moisture values the FHI tends to be independent from this parameter. The elevation also has the same effect as 10-h fuel moisture on fire danger, even if its values are more spread out (Fig. 3c). At lower elevations fire danger increases more rapidly, while at higher elevation fire danger tends to become constant depending more on climatic anomalies. Length of fire season and fuel vary with elevation due to differences in amount of precipitation received, dates of snow melt, and greening and curing dates of vegetation.

Aspect is the direction a slope is facing, measured clockwise in degrees from 0 to 360. Aspect affects fire potential through variations in the amount of solar radiation and wind that different aspects receive. The PD values of FHI with respect to terrain aspect (Fig. 3d) are negative for values 0–200° (north, east, and south) and positive for values 200–360° (southwest, west, and north). Slopes facing eastwards have the maximum negative influence (e.g., earlier heating and cooling and generally lee sides of mountains). South to southwest facing slopes typically have the greatest number of fires and longest fire season due to high solar radiation and vegetation water stress. Northerly aspects (having both 0 and 360° values) may support low frequency but high intensity fire regimes due to more moist and heaviest fuel conditions (e.g., Agee 1994; Larjavaara et al. 2004).

The strong influence of the 10-h fuel moisture is indicated by the more evenly formed plot of the values of PD in Fig. 3b that clearly show its effects. This is further supported by the very high values of PD values of FHI with respect to 10-h fuel moisture (Table 3). Indeed, the 10-h fuel (i.e., fuel particles <2.5 cm in diameter) moisture content is one of the critical parameters of fuel properties related to fire ignition and behavior. Drying of the fine dead fuels does increase ignition probability and potential fire behavior (Chuvieco et al. 1999a). Consequently, a lot of operational fire danger systems include its calculation (Deeming et al. 1977; Stocks et al. 1989). For the elevation above sea level, fire occurrence is affected negatively at higher altitudes where (i) climatic conditions tend not to lead to vegetation water stress and (ii) fire ignition causes (e.g., human activities) are not close.

3.3 Fire Risk Index (FRI)

The importance of human impact variables that were taken into consideration in the FRI, according to the four methods used, is shown in Table 4. In general, there was no total agreement among the methods regarding the ranking of a variable's influence. This is explained by high covariance from the larger number of variables in FRI compared to the other two indices, and the fact that there was not a prevailing influence from any individual input parameter (e.g., in the case of 10-h fuel moisture content in FHI), but all of them have had comparable effects. The different computation of all the methods leads to different results in magnitude (Gevrey et al. 2003); thus the association between rank orders was further evaluated with Spearman's rho (ρ) in the next section. However, some similarities regarding the magnitude of influence were observed as in the case of the month parameter in which the frequency of fire ignitions was ranked first according to WP, and PD, second according to PI and high in LR (Table 4). Also, the proximity to recreation areas has been classified last in WP and second to last in LR and PD. With the exception of proximity from recreational areas, all other variables have had the same kind of influence according to all methods, as denoted by the signs of their respective statistics in Table 4. This can also be noticed from Fig. 4, where the PD of the FRI neural network response with respect to each input variable are presented.

Table 4 Results of the four methods applied for the sensitivity analysis of the input variable in Fire Risk Index

Variable	PI (%)	WP		LR		PD (SSD)
		Mean	SD	*B*	Sig.	
Primary road network	17.3	−3.977	6.844	−2.462	0.031	293.830
Secondary road network	6.0	−1.944	3.067	−2.002	0.023	48.318
Power lines	6.3	−1.702	2.922	−1.538	0.216	60.052
Urban areas	9.1	−9.183	12.857	−2.319	0.038	489.701
Landfills	9.2	4.750	6.157	0.841	0.345	175.973
Recreation areas	12.1	0.247	0.248	−0.489	0.471	41.051
Agricultural land	12.8	1.608	3.491	1.600	0.115	99.546
Month	14.3	32.965	12.690	1.724	0.000	1304.499
Day of the week	12.9	3.186	6.010	0.392	0.201	36.074
Total	100.00					

PI percentage of influence, *WP* weight product, *LR* logistic regression model, *PD* partial derivatives

The PD values of FRI with respect to distance from recreational areas are spread over the domain of input values (0–1), resulting at the same time in both negative and positive PD values (Fig. 4f). A reason for this effect can be attributed to near-zero influence of this variable on fire risk, while it can have some effects only by interacting with other inputs. Regarding the kind of influence, potential fire ignition decreases away from main and secondary roads, power lines, and urban areas (Table 4, Fig. 4a–d). It is an interesting finding that fire danger does not increase as we move close to landfills and agricultural lands (Fig. 4e, g) in our study area, contrary to the general perception that these human activities constitute major potential causes of fire ignition. Also, weekdays seem not to affect fire danger as opposed to weekends when fire danger increases significantly (Fig. 4i). Our results from the analysis of the FRI input variables explain well the human risk factors of fire history records in Lesvos Island (Fig. 1). August and September are the most critical months in wildland fire occurrence in the island of Lesvos, followed by the proximity to urban areas and main roads. Demographics and urban growth lead to pressures on wildlands and result in more ignition sources (Chuvieco et al. 1999b) in the vicinity of urban areas, while the dense human presence on or close to main roads means more possible causes of ignitions by accident or negligence (Thompson 2000). Recreational areas include beaches and places of religious and archeological attraction, and they are ranked last on fire risk levels since visitors tend not to cause fire ignitions. In the Mediterranean Basin, places with conflicts in the rural and urban interface seem to have a higher potential of ignition risk, especially during the months of August and September when the weather conditions and fire operational limitations may support severe wildland fires.

4 Conclusions

One of proposed methods for the study of fire ignition potential is ANNs that show a significant ability in pattern recognition of complex models, including natural phenomena (Chuvieco et al. 1999b; Vasconcelos et al. 2001; Vasilakos et al. 2007; Vega-García et al. 1996). Contrary to their modeling flexibility, ANN's analytical capabilities of directly

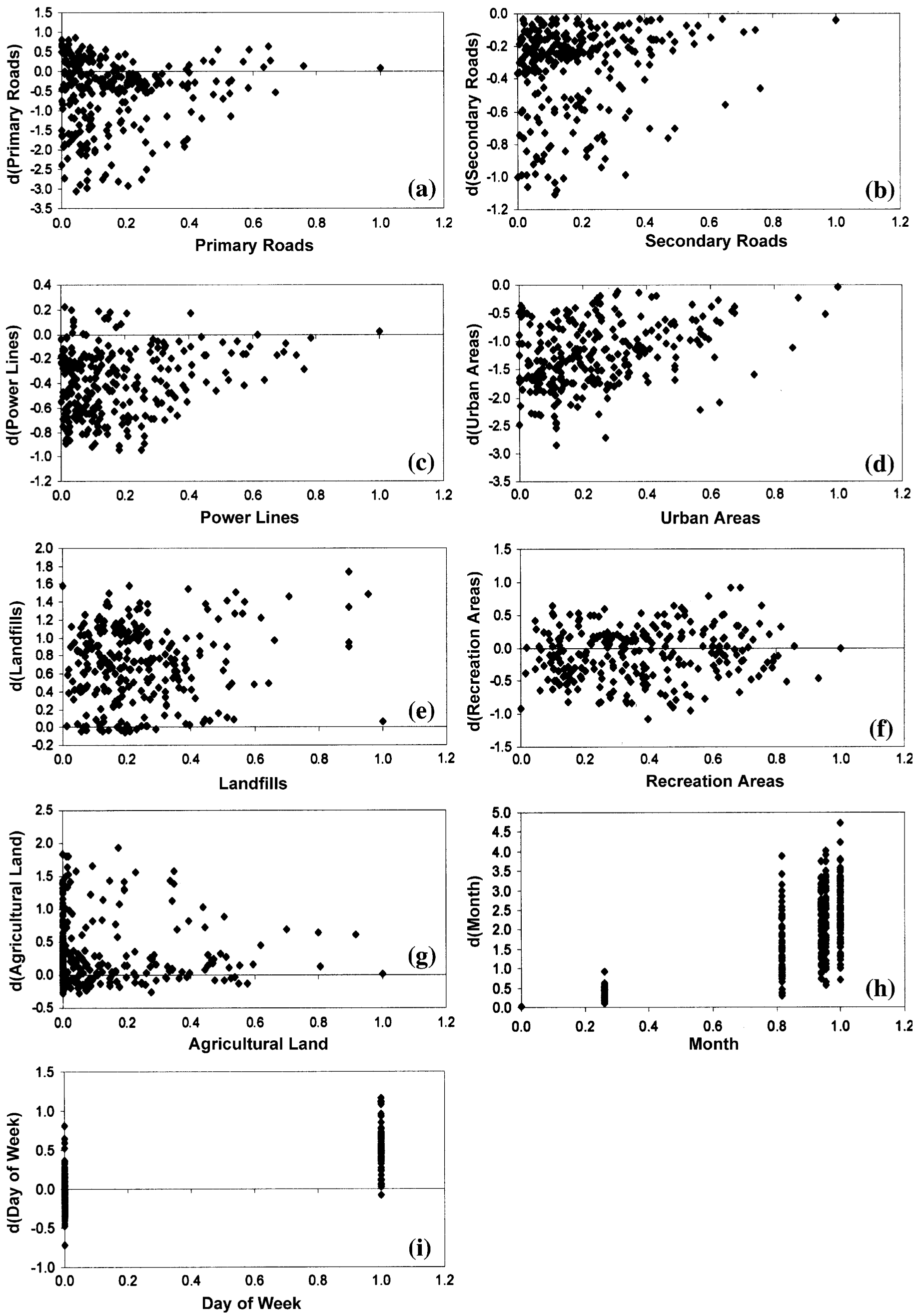

Fig. 4 Partial derivatives of the FRI neural network response with respect to each input variable

measuring variables' influence are limited. Various explanatory methods have been explored to dismantle this "black box." The scope of the current research was to apply four explanatory methods to estimate variables' importance on a fire ignition danger scheme.

Table 5 Spearman's rho (ρ) of ranking methods for Fire Weather Index (FWI), Fire Hazard Index (FHI), and Fire Risk Index (FRI)

Index	Method	PI	WP	LR	PD	Mean
FWI	PI	1.000	1.000	0.200	1.000	0.733
	WP	1.000	1.000	0.200	1.000	0.733
	LR	0.200	0.200	1.000	0.200	0.200
	PD	1.000	1.000	0.200	1.000	0.733
FHI	PI	1.000	0.400	0.400	0.800	0.533
	WP	0.400	1.000	0.400	0.200	0.333
	LR	0.400	0.400	1.000	0.800	0.533
	PD	0.800	0.200	0.800	1.000	0.600
FRI	PI	1.000	0.250	0.033	0.150	0.144
	WP	0.250	1.000	0.417	0.850	0.506
	LR	0.033	0.417	1.000	0.550	0.333
	PD	0.150	0.850	0.550	1.000	0.516

In the calculation of mean, the ρ coefficient between the same methods is excluded

PI percentage of influence, *WP* weight product, *LR* logistic regression model, *PD* partial derivatives

In our real-data research study, the relative performance of each method was statistically epitomized with the Spearman's rank correlation coefficient rho (ρ) and the results for each index can be seen in Table 5 (Norusis 1990). PI, WP, and PD showed similar significant performance in calculation of ranking variables of FWI (mean Spearman's ρ coefficients of 0.733 in Table 5), while LR was not able to classify the input parameters especially rainfall and wind speed (Table 2). The input variables that composed FHI were better classified by PD (mean Spearman's ρ coefficients of 0.600 in Table 5); it is worth mentioning that all methods ranked first the 10-h fuel moisture content, which had a significant influence on the calculation of this index (Table 3). The WP and PD methods have shown almost similar performance in the estimation of rankings for FRI (mean Spearman's ρ coefficients of 0.506 and 0.516 in Table 5, respectively), where nine-paired ranks have existed (Table 4). Overall, PD have shown considerable stability and credibility on their performance (even with inputs having close influence); performance of the rest of the methods was affected by the number of input parameters and the magnitude of their importance.

Conclusively, Table 6 summarizes the ranking of the variables' influence according to the results discussed above. In regard to weather conditions, the most significant parameter for fire ignition in the study area is the occurrence of rainfall in the last 24 h followed by temperature, wind speed, and relative humidity. The largest positive change of fire danger is caused by the higher values of temperature, and the biggest rate of negative change of fire danger is caused by the occurrence of rainfall in the last 24 h (Fig. 2). With respect to the vegetation and topography fire hazard, 10-h fuel moisture content is the most significant variable followed by fuel models, aspect, and elevation (Table 6). Regarding human presence and socioeconomic impacts, the FRI (Table 6) is mostly influenced by month of the year followed by proximities to urban areas, landfills and main roads; distance from recreational areas seems to have the smallest effect. The FRI had the biggest rate/slope of positive change by months of the year with high percentage of historic fire ignitions, and the biggest rate of negative change by proximity to main roads and urban areas (Fig. 4).

Table 6 Ranking of input variables' influence for each fire index

Index	Variable	PI	WP	LR	PD
FWI	Temperature	2	2	2	2
	Wind speed	3	3	1	3
	Relative humidity	4	4	4	4
	Rain	1	1	3	1
FHI	Fuel models	3	4	2	2
	10-h Fuel moisture	1	1	1	1
	Elevation	4	2	3	4
	Aspect	2	3	4	3
FRI	Primary road network	1	4	1	3
	Secondary road network	9	6	3	7
	Power lines	8	7	6	6
	Urban areas	7	2	2	2
	Landfills	6	3	7	4
	Recreation areas	5	9	8	8
	Agricultural land	4	8	5	5
	Month	2	1	4	1
	Day of the week	3	5	9	9

FWI Fire Weather Index, *FHI* Fire Hazard Index, *FRI* Fire Risk Index, *PI* percentage of influence, *WP* weight product, *LR* logistic regression model, *PD* partial derivatives

Regarding the comparison of the different methods used in this research, the use of PD as an explanation of a variable's influence in a neural network that models wildland fire ignitions can provide useful results to forest management and other agencies dealing with forest fires. As a result, preventive measures can be focused on the factors identified as important and enhance the reduction of fire ignitions and the mitigation of likely damages.

Acknowledgments This research was funded by the European Union within the RTD project "Automated Fire and Flood Hazard Protection System/AUTO-HAZARD PRO" (EVG1-CT-2001-00057). The authors would like to thank their colleagues at the Geography of Natural Disasters Laboratory, University of the Aegean, and the Greek Fire Brigade and Forest Service authorities for their cooperation. Dr. Philip N. Omi of Colorado State University, USA; Dr. Peter F. Moore of Sustainable Forestry Management Ltd, Australasia; and two anonymous referees are acknowledged for their review comments and suggestions in an earlier version of the manuscript.

References

Agee JK (1994) Fire and weather disturbances in terrestrial ecosystems of the eastern Cascades. General technical report PNW-GTR-320, USDA, Forest Service, Pacific Northwest Research Station, Portland

Alcázar J, Garcia CV, Grauet M, Pemán J, Fernández Á (1998) Human risk and fire danger estimation through multicriteria evaluation methods for forest fire prevention in Barcelona, Spain. In: Proceedings of the 3rd international conference on forest fire research and 14th conference on fire and forest meteorology, Luso, 16–20 Nov 1998, pp 2379–2387

Andersson FO, Aberg M, Jacobsson SP (2000) Algorithmic approaches for studies of variable influence, contribution and selection in neural networks. Chemom Intell Lab Syst 51:61–72. doi:10.1016/S0169-7439(00)00057-5

Andrews PL, Bevins CD, Seli RC (2003) BehavePlus fire modeling system user's guide, v. 2.0. General technical report RMRS-GTR-106WWW, USDA, Forest Service, Rocky Mountain Research Station, Ogden

Bishop CM (1995) Neural networks for pattern recognition. Oxford University Press, Oxford

Chen K, Blong R, Jacobson C (2003) Towards an integrated approach to natural hazards risk assessment using GIS: with reference to bushfires. Environ Manag 31(4):546–560. doi:10.1007/s00267-002-2747-y

Chou YH (1992a) Management of wildfires with a Geographical Information System. Int J Geogr Inf Syst 6(2):123–140. doi:10.1080/02693799208901900

Chou YH (1992b) Spatial autocorrelation and weighting functions in the distribution of wildland fires. Int J Wildland Fire 2(4):169–176. doi:10.1071/WF9920169

Chou YH, Minnich RA, Chase RA (1993) Mapping probability of fire occurrence in San Jacinto Mountains, California, USA. Environ Manag 17(1):129–140. doi:10.1007/BF02393801

Chuvieco E, Congalton RG (1989) Application of remote sensing and Geographic Information System to forest fire hazard mapping. Remote Sens Environ 29:147–159. doi:10.1016/0034-4257(89)90023-0

Chuvieco E, Salas J (1996) Mapping the spatial distribution of forest fire danger using GIS. Int J Geogr Inf Syst 10(3):333–345. doi:10.1080/02693799608902082

Chuvieco E, Deshayes M, Stach N, Cocero D, Riaño D (1999a) Short-term fire risk: foliage moisture content estimation from satellite data. In: Chuvieco E (ed) Remote sensing of large wildfires in the European Mediterranean Basin. Springer-Verlag, Berlin

Chuvieco E, Salas J, Carvacho L, Rodríguez-Silva F (1999b) Integrated fire risk mapping. In: Chuvieco E (ed) Remote sensing of large wildfires in the European Mediterranean Basin. Springer-Verlag, Berlin

Dedecker AP, Goethals PL, D'heygere T, Gevrey M, Lek S, De Pauw N (2005) Application of artificial neural network models to analyse the relationships between *Gammarus pulex* L. (Crustacea, Amphipoda) and river characteristics. Environ Monit Assess 111(1–3):223–241. doi:10.1007/s10661-005-8221-6

Deeming JE, Burgan RE, Cohen JD (1977) The national fire-danger rating system—1978. General technical report INT-39, USDA, Forest Service, Intermountain Forest and Range Experiment Station, Ogden

Dimopoulos I, Bourret P, Lek S (1995) Use of some sensitivity criteria for choosing networks with good generalization ability. Neural Process Lett 2(6):1–4. doi:10.1007/BF02309007

Dimopoulos I, Chronopoulos J, Chronopoulou-Sereli A, Lek S (1999) Neural network models to study relationships between lead concentration in grasses and permanent urban descriptors in Athens city (Greece). Ecol Model 120:157–165. doi:10.1016/S0304-3800(99)00099-X

Fausett L (1994) Fundamentals of neural networks: architecture, algorithms and applications. Prentice-Hall Inc, USA

Finklin AI, Fischer WC (1990) Weather station handbook—an interagency guide for wildland mangers. National Wildfire Coordinating Group, NFES no. 2140, Boise

Garson GD (1991) Interpreting neural network connection weights. Artif Intell Expert 6:47–51

Gevrey M, Dimopoulos I, Lek S (2003) Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecol Model 160(3):249–264. doi:10.1016/S0304-3800(02)00257-0

Goh ATC (1995) Back-propagation neural networks for modeling complex systems. Artif Intell Eng 9:143–151. doi:10.1016/0954-1810(94)00011-S

Hampshire JBII, Pearlmutter BA (1990) Equivalence proofs for multi-layer perceptron classifiers and the Bayesian discriminant function. In: Touretzky D, Elman J, Sejnowski T, Hinton G (eds) Proceedings of the 1990 Connectionist Models Summer School. Morgan Kaufmann, CA

Henderson M, Kalabokidis K, Marmaras E, Konstantinidis P, Marangudakis M (2005) Fire and society: a comparative analysis of wildfire in Greece and the United States. Hum Ecol Rev 12(2):169–182

Hoffmann AA, Schindler L, Goldammer JG (1999) Aspects of a fire information system for East Kalimantan, Indonesia. In: Proceedings of the 3rd international symposium on Asian tropical forest management, Samarinda, 20–23 Sept 1999

Hosmer D, Lemeshow S (1989) Applied logistic regression. A much-cited recent treatment utilized in SPSS routines. Wiley, NY

Howes P, Crook N (1999) Using input parameter influences to support the decisions of feedforward neural networks. Neurocomputing 24:191–206. doi:10.1016/S0925-2312(98)00102-7

Ibarra AA, Gevrey M, Park Y, Lim P, Lek S (2003) Modelling the factors that influence fish guilds composition using a back-propagation network: assessment of metrics for indices of biotic integrity. Ecol Model 160:281–290. doi:10.1016/S0304-3800(02)00259-4

Jordan MI (1995) Why the logistic function? A tutorial discussion on probabilities and neural networks. MIT computational cognitive science report 9503. http://www.cs.berkeley.edu/~jordan

Kalabokidis KD, Koutsias N (2000) Geographic multivariate analysis of spatial fire occurrence. Geotech Sci Issues 11(1):37–47

Kalabokidis KD, Gatzojannis S, Galatsidas S (2002a) Introducing wildfire into forest management planning: towards a conceptual approach. For Ecol Manag 158(1–3):41–50. doi:10.1016/S0378-1127(00)00715-5

Kalabokidis KD, Konstantinidis P, Vasilakos C (2002b) GIS analysis of physical and human impact on wildfire patterns. In: Viegas DX (ed) Proceedings of the IV international conference on forest fire research & wildland fire safety, Coimbra. Millpress, Rotterdam

Kalabokidis KD, Koutsias N, Konstantinidis P, Vasilakos C (2007) Multivariate analysis of landscape wildfire dynamics in a Mediterranean ecosystem of Greece. Area 39(3):392–402. doi:10.1111/j.1475-4762.2007.00756.x

Larjavaara M, Kuuluvainen T, Tanskanen H, Venäläinen A (2004) Variation in forest fire ignition probability in Finland. Silva Fenn 38(3):253–266

Masters T (1993) Practical neural networks recipes in C++. Academic Press, USA

Mendenhall W, Sincich T (1996) A second course in statistics: regression analysis. Prentice-Hall Inc, USA

Montano JJ, Palmer A (2003) Numeric sensitivity analysis applied to feedforward neural networks. Neural Comput Appl 12:119–125. doi:10.1007/s00521-003-0377-9

Nord LI, Jacobsson SP (1998) A novel method for examination of the variable contribution to computational neural network models. Chemom Intell Lab Syst 44:153–160. doi:10.1016/S0169-7439(98)00118-X

Norusis MJ (1990) SPSS/PC + Advanced Statistics™ 4.0 for the IBM PC/XT/AT and PS/2. SPSS Inc, Chicago

Olden JD, Jackson DA (2002) Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. Ecol Model 154:135–150. doi:10.1016/S0304-3800(02)00064-9

Olden JD, Joy MK, Death RG (2004) An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. Ecol Model 178:389–397. doi:10.1016/j.ecolmodel.2004.03.013

Özesmi SL, Özesmi U (1999) An artificial neural network approach to spatial habitat modeling with interspecific interaction. Ecol Model 116:15–31. doi:10.1016/S0304-3800(98)00149-5

Papadokonstantakis A, Lygeros A, Jacobsson SP (2006) Comparison of recent methods for inference of variable influence in neural networks. Neural Netw 19(4):500–513. doi:10.1016/j.neunet.2005.09.002

Park YS, Rabinovich J, Lek S (2007) Sensitivity analysis and stability patterns of two-species pest models using artificial neural networks. Ecol Model 204(3–4):427–438. doi:10.1016/j.ecolmodel.2007.01.021

Preisler HK, Brillinger DR, Burgan RE, Benoit JW (2004) Probability based models for estimation of wildfire risk. Int J Wildland Fire 13(2):133–142. doi:10.1071/WF02061

Pyne SJ, Andrews PL, Laven RD (1996) Introduction to wildland fire, 2nd edn. Wiley, NY

Rumelhart DE, McClelland JL (1986) Parallel distributed processing: explorations in the microstructure of cognition. The MIT Press, USA

Salas FJ, Chuvieco E (1994) GIS applications to forest fire risk mapping. Wildfire 3:7–13

Sarle WS (1997) Neural network FAQ. Periodic posting to the Usenet newsgroup comp.ai.neural-nets. ftp://ftp.sas.com/pub/neural/FAQ.html

Schroeder MJ, Buck CC (1970) Fire weather. Agriculture handbook 360. USDA Forest Service, Washington, DC

Stocks BJ, Lawson BD, Alexander ME, Van Wagner CE, McAlpine RS, Lynham TJ, Dubé DE (1989) The Canadian forest fire danger rating system: an overview. For Chron 65:258–265

Taylor SW, Alexander ME (2006) Science, technology and human factors in fire danger rating: the Canadian experience. Int J Wildland Fire 15:121–135. doi:10.1071/WF05021

Tchaban T, Taylor MJ, Griffin A (1998) Establishing impacts of the inputs in a feedforward network. Neural Comput Appl 7:309–317. doi:10.1007/BF01428122

Thompson WA (2000) Using forest fire hazard modeling in multiple use forest management planning. For Ecol Manag 134:163–176. doi:10.1016/S0378-1127(99)00255-8

Trewartha GT, Horn LH (1980) An introduction to climate, 5th edn. McGraw-Hill, NY

Van Wagner CE (1987) Development and structure of the Canadian Forest Fire Weather Index System. Forestry technical report 35, Canadian Forestry Service, Ottawa

Vasconcelos MJP, Silva S, Tome M, Alvim M, Pereira JMC (2001) Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks. Photogramm Eng Remote Sens 67(1):73–81

Vasilakos C, Kalabokidis K, Hatzopoulos J, Kallos G, Matsinos J (2007) Integrating new methods and tools in fire danger rating. Int J Wildland Fire 16(3):306–316. doi:10.1071/WF05091

Vega-García C, Lee B, Wooddard PM, Titus SJ (1996) Applying neural network technology to human-caused wildfire occurrence prediction. Artif Intell Appl 10(3):9–18

Viegas X, Bovio G, Ferreira A, Nosenzo A, Bernard S (1999) Comparative study of various methods of fire danger evaluation in Southern Europe. Int J Wildland Fire 9(4):235–246. doi:10.1071/WF00015

Yoon Y, Guimaraes T, Swales G (1994) Integrating artificial neural networks with rule-based expert systems. Decis Support Syst 11:497–507. doi:10.1016/0167-9236(94)90021-3

Yuan M (1997) Use of knowledge acquisition to build wildfire representation in Geographical Information Systems. Int J Geogr Inf Sci 11(8):723–745. doi:10.1080/136588197242059