Journal of Plant Ecology Advance Access published January 13, 2015

Journal of
Plant Ecology

PAGES 1–11

doi:10.1093/jpe/rtu041

available online at
www.jpe.oxfordjournals.org

Historic distribution and driving factors of human-caused fires in the Chinese boreal forest between 1972 and 2005

Futao Guo[1,2], *John L. Innes*[2], *Guangyu Wang*[2,*], *Xiangqing Ma*[1], *Long Sun*[3], *Haiqing Hu*[3] *and Zhangwen Su*[1]

[1] *Faculty of Forestry, Fujian Agriculture and Forestry University, No. 15 Shangxiadian Road, Fuzhou 350002, China*
[2] *Sustainable Forest Management Laboratory, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, British Columbia V6T 1Z4, Canada*
[3] *Faculty of Forestry, Northeast Forestry University, No. 26 Hexing Road, Harbin, Heilongjiang Province 150040, China*
*Correspondence address. Sustainable Forest Management Laboratory, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, British Columbia V6T 1Z4, Canada. Tel: (+1)-604-822-2681; Fax: (+1)-604-822-8645; E-mail: guangyu.wang@ubc.ca

Abstract

Aims

The pattern and driving factors of forest fires are of interest for fire occurrence prediction and forest fire management. The aims of the study were: (i) to describe the history of human-caused fires by season and size of burned area over time; (ii) to identify the spatial patterns of human-caused fires and test for the existence of 'hotspots' to determine their exact locations in the Daxing'an mountains; (iii) to determine the driving factors that determine the spatial distribution and the possibility of human-caused fire occurrence.

Methods

In this study, *K*-function and Kernel density estimation were used to analyze the spatial pattern of human-caused fires. The analysis was conducted in S-plus and ArcGIS environments, respectively. The analysis of driving factors was performed in SPSS 19.0 based on a logistic regression model. The variables used to identify factors that influence fire occurrence included vegetation types, meteorological conditions, socioeconomic factors, topography and infrastructure factors, which were extracted and collected through the spatial analysis mode of ArcGIS and from official statistics, respectively.

Important Findings

The annual number of human-caused fires and the area burnt have declined since 1987 due to the implementation of a forest fire protection act. There were significant spatial heterogeneity and seasonal variations in the distribution of human-caused fires in the Daxing'an Mountains. The heterogeneity was caused by elevation, distance to the nearest railway, forest type and temperature. A logistic regression model was developed to predict the likelihood of human-caused fire occurrence in the Daxing'an Mountains; its global accuracy attained 64.8%. The model was thus comparable to other relevant studies.

Keywords: anthropogenic fire, spatial distribution, *K*-function, Kernel intensity estimation, boreal forest, logistic regression

Received: 2 March 2014, Revised: 5 November 2014, Accepted: 22 November 2014

INTRODUCTION

Forest fires, whether caused by lightning or human activities, are an important part of the ecology of the boreal forest. Fire is one of the primary agents of forest renewal and succession (Podur *et al.* 2003), but it can have devastating effects on forest ecosystems including loss of wildlife habitat and forest resources, high management costs and losses to communities (Chas-Amill *et al.* 2010). The fire regimes of many forest ecosystems are largely caused and shaped by human settlement and behavior (Bergeron *et al.* 2004). Human activities are responsible for most wildfire ignitions worldwide (Prestemon and Butry 2005). Forest fires are a global issue: ~50 000 fires sweep through >700 000 ha of forest and other wooded land in the Mediterranean Basin alone each year (Dimitrakopoulos and Mitsopoulos 2006). On average, >100 000 wildfires occur

annually in the USA, burning 1.6–2 million hectares of land, and 10 000 fires burn 2.5 million ha of wildland areas in Canada each year (BC Ministry of Forests and Range Wildfire Management Branch 2009). In China, >10 000 forest fires occur on average every year, burning 820 000 ha annually (Zhong *et al.* 2003).

Chinese boreal forests, geographically distributed in the Daxing'an Mountains of northeastern China, are the most southern part of the global boreal forest biome. The number of human-caused forest fires in this area accounts for 60–80% of the total number of fires from all ignition causes. Forest fire prevention and protection of forest resources from fires is a critical for forestry development in China due to the scarcity of forest. The total forest cover in China is only two-thirds of the global average value, and the *per capita* forest area is 0.145 ha, less than one-fourth of the world average (China Forest Resources Report 2009). In order to decrease forest degradation by fire effectively, especially human-caused fires, understanding the spatial and temporal characteristics of human-caused fires and the potential drivers of fire occurrence are important. The temporal and spatial distributions of forest fires have been studied in different countries and regions (Amatulli *et al.* 2007; Grala and Cooke 2010; Gralewicz *et al.* 2012). There are also some studies of human-caused fires and their driving factors (Martínez *et al.* 2009; Padilla and Vega-Garcia 2011). However, the results from these studies vary widely due to differences in fire regimes, vegetation and other factors, making generalizations difficult. In China, there are few studies of human-caused fires. The studies that exist either do not fully consider the influence of socioeconomic factors or do not take into account the variation in human infrastructure (Liu *et al.* 2012; Tian *et al.* 2011).

To fill this knowledge gap, we applied a spatial point pattern (SPP) approach to examine the spatial pattern of human-caused fires (Cressie 1993). *K*-function, *L*-function and Kernel intensity estimation are often used to deal with nearest statistic and intensity estimation within the SPP approach (de la Riva *et al.* 2004; Genton *et al.* 2006). The logistic regression was used to test the factors driving the probability of fire occurrence in China's boreal forest. The objectives of the study were: (i) to describe the history of human-caused fires by season and area burnt over time; (ii) to identify the spatial patterns of human-caused fires and test for the existence of 'hotspots' to determine their exact locations in the Daxing'an mountains and (iii) to determine the driving factors affecting the spatial distribution and the possibility of human-caused fires.

MATERIALS AND METHODS

Study site

China's boreal forests, geographically located in the Daxing'an Mountains of northeastern China, are the southernmost part of the global boreal forest biome (Jiang *et al.* 2002). In this study, we selected the Daxing'an Mountains of Heilongjiang Province as our study area. The study area was located 50°10–53°33′N and 121°12′–127°00′E, and had a total area of 8.46×10^6 ha (Fig. 1). The climate is typical of the cold temperate zone. The dominant species is Dahurian larch (*Larix gmelinii* Rupr.), which is normally accompanied by white birch (*Betula platyphylla* Suk.), Mongolian pine (*Pinus sylvestris* L. var. *mongolica* Litv.) and Mongolian oak (*Quercus mongolica* Fischer ex Ledebour). The mean annual temperature lies between −2 and 4°C, with a range extending from −52.3 to

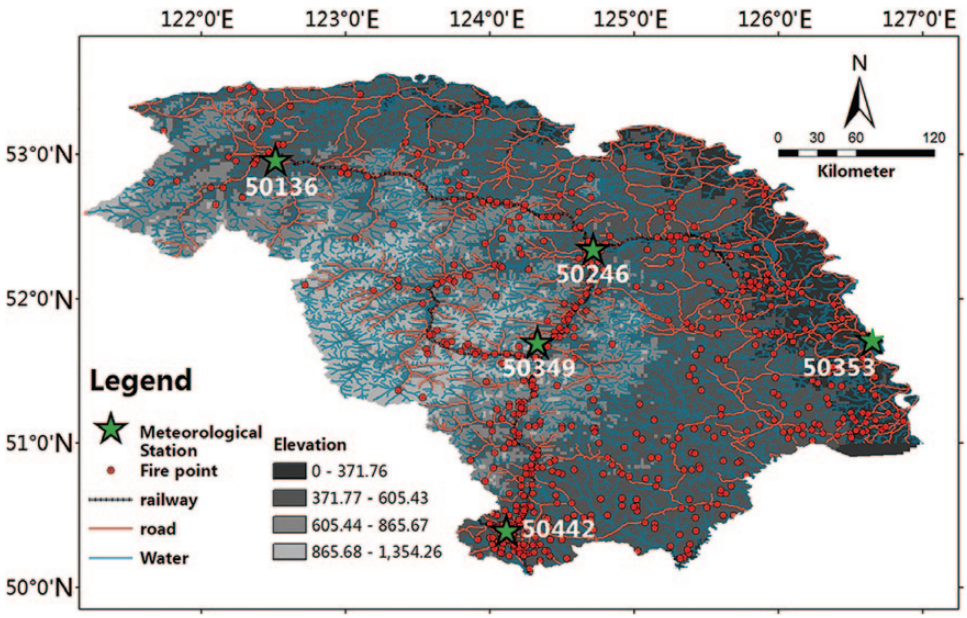

Figure 1: location of the study region together with fire points (dots) and meteorological stations (stars).

39.0°C. The mean total annual precipitation is between 350–500 mm (Guo *et al.* 2011). It is the largest natural forest area in China, and provided 126 million m^3 of timber harvest from 1964 to 2005. It is exposed to extremely high fire risk and has the largest average annually burned area in China. There were >1000 forest fires, including >600 human-caused fires, in the period between 1980 and 2005, with the total area of burned forest amounting to 1 300 000 ha. Dendrochronological studies have indicated that the historical fire regime was characterized by frequent, low-intensity surface fires and the fire return interval ranged from 30 to 120 years (Xu *et al.* 1997).

Data collection

Fire record

Our analysis used human-caused fire data for the Daxing'an Mountains from 1972 to 2005, provided by the Fire Prevention Office of Jiagedaqi city (FPOJ). One of FPOJ's main responsibilities is the collection of fire occurrence information, such as fire location, size, causes and date of occurrence of forest fires in the Daxing'an Mountains. The definition of human-caused fire in this study did not include the controlled prescribed burns and other actions taken by government or forest management agencies, but did contain prescribed burns that had got out of control and caused an unexpectedly large burned area. Other important causes of ignition included smoking, hunting, fireworks, escaped fire from locomotives and residents' homes. The data were acquired in a geo-database format (ESRI [Environmental System Research Institute] data storage and management framework) and contained geographically referenced point locations of all non-structure fires in the Daxing'an Mountains. Before 1990, the records of fire location were determined by the fire chief, who identified each fire location through a combined approach of fixed observation points in the forest and the Terrain and Forest Instruction Map (1:100 000). The location of fires after 1990 were recorded by GPS [Global Positioning System]. The dataset from FPOJ was used to determine the spatial patterns of human-caused fire occurrence and statistics for burned area per fire. This dataset was also used to identify temporal characteristics of the human-caused fires, including annual and monthly trends of human-caused fire frequencies and sizes.

Vegetation types

A digital vegetation map of China with 1 km resolution was downloaded from The Cold and Arid Regions Science Data Center, China (http://westdc.westgis.ac.cn/). The map includes 20 vegetation functional types, such as needleleaf evergreen tree (temperate), needleleaf deciduous tree, broadleaf evergreen tree (temperate), broadleaf deciduous tree (tropical), etc. We clipped the digital map using the boundary of the Daxing'an mountains region. In addition, we grouped the polygons into eight categories, including needleleaf deciduous tree (covering 30.6% of the territory of the Daxing'an mountains), broadleaf deciduous tree (covering 12.3%), needleleaf evergreen tree (covering 11.6%), broadleaf deciduous shrub (covering 7.45%), grass (covering 29.3%), crop (covering 3.72%), permanent wetlands (covering 4.5%) and water bodies (covering 0.52%). The vegetation types for each fire point and non-fire point were extracted from the vegetation map layer in ArcGIS 10.0.

Meteorological factors

The mean temperature, relative humidity and precipitation data with a spatial resolution of 0.085° and monthly temporal resolution for the period 1972–2005 were generated by the International Institute of Earth System Science, Nanjing University, China, based on the HADCM2 model, which is a coupled ocean atmosphere model produced by the UK Met Office Hadley Centre. The process of data generation included converting HADCM2 model data to text, extracting China region data, combining with a Digital Elevation Model (DEM) and interpolating using ANUSPLIN software package. The data are available at the Earth System Science Data Sharing Platform, China (http://geodata.nju.edu.cn/). In addition, other weather-related variables including daily rainfall, mean wind speed, mean temperature, minimum temperature, maximum temperature, mean humidity, minimum humidity were provided by the China Meteorological Data and Sharing Network (http://cdc.cma.gov.cn/). There are five national level weather stations located in the study area (Fig. 1) (station ID: 50246, 50247, 50349, 50353 and 50442), and we assigned the values of weather variables for each ignition and non-ignition point based on the nearest weather station.

Socioeconomic factors

Socioeconomic factors, including annual funding for forest fire prevention, population density, *per capita* gross domestic product (GDP) and the number of fire towers, were obtained from secondary sources (see online supplementary Table S1). The effect of inflation was eliminated while calculating the Annual Funding for Forest Fire Prevention and Capita GDP.

Topographic factors

We collected a 1:250 000 Digital Line Graphic (DLG) map from the National Administration of Surveying, Mapping and Geoinformation of China. The elevation, aspect and slope of each fire point and non-fire point (control point) were derived from the DEM layer of the DLG map in an ArcGIS 10.0 environment, employing spatial analysis.

Infrastructure factors

Infrastructure variables included the interannual variation of the infrastructure in the study area, such as the number of fixed inspection stations and the length of burned line for fire prevention, and also involved the distance of the domain to the nearest railway, road, river and area of settlement. The relevant information was extracted from the Local Chronicles of Forest Fire Prevention of the Daxing'an Mountains and the DLG map.

The information about the independent variables in this study has been summarized in online supplementary Table S1.

Data analysis

Spatial statistical methods and logistic regression modeling were used to examine the spatial pattern and the driving factors for fire ignition in the Daxing'an Mountains, respectively. The analytical process for the study is shown schematically in Fig. 2.

Analysis of spatial pattern

Ripley's $K(r)$ function is a nearest-neighbor statistical method and has been widely used to describe the relationships between two or more point patterns (Peterson and Squiers 1995; Stoyan and Penttinen 2000). The K-function can be simplified to:

$$\hat{K}(r)=\frac{1}{\lambda^2 A}\times\sum_{i\neq j}\sum I_r(d_{ij}) \quad (1)$$

where, d_{ij} is the distance between the ith and jth observed fires, A is the size of the study area, r is the distance from any fire point (in this study r can be considered as measuring radius), λ is the intensity or mean number of fires per unit area, and $I_r(d_{ij})$ is an indicator function with a value of 1 if $d_{ij} \leq r$ and 0 otherwise (Dixon 2002). In practice, transformation of the K-function is usually made to increase the intuitive judgment on the point patterns, termed the L-function.

$$\hat{L}(r)=\sqrt{\frac{\hat{K}(r)}{\pi}} \quad (2)$$

where, $\hat{L}(h)$ for a random process is $\sqrt{\frac{\pi r^2}{\pi}}=r$, the 1:1 line. If $\hat{L}(r)$ is above this line, the SPP tends to be clustered, and if $\hat{L}(r)$ is below this line, the SPP is likely to be regular. As a test of significance, a random spatial point pattern was simulated 99 times and the $\hat{L}(r)$ function plotted (Podur *et al.* 2003). The L-function was calculated using the spatial module of S-Plus (Venables and Ripley 1999).

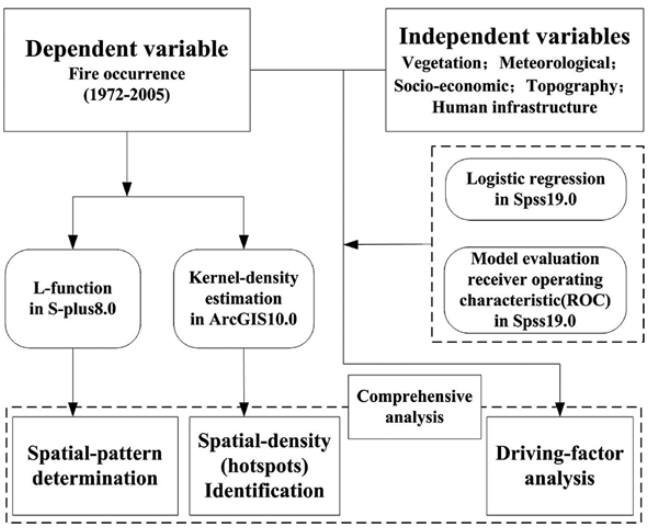

Figure 2: the analysis process of the study.

In this study, we used intervals of 5 years to calculate the patterns of fire occurrence and determine the spatial density. There were two reasons to do so: (i) sufficient samples were obtained to enable a statistical analysis (Lafon and Grissino-Mayer 2007; Nielsen and Wendroth 2003), and (ii) one year's forest fires could interact with a former year's fire and later years through consuming the forest fuel (Collins *et al.* 2009). In addition, the results of fire patterns and spatial density could be affected by the chosen length of intervals. The selection of the interval length depends on the purpose and the temporal scale of the research. Various intervals have been used in previous studies (Lafon and Grissino-Mayer 2007; Podur *et al.* 2003). In this study, we attempted various intervals for the spatial analysis, from 1 to 5 years; however, the results showed significant differences in the spatial patterns of human-caused fire at the 5-year scale.

Analysis of spatial density

The spatial intensity of a point process can be calculated using a Kernel density estimator (Gavin *et al.* 1993):

$$\hat{\lambda}_\tau(z)=\frac{1}{n\tau}\sum_{i-1}^{n}k\left(\frac{z-z_i}{\tau}\right) \quad (3)$$

where, $\hat{\lambda}(z)$ is the estimated spatial intensity of fires, $z_1 \ldots z_n$ are the locations of n observed fires, $k(.)$ is the kernel, a symmetric but not necessarily positive function that integrates to one, and τ is the bandwidth and determines the radius of a disc centered on z within which points Z_i will contribute to the intensity $\hat{\lambda}(z)$.

The same time intervals as 'fire occurrence pattern' were adopted for the spatial density analysis. The spatial analyst mode in ArcGIS 10.0 was used to conduct the Kernel estimation of the spatial density of human-caused fires. According to function (3), the bandwidth (named 'search radius' in ArcGIS 10.0) τ is a key to kernel estimation. There are several methods to find the appropriate size of the bandwidth. One is the selection of the bandwidth subjectively by eye. This method is based on looking at several density estimates over a range of bandwidths and selecting the density that is the most suitable in some sense (Nazan Kuter *et al.* 2011). Another method involves mean random distance (RD mean) calculations (de la Riva *et al.* 2004; Koutsias *et al.* 2004). The results of the calculations are determined by the number of ignition points and the polygon size. We tried different bandwidths (i.e. 10, 30, 50 and 70 km) for the kernel estimation. However, only the 50 km bandwidth showed a smooth spatial density map. On the other hand, we obtained a range of bandwidths from 39 to 66 km by RD mean calculation. We chose 50 km as the bandwidth in this study since it generated a smooth density map and also fell within the range of bandwidths obtained by the RD mean method.

Prediction of forest fire occurrence

We predicted fire occurrence using a binary logistic regression model that requires both the presence and absence of fire ignition. Presence and absence were represented by 1 and 0 in binary logistic regression model, respectively. For presence, we

used the fire record data to generate a point cover based on fire ignition locations using ArcGIS 10.0. For absence, we randomly generated the same number as ignition points (Catry *et al.* 2009; Chang *et al.* 2013) within the study area. The independent variables were extracted for each fire ignition and non-ignition point except for the daily weather variables. The values of the daily weather variables recorded by weather stations close to the fire ignition points were assigned to those ignition points. For the randomly generated non-ignition points, we chose a date for each non-ignition point from the dates of the ignition points, and assigned the daily weather variables for each non-ignition point according to the nearest weather station.

In addition, we randomly selected 70% of ignition points (428) and 70% of non-ignition points (418) as the building dataset to fit the binary logistic regression model, and the remaining points (30%) as the validation dataset to perform model validation. All analyses were performed using SPSS 19.0 software.

We used the Wald test (Legendre and Legendre 1998) to assess the significance of each variable ($P < 0.05$). The overall significance was evaluated according to the Hosmer and Lemeshow goodness-of-fit test (Hosmer and Lemeshow 2000). In order to assess the predictive ability of the logistic model, 2×2 classification tables of observed vs. predicted values were constructed using both the building and validation dataset. The Yueden criterion was applied to determine the probability threshold (cut-off point) at and above which the occurrence of fire ignition is accepted, and below which it is considered that no fire occurred (Garcia *et al.* 1995). This approach has also been used in other relevant studies (Catry *et al.* 2009; Chang *et al.* 2013). In addition, an alternative method called receiver operating characteristic (ROC) analysis (Fielding and Bell 1997) was applied to assess the predictive ability of the model. The ROC curve was obtained by plotting sensitivity vs. specificity for various probability thresholds. The area under the curve (AUC) can reflect the performance of model (Jiménez-Valverde 2012). This method has already been used to assess the performance of logistic models in wildfire risk estimation (Chang *et al.* 2013; del Hoyo *et al.* 2011). An AUC value between 0.5 and 0.69 can be considered as poor performance; a value between 0.7 and 0.79 indicates reasonable performance; a value of 0.8 or higher indicates excellent performance (del Hoyo *et al.* 2011).

RESULTS

Temporal characteristics

On average, there were 25 human-caused fires annually that burned 77 193 ha. The largest number of fire occurrences was in 1986, with 65 fires, whereas the lowest number was in 1992, with only two fires (Fig. 3). The largest burnt area was in 1987, when 758 701 ha burnt. The smallest burnt area was in 1995, with only 0.4 ha burnt. There is no apparent trend related to the number of human-caused fires before 1987, but the number decreased from 1987 onward until another peak event appeared in 2003. The area burnt had a declining trend in 1972–86 and 1987–2002; with the exception of peak fire events in 1987 and later in 2003 (Fig. 3). On a monthly basis, the largest human-caused fires occurred in May, followed by April and June (Table 1). Although September and October did not account for a large percentage of fire incidences, the area burnt in both months was still considerable.

Fire occurrence patterns

Analysis of fire occurrence patterns revealed clusters in 1972–76, 1977–81, 1982–86 and 1997–2001 at a scale of 200 km and in 1997–2001 at a scale of 100 km, while no obvious peak clustering was found in 1987–91. However, the pattern of fire occurrence was random with the $\hat{L}(h)$ values equal to h, the 1:1 line in 1992–96 (Fig. 4). The simulation tests also revealed that the pattern of fire occurrences was more clustered than random. The analysis of fire occurrence pattern further showed regularity at larger spatial scales; i.e. patterns of regularity appeared at distance scales of >400 km in 1972–76 and at >370 km in 1977–81. The distances of fire occurrence patterns changed from clustered to regular at ranges between 200 and 400 km. The total area of the Daxing'an Mountains is 84 600 km^2 and treating the area as a circle, the radius would be 164 km (below the 200 km threshold); thus the major fire pattern would be clustered at a spatial scale of 164 km in our study area.

Spatial intensity of human-caused fires

Analysis of the spatial intensity of human-caused fires showed similar locations of 'hotspots' for the periods 1972–76, 1977–81, 1982–86 and 1987–91 (Fig. 5). The core area for the 'hotspots' was: 124°10′E, 50°23′N, in the administrative jurisdiction of Jiagedaqi. There were no obvious 'hotspots' in 1992–96, and the patterns of human-caused fire points were random under a bandwidth of 200 km. The 'hotspots' in 1997–2001 were located at 52°09′N, 125°55′, north of Hanjiayuan, and differed from other time intervals. The number of 'hotspots' in 2002–05 was more than other time intervals and mainly located in the south of the Daxing'an Mountains, with core area coordinates of 51°35′N, 125°25′E; 51°13′N, 125°27′E; 51°07′N, 124°38′E; 50°34′N, 125°26′E and 51°16′N, 123°41′E.

Prediction of forest fire occurrence

A logistic regression model was developed using several explanatory variables to predict the likelihood of forest fire occurrence. The first modeling attempt involving all explanatory variables had a significant overall accuracy ($P < 0.001$); and elevation, forest type, monthly mean temperature and daily mean temperature had significant impacts on the likelihood of human-caused fire occurrence (Table 2). We then removed the non-significant independent variables one by one to obtain the best model in which all the independent variables were significant at $P = 0.05$ level. The best logistic regression model (Table 3) indicated that elevation, forest type and distance to the nearest railway were the most influential variables explaining the spatial pattern of natural fire ignition ($P < 0.001$), followed by monthly mean temperature and daily minimum temperature. The Hosmer and Lemeshow

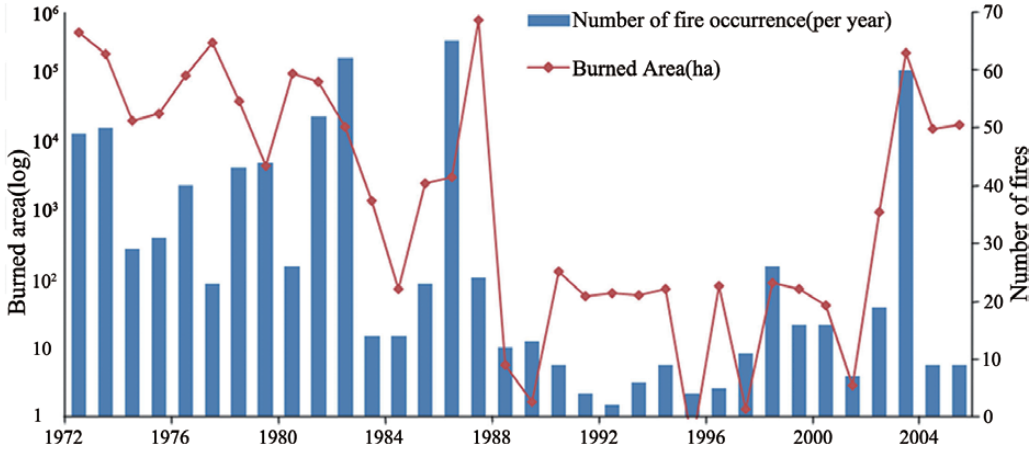

Figure 3: the annual burnt area and human-caused fire frequency in Daxing'an Mountains during 1972–2005.

Table 1: summary of monthly wildfire frequency and area burned in Daxing'an mountains

Month	Average number of fires per year	Percentage of occurrence	Average burned area (ha)	Percentage of burned area
January	0	0	0	0
February	0	0	0	0
March	1.1	3.51	20.92	0.03
April	7.0	22.36	5113.28	6.32
May	9.0	28.75	62 996.07	77.85
June	5.6	19.89	293.45	0.36
July	1.5	4.79	19.28	0.02
August	1.1	3.51	0.001	0
September	3.0	9.58	8383.39	10.36
October	3.0	9.58	4090.02	5.05
November	0	0	0	0
December	0	0	0	0

goodness-of-fit test showed sufficient fit of the regression model to the data (χ^2 = 13.181, df = 8, P = 0.106). The AUC was 0.784 for the prediction (Fig. 6a) and 0.706 for the validation dataset (Fig. 6b). The classification accuracy was also evaluated using the validation dataset, and the goodness-of-fit test showed sufficient fit of the regression model to the validation data (χ^2 = 14.964, df = 8, P = 0.05). The contingency matrix showed that the regression model correctly classified 74.7 and 64.8% of all observations in the prediction dataset (cross-validation) and validation data (independent test set), respectively (Table 4). The best logistic regression model for predicting the likelihood of fire occurrence in the Daxing'an Mountains would be:

$$p = 1 / \left\{ 1 + e^{-\left(\begin{array}{l} 2.877 - 0.004\,Elev - 0.055\,Forest\ type + 0.089\,Mmean \_ \\ Tem - 0.000012\,Dis\_railway - 0.009\,Dmin\_Tem \end{array}\right)} \right\} \quad (4)$$

where, p is the likelihood of fire occurrence, Elev is the elevation, Mmean_Tem is the Daily mean temperature, Dis_Railway is the distance to the nearest railway and Dmin_Tem is the Daily minimum temperature.

DISCUSSION

This study provides evidence that fire occurrence in the Daxing'an Mountains is temporally variable, with a major peak occurrence in 2003 (Fig. 3). This unusual peak is related to fire that escaped from a prescribed burn under conditions of dry weather and large amounts of fuel. The fire spread very quickly and resulted in a large burnt area. However, the overall trend in annual human-caused fires has decreased since 1987, with the main cause for this declining trend being the implementation of the Forest Fire Prevention Acts (FFPA) which were first enacted by the State Council of China in March 1988. As can be seen in Fig. 3, a remarkable result was achieved through the implementation of the FFPA, which resulted in a dramatic decline in the number of human-caused fires in the same year as the FFPA took effect. The positive effect of forest fire management policies on the reduction in fire occurrence has also been stressed in other studies (Galiana *et al.* 2013; Gianni *et al.* 2013).

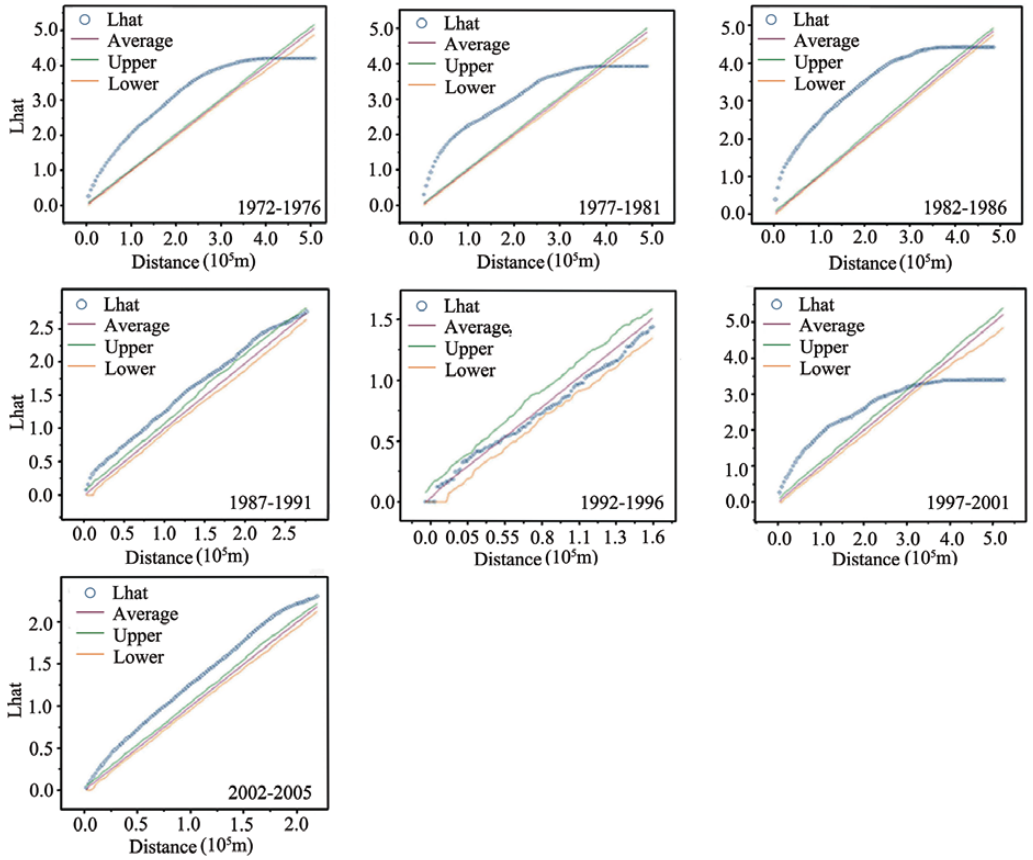

Figure 4: the pattern of human-caused fire occurrence in Daxing'an Mountains during 1972–2005 with 5-year time interval. Lhat was the estimated value of *L*-function. The upper and lower lines define the 95% confidence interval. The average was calculated from the mean of upper and lower lines.

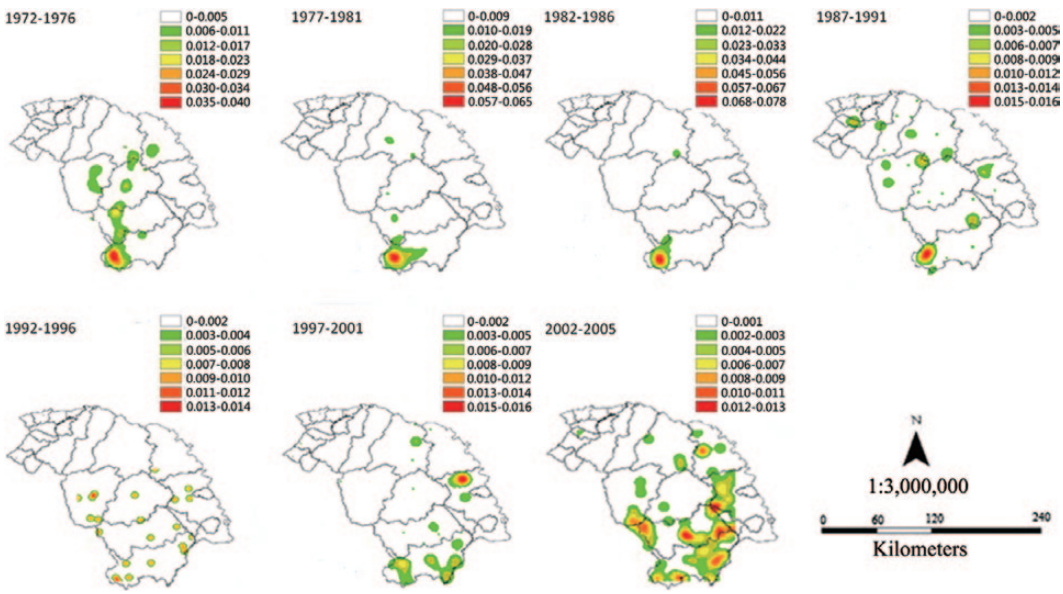

Figure 5: spatial intensity of human-caused fires in Daxing'an Mountains during 1972–2005 with 5-year time interval. The intensity represented the number of human-caused fire per square kilometer based on 50 km bandwidth scale.

Results from the analysis of the spatial pattern and intensity of fires clearly show the existence of hotspots in China's boreal forest. The main hotspot is located in the southwest of the Daxing'an Mountains; consistent with a previous study (Liu *et al.* 2012). The main factors influencing the occurrence of human-caused fires in the Daxing'an Mountains are elevation,

Table 2: results of the multivariate logistic regression model for predicting the likelihood of human-caused fire occurrence in Daxing'an mountains (for all variables)

Variables	df	Coefficients	Standard error	Wald chi square	Pr > |t|
Elevation	1	−0.004	0.001	54.935	<0.0001
Slope	1	−0.001	0.025	0.002	0.969
Forest type	1	−0.060	0.015	16.448	<0.0001
Mmean_Pre	1	0.008	0.119	0.004	0.949
Mmean_RH	1	0.015	0.022	0.471	0.493
Mmean_Tem	1	0.170	0.048	12.486	<0.0001
Dis_Railway	1	0.000	0.000	1.870	0.172
Dis_River	1	0.000	0.000	0.907	0.341
Dis_Road	1	0.000	0.000	0.002	0.962
Dis_Settlement	1	0.000	0.000	1.308	0.253
D_RF	1	0.000	0.001	0.102	0.750
Dmean_WS	1	0.015	0.021	0.488	0.485
Dmean_Tem	1	0.018	0.007	6.274	0.012
Dmean_RH	1	−0.021	0.025	0.715	0.398
Dmin_Tem	1	−0.014	0.017	0.619	0.431
Dmax_Tem	1	−0.013	0.018	0.561	0.454
Capita gross domestic product	1	0.000	0.000	0.012	0.913
Funding	1	0.000	0.000	0.044	0.833
Tower	1	0.003	0.010	0.094	0.759
F_Ins	1	0.001	0.003	0.069	0.793
Len_Bline	1	0.000	0.000	2.188	0.139
Len_Road	1	0.000	0.002	0.002	0.962
Den_Pop	1	−0.295	0.547	0.292	0.589
Constant	1	7.302	15.583	0.220	0.639

Full-model goodness-of-fit statistic = 13.181 (df = 8, $P = 0.106$); AUC = 0.784 ± 0.018 ($P < 0.001$). The meaning and description of each variable showed in Table 1.

Table 3: results of the multivariate logistic regression model for predicting the likelihood of human-caused fire occurrence in Daxing'an mountains based on selected significant variables

Variables	df	Coefficients	Standard error	Wald chi square	Pr > |t|
Elevation	1	−0.004	0.001	67.234	<0.0001
Forest type	1	−0.055	0.014	14.753	<0.0001
Monthly mean temperature	1	0.089	0.028	10.056	0.002
Distance to the nearest railway	1	0.000012	0.000	29.031	<0.0001
Daily minimum temperature	1	−0.009	0.003	7.941	0.005
Constant	1	2.877	0.399	51.875	<0.0001

Full-model goodness-of-fit statistic = 14.964 (df = 8, $P = 0.05$); AUC = 0.706 ± 0.018 ($P < 0.001$). The meaning and description of each variable showed in Table 1.

forest type, distance to the nearest railway and monthly mean temperature. Elevation and forest type were negatively related to the likelihood of fire occurrence, while distance to the nearest railway and the mean monthly temperature were positively related to fire occurrence. We found that the hotspots of human-caused fires are located in low elevation areas with a high density of railways and in the southwest corner of the Daxing'an Mountains, which has relatively higher temperatures compared to other places within the study area. Climatic factors could contribute to the existence of hotspots (Gralewicz *et al.* 2012; Joanne and Martin 2009; Lafon and Grissino-Mayer 2007; Wotton *et al.* 2010). Our findings are consistent with recent studies that have emphasized the importance of topographic, vegetation and climatic factors as drivers of human-caused fire

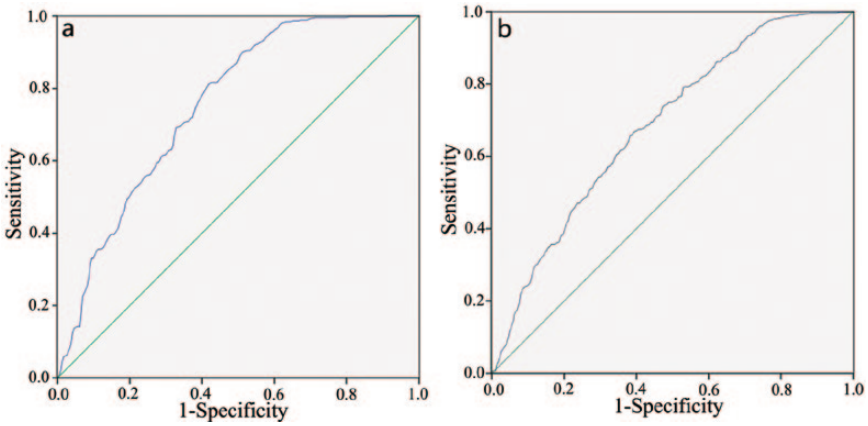

Figure 6: The ROC curves for the logistic regression. (**a**) Prediction dataset; (**b**) Validation dataset.

Table 4: contingency tables for the prediction and validation datasets of multivariate logistic regression models for human-caused fire, with cut-points of 0.3018 according to the Yueden criterion

	Predicted							
	Building dataset				Validation dataset			
	Fire	0	1	Percentage correct	Fire	0	1	Percentage correct
Observed	0	260	142	64.7	0	102	92	52.6
	1	67	375	84.8	1	39	130	76.9
Overall percentage				74.7				64.8

(Chang *et al.* 2013; Liu *et al.* 2012; Martínez *et al.* 2009; Syphard *et al.* 2007; Vilar *et al.* 2010). As a whole, the accuracy of our model (64.8%) is similar (Chang *et al.* 2013) and/or comparable to the global accuracy of other logistic regression models developed to predict fire occurrence (de Vasconcelos *et al.* 2001; del Hoyo *et al.* 2011; Lozano *et al.* 2007).

In this study, socioeconomic factors such as the density of population and GDP did not show a significant influence on the occurrence of human-caused fires in the Daxing'an Mountains. Our findings agree with those of Martínez *et al.* (2009) but disagree with others (Chang *et al.* 2013; Syphard *et al.* 2007). In our case, a possible explanation could be the implementation of a strict fire prevention policy, which requested all residents to move out of the core zone of the forest and prohibited entry into the forest during the fire season. Under these conditions, the influence of the density of population on fire occurrence may be limited.

CONCLUSIONS

In this study, the temporal and spatial distributions of human-caused fires in China's boreal forests from 1972 to 2005 were analyzed and the factors affecting fire occurrence were determined. There is significant spatial heterogeneity in the occurrence of fires, and seasonal variations in the distribution of human-caused fires have occurred (April and May appear to be the peak fire season). The annual number of human-caused fires has trended downward since 1987 due to the implementation of forest fire protection act. The study also provides evidence about the existence of hotspots of human-caused fire in the Daxing'an Mountains within a 50 km bandwidth. Elevation, distance to the nearest railway, forest type and mean monthly temperature play an important role on the spatial distribution of human-caused fires in the study area, and the logistic regression model that was built to predict the possibility of fire occurrence has a reasonable performance. We recommend other factors such as the normalized difference vegetation index, land use, the Pacific Decadal Oscillation and the El Niño—Southern Oscillation should also be taken into account to improve the predictive ability of the logistic regression model.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of Plant Ecology* online.

FUNDING

Asia-Pacific Forests Net (APFNET/2010/FPF/001).

ACKNOWLEDGEMENTS

We thank Prof. Mulualem Tigabu and Dr Howard Harshaw for the revision of the manuscript. This work was based on the project of 'Adaptation of Asia-Pacific Forests to Climate Change'.
Conflict of interest statement. None declared.

REFERENCES

Amatulli G, Pérez-Cabello F, de la Riva J (2007) Mapping lightning/human-caused wildfires occurrence under ignition point location uncertainty. *Ecol Modelling* **200**:321–33.

BC Ministry of Forests and Range Wildfire Management Branch (2009) Climate change and fire management research strategy, 2010.

Bergeron Y, Gauthier S, Flannigan M, *et al.* (2004) Fire regimes at the transition between mixed wood and coniferous boreal forest in Northwestern Quebec. *Ecology* **85**:1916–32.

Catry FX, Rego FC, Bação FL, *et al.* (2009) Modeling and mapping wildfire ignition risk in Portugal. *Int J Wildland Fire* **18**:921–31.

Chang Y, Zhu Z, Bu R, *et al*. (2013) Predicting fire occurrence patterns with logistic regression in Heilongjiang Province China. *Landsc Ecol* **28**:1989–2004.

Chas-Amill ML, Touza J, Prestemon JP (2010) Spatial distribution of human-caused forest fires in Galicia (NWSpain). In Perona G, Brebbia CA (eds). *WIT Transactions on Ecology and the Environment*, **Vol. 137**. Southampton: WIT Press, www. witpress.com, ISSN 1743–3541 (online).

China Forest Resources Report – The seventh national forest resources inventory (2009), China: Administration of State Forestry.

Collins BM, Miller JD, Thode AE, *et al.* (2009) Interactions among wildland fires in a long-established Sierra Nevada natural fire area. *Ecosystems* **12**:114–28.

Cressie NAC (1993) *Statistics for Spatial Data, (Revised Edition)*. New York: Wiley Publishing.

de la Riva J, Pérez-Cabello F, Lana Renault N, *et al.* (2004) Mapping wildfire occurrence at a regional scale. *Remote Sens Environ* **92**:363–9.

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

del Hoyo VL, Martín Isabel M, Martínez Vega F (2011) Logistic regression models for human-caused wildfire risk estimation: analysing the effect of the spatial accuracy in fire occurrence data. *Eur J Forest Res* **130**:983–96.

Dimitrakopoulos AP, Mitsopoulos ID (2006) *Global Forest Resources Assessment 2005-Report on Fires in the Mediterranean Region*. Forestry Department, FAO.

Dixon PM (2002) Ripley's K function. In El-Shaarawi AH, Piegorsch WW (eds). *Encyclopedia of Environmetrics*. Chichester: John Wiley& Sons, 1796–803.

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv* **24**:38–49.

Galiana L, Aguilar S, Lázaro A (2013) An assessment of the effects of forest-related policies upon wildland fires in the European Union: applying the subsidiarity principle. *Forest Policy Econ* **29**:36–44.

Garcia CV, Woodard P, Titus S, *et al.* (1995) A logit model for predicting the daily occurrence of human caused forest-fires. *Int J Wildland Fire* **5**:101–11.

Gavin J, Haberman S, Verrall R (1993) Moving weighted average graduation using kernel estimation. *Insur Mathem Econ* **12**:113–26.

Genton MG, Butry DT, Gumperta M, *et al.* (2006) Spatio-temporal analysis of wildfire ignitions in the St. Johns River water management district. *Int J Wildland Fire* **15**:87–97.

Gianni BP, Zumbrunnen T, Bürgi M, *et al.* (2013) Fire regime shifts as a consequence of fire policy and socio-economic development: an analysis based on the change point approach. *Forest Policy Econ* **29**:7–18.

Grala K, Cooke WH III (2010) Spatial and temporal characteristics of wildfires in Mississippi, USA. *Int J Wildland Fire* **19**:14–28.

Gralewicz NJ, Nelson TA, Wulder MA (2012) Spatial and temporal patterns of wildfire ignitions in Canada from 1980 to 2006. *Int J Wildland Fire* **21**:230–42.

Guo FT, Hu HQ, Sun L, *et al.* (2011) Spatial patterns of lightning-ignited forest fires in Daxing'an mountains, Heilongjiang province, China, 1973–1997. *Adv Mat Res* **183–185**:2268–74.

Hosmer DW, Lemeshow S (2000) *Applied Logistic Regression*. New York: Wiley.

Jiang H, Apps MJ, Peng CH, *et al.* (2002) Modelling the influence of harvesting on Chinese boreal forest carbon dynamics. *Forest Ecol Manag* **169**:65–82.

Jiménez-valverde (2012) Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecol Biogeogr* **21**:498–507.

Joanne H, Martin O (2009) *Weather-Based Estimation of Wildfire Risk*, SFB 649 discussion paper 2009; No. 2009,032.

Koutsias N, Kalabokidis KD, Allgöwer B (2004) Fire occurrence patterns at landscape level: beyond positional accuracy of ignition points with kernel density estimation methods. *Nat Resour Model* **17**:359–76.

Kuter N, Yenilmez F, Kuter S (2011) Forest fire risk mapping by kernel density estimation. *Croat J For Eng* **32**:599–610.

Lafon CW, Grissino-Mayer HD (2007) Spatial patterns of fire occurrence in the central Appalachian mountains and implications for wildland fire management. *Phys Geog* **28**:1–20.

Legendre P, Legendre L (1998) *Numerical Ecology*. Amsterdam: Elsevier, 723.

Liu ZH, Yang J, Chang Y, *et al.* (2012) Spatial patterns and drivers of fire occurrence and its future trend under climate change in a boreal forest of Northeast China. *Glob Change Biol* **18**:2041–56.

Lozano FJ, Suárez-Seoane S, de Luis E (2007) Assessment of several spectral indices derived from multi-temporal landsat data for fire occurrence probability modelling. *Remote Sens Environ* **107**:533–44.

Martínez J, Vega-Garcia C, Chuvieco E (2009) Human-caused wildfire risk rating for prevention planning in Spain. *J Environ Manage* **90**:1241–52.

Nielsen DR, Wendroth O (2003) Spatial and temporal statistics. Sampling field soils and their vegetation. In Nielsen DR, Wendroth O (eds). *GeoEcology Textbook*. Reiskirchen: Catena.

Padilla M, Vega-Garcia C (2011) On the comparative importance of fire danger rating indices and their interaction with spatial and temporal fire occurrences in Spain. *Int J Wildland Fire* **20**:46–58.

Peterson CJ, Squiers ER (1995) An unexpected change in spatial pattern across 10 years in an Aspen-White Pine forest. *J Ecol* **83**:847–55.

Podur J, Martell DL, Csillag F (2003) Spatial patterns of lightning-caused forest fires in Ontario, 1976–1998. *Ecol Model* **164**:1–20.

Prestemon JP, Butry DT (2005) Time to burn: modeling wildland arson as an autoregressive crime function. *Am J Agric Econ* **87**:756–70.

Stoyan D, Penttinen A (2000) Recent applications of point process methods in forestry statistics. *Stat Sci* **15**:61–78.

Syphard AD, Radeloff VC, Keeley JE, *et al.* (2007) Human influence on California fire regimes. *Ecol Appl* **17**:1388–402.

Tian XR, Su LF, Zhao FJ, *et al.* (2011) Future impacts of climate change on forest fire danger in northeastern China. *J Forestry Res* **22**:437–46.

Venables WN, Ripley BD (1999) *Modern Applied Statistics With S-Plus*. New York: Springer Press, 500.

Vilar L, Woolford DG, Martell DL, *et al.* (2010) A model for predicting human-caused wildfire occurrence in the region of Madrid, Spain. *Int J Wildland Fire* **19**:325–37.

Wotton BM, Nock CA, Flannigan MD (2010) Forest fire occurrence and climate change in Canada. *Int J Wildland Fire* **19**:253–71.

Zhong MH, Fan WC, Liu TM, *et al.* (2003) Statistical analysis on current status of China forest fire safety. *Fire Safety J* **38**:257–69.

Xu CH, Li ZD, Qiu Y (1997) Fire disturbance history in virgin forest in northern region of Daxing'anling mountains. *Acta Ecologica Sinica* **17**:3–9.