Title
Modeling the spatial patterns of human wildfire ignition in Yunnan province, China

Authors
Ye, Jiangxia
Wu, Mingshan
Deng, Zhongjian
et al.

Publication Date
2017-12-01

DOI
10.1016/j.apgeog.2017.09.012

Peer reviewed

Applied Geography 89 (2017) 150–162

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog

Modeling the spatial patterns of human wildfire ignition in Yunnan province, China

Jiangxia Ye[a,1,2], Mingshan Wu[b], Zhongjian Deng[c], Shengji Xu[d], Ruliang Zhou[e,**], Keith C. Clarke[f,*]

[a] College of Forestry, Southwest Forestry University, Kunming, 650224, China
[b] College of Mathematics and Physics, Southwest Forestry University, Kunming, 650224, China
[c] Institute of Yunnan Biodiversity, Southwest Forestry University, Kunming, 650224, China
[d] Southern Aviation Forest Bureau, State Forestry Administration, Kunming, 650021, China
[e] College of Geography, Southwest Forestry University, Kunming, 650224, China
[f] Department of Geography, University of California, Santa Barbara, CA, 93117, USA

ARTICLE INFO

Keywords:
Wildfire
Human ignition
Weights-of-evidence
Probability
Integrated prediction

ABSTRACT

Despite wildfire being an important regulator of dryland ecosystems, uncontrolled wildfire can be harmful to both forest ecosystems and human society, and wildfire prevention and control continue to raise worldwide concern. Wildfire management depends on knowledge of wildfire ignitions, both for cause and location. The regimes and factors influencing wildfire ignition have been studied at length. Humans have a profound effect on fire regimes and human activity is responsible for igniting the largest number of fires in our study area. Understanding the spatial patterns of ignitions is foremost to achieving efficiency in wildfire prevention. Previous studies mainly concentrate on overall wildfire risk integrating numerous factors simultaneously, yet the importance of human factors on ignition has not received much attention. In this study, we mapped human accessibility to explore the influence of human activity on wildfire ignition in a simple and straightforward way. A Bayesian weights-of-evidence (WofE) method was developed based on fire hotspots in China's Yunnan province extracted from satellite images and verified as known wildfires for the period 2007–2013. We considered a set of factors that impact fire ignition as associated with human accessibility: the locations of settlements, roads, water and farmland susceptible to human wildfire ignition. Known points of likely wildfire ignition were selected as training samples and all suspected thematic maps of the factors were taken as explanatory layers. Next, the weights of each layer in terms of its explanatory power were computed and used to generate evidence based on a threshold to pass a statistical test. The conditional independence (*CI*) of each layer was checked with the Agterberg-Cheng test. Finally, the posterior probability was calculated and its precision validated using samples of both presence and absence by withheld validation data. A comparison of WofE models was made to test the predictability. Results show proximity to villages, roads and farmland are strongly associated with human wildfire ignition and that wildfire more often occurs at an intermediate distance from high-density human activity. The WofE method proved more powerful than logistic regression, improving predictive accuracy by 10% and was more straightforward in presenting the association of dependence and independence. In addition, WofE with 1000 m buffer bands is more robust in predicting human wildfire ignition risk than binary or 100 m buffers for the ecoregion studied. Our results are significant for advising practical wildfire management and resource allocation, evaluation of human ignition control and also provides a foundation for future efforts toward integrated wildfire prediction.

* Corresponding author.
** Corresponding author.
E-mail addresses: yjx125@163.com (J. Ye), wms616@163.com (M. Wu), 43946916@qq.com (Z. Deng), xushengji@yahoo.com (S. Xu), Zhou_ruliang@163.com (R. Zhou), kcclarke@ucsb.edu (K.C. Clarke).
[1] Present address. University of California, Santa Barbara, CA, 93117, USA.
[2] Permanent address: College of Forestry, Southwest Forestry University, Kunming, China.

https://doi.org/10.1016/j.apgeog.2017.09.012
Received 19 May 2017; Received in revised form 21 September 2017; Accepted 29 September 2017
0143-6228/

1. Introduction

Despite wildfire being an important regulator of ecosystems by influencing vegetation succession, shaping biomass distribution and maintaining biological diversity (Bond & Keeley, 2005; Bond, Woodward, & Midgley, 2005; Bowman et al., 2011; Crisp, Burrows, Cook, Thornhill, & Bowman, 2011; Simon et al., 2009), uncontrolled wildfire is usually destructive to both forest ecosystems and to human society by causing natural resource degradation, economic disruption and loss of life, and reduced biodiversity (Cameron et al., 2009; Johnston, 2009; Kwak et al., 2012; Rodrigues, de la Riva, & Fotheringham, 2014). Especially in managed ecosystems, wildfire prevention and control continues to be a worldwide concern. The regimes and factors influencing wildfire ignition have been studied at length (Cardille, Ventura, & Turner, 2001; Harrison, Marlon, & Bartlein, 2010; Maingi & Henry, 2007; Plucinksi, 2011). Wildfire occurrence is attributed to weather and climate, fuel condition and a source of ignitions (Gralewicz, Nelson, & Wulder, 2012a; Malamud, Millington, & Perry, 2005), based on which, wildfire risk or susceptibility can be assessed (Dickson et al., 2006; Guo et al., 2017; Hawbaker et al., 2013; Xu, Zhang, Chen, Wu, & Li, 2016). Although ignition is an integral component of wildfire factors, it is crucial in terms of wildfire occurrence, because the possibility of wildfire is minimal, no matter how dry the weather conditions and how high the forest flammability without ignition. Humans have a profound affect on fire regimes by being a source of ignitions (Fusco, Abatzoglou, Balch, Finn, & Bradley, 2016), and human activity is responsible for igniting a majority of all fires (Benali et al., 2017; FAO, 2007; Prestemon & Butry, 2005; Román-Cuesta, Gracia, & Retana, 2003). According to statistics, the main causes of wildfires in China are related to human activities, with lightning accounting for only 0.38% of the total (Zhong, Fan, Liu, & Li, 2003). Consequently, an improved understanding of wildfire risk should address the patterns of human activity and its relation to fire ignition (Dickson et al., 2006; Narayanaraj & Wimberly, 2012; Prestemon, Pye, Butry, Holmes, & Mercer, 2002). Significant research effort has been undertaken to explore the relationship between wildfire and its causative factors with the goal of building predictive models (Cardille et al., 2001; Chas-Amil, Prestemon, McClean, & Touza, 2015; Maingi & Henry, 2007; Narayanaraj & Wimberly, 2012; Romero-Calcerrada, Barrio-Parra, Millington, & Novillo, 2010; Román-Cuesta et al., 2003; Salis et al., 2013; Syphard et al., 2007; Watts & Hall, 2016; Ye, Wang, Guo, & Li, 2017), and has concluded that wildfire tends to occur in areas near human infrastructure on the human-wildland interface (Zhang, Lim, & Sharples, 2016), and frequently exhibits nonlinear relationships (Hawbaker et al., 2013). However, previous studies mainly concentrate on overall wildfire risk integrating numerous factors simultaneously, yet the importance of human factors on ignition has not received much attention. Most fire literature examines biological and physical wildfire factors, such as topography, wind, humidity and fuel load (Romero-Calcerrada et al., 2010; Yang, He, & Shifley, 2008). With an increasing concern in studies of the anthropogenic impacts on wildfire regime, some researchers have tried to isolate the human variables in a quantitative way to figure out the patterns of human influence that cause wildfire ignition (Catry, Rego, Bação, & Moreira, 2010; Fusco et al., 2016; Romero-Calcerrada et al., 2010). The modeling of human activity and its patterns of influence in a more explicit spatial way offer a new level of explanation for local government decision making in wildfire prevention. This implies wildfire prevention, rather than firefighting and management after ignition.

With the development of remote sensing and Geographic Information Systems, it is feasible to model human variables and their impacts on wildfire ignition spatially. Yet due to the variety of human motivations and behavior, modeling human activity remains a difficult and complicated problem (Song, Wang, Satoh, & Fan, 2006). Nevertheless, there are clearly empirical associations between wildfire ignition points and certain aspects of the human footprint. Humans usually have an extent of mobility and a geographic range, which is largely determined by the infrastructure and settlements. Consequently, previous work associated with human activity has commonly utilized land cover, distance or proximity to roads, settlements or other infrastructure as straight distance for buffer analyses (Fusco et al., 2016; Gralewicz et al., 2012a; Guo et al., 2017; Hawbaker et al., 2013; Kwak et al., 2012; Maingi & Henry, 2007; Romero-Calcerrada et al., 2010; Zhang et al., 2016). However, the effects of different factors on fire occurrence can vary among ecosystems and across spatial scales (Catry et al., 2010). Additionally, numerous approaches have been employed to estimate wildfire ignition probability. Logistic regression is the most extensively used method due to its flexibility and robustness to non-normally distributed variables (Catry et al., 2010; Curt, Fréjaville, & Lahaye, 2016; Guo et al., 2014, 2016b; Legendre & Legendre, 2012; Rodrigues et al., 2014). However, logistic regression generally produces a result based on approximate linear relations between map layers (Agterberg & Cheng, 2002; Guo et al., 2016c). Consequently, an improved spatial prediction model of human activity should lead to a better understanding of the spatial and temporal patterns of human-caused wildfire ignition.

In this study, we applied an objective model using weights-of-evidence (WofE) to identify the extent of human impacts on wildfire ignition. To better investigate the patterns of human influence, we assumed that wildfire occurrence is mainly determined by human variables. Several steps were required. First, known points of historical wildfire ignition locations were selected as the training samples and all suspected thematic maps of human activities were taken as explanatory layers. Next, the weights were computed and evidence generated at different scales and for categories based on statistical significance. Conditional independence was examined with the Agterberg-Cheng test. Accordingly, the posterior probability was calculated and its precision validated for both presence and absence of wildfire ignition. Lastly, we analyzed the predictive power of the model compared with logistic regression and with different variable patterns.

2. Materials and methods

2.1. Study area

Yunnan province is located in the southwestern border of China between 21°09′-29°15′ N and 97°32′-106°12′E (Fig. 1). It is ranked second among the forested regions of China with abundant forest resources. The area has a highly diverse gene pool of plants and animals, and is among the top 25 global biodiversity hotspots (Myers, Mittermeier, Mittermeier, Da Fonseca, & Kent, 2000; Yuming, Kun, Jiming, & Shengji, 2004). Yunnan has rugged topography, underdeveloped transportation, is relatively poor, has multi-ethnic inhabitants and is highly populated. Additionally, the northerly winter monsoon in this region is usually obstructed by high mountains, generating a warm, dry winter and moderately hot humid summer monsoon type climate (Li et al., 2017). There is a record of continuous droughts in recent years, which have made Yunan also a region of frequent and severe wildfire occurrence, among the most in China (Xu et al., 2007; Zhong et al., 2003). Wildfires in Yunnan mainly occur in winter and spring from December to May, concentrated in spring from mid-February to mid-May (Cao, Wang, & Liu, 2017; Chen, Fan, Niu, & Zheng, 2014). Wildfire in Yunnan has shown a slight upward trend during recent years (Zhao, Shu, Tiao, & Wang, 2009), mostly caused by human activity (Chen, Pereira, Masiero, & Pirotti, 2017; Tian, Zhao, Shu, & Wang, 2013) including arson, fire misuse, and the tradition of honoring ancestors around Tomb Sweeping Day by burning imitation currency. Additionally, the slash-and-burn farming cultivation and the mosaic of farmland and forest aggravate this situation. State and local governments conduct large scale ground patrols to check the ignitions induced by human activities at the peak of wildfire occurrence season every year.

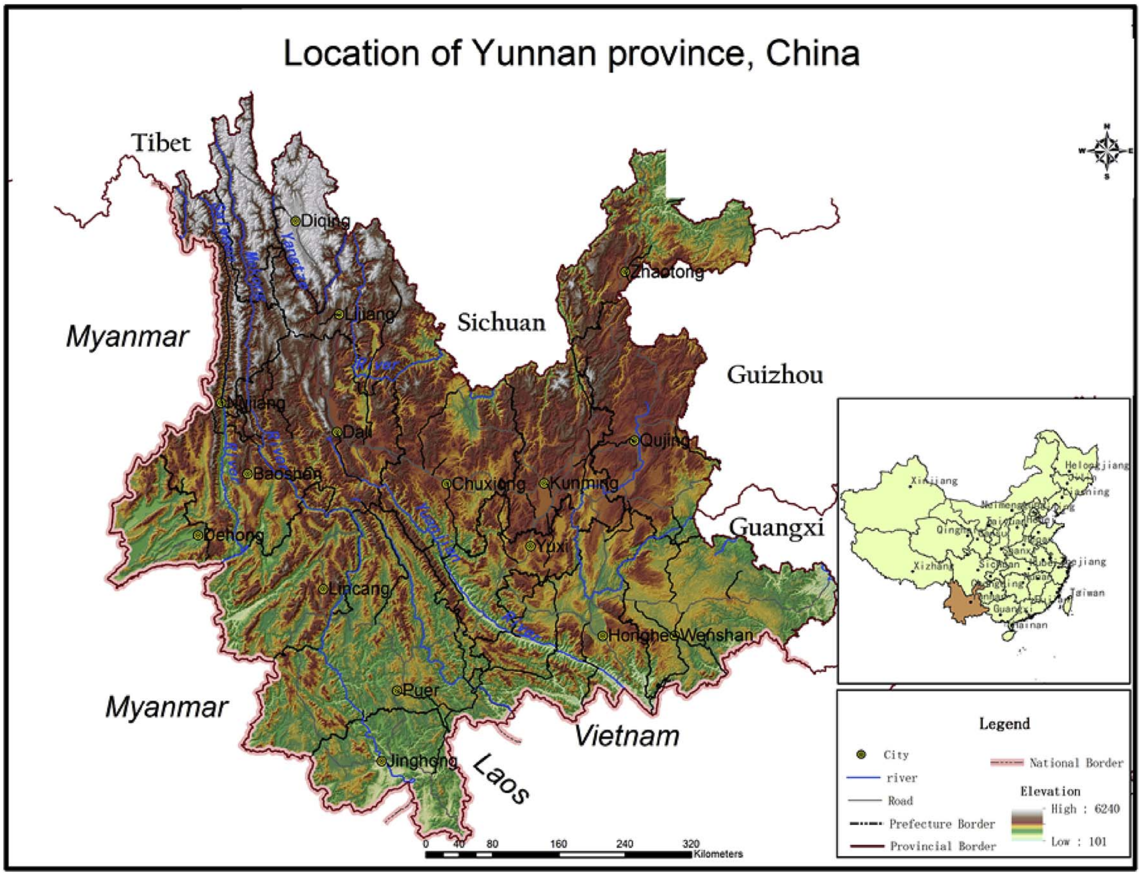

Fig. 1. Location of the study area.

2.2. Wildfire points

A location layer is necessary to serve as the training points in the WofE model computations. Wildfire statistics in 2007–2013 were collected from the wildfire satellite monitoring center, of the State Forest Administration. The data consisted of fire hotspots derived from MODIS, NOAA, and FY remote sensing satellite images collected in the Forest Monitoring System, and later verified as actual fire locations by ground verification, since hotspots can include the heat sources of ongoing active fires. The data showed 15,614 hotspot points depicted as high-temperature sites from 2007 to 2013, with 6343 points identified as active fires. Due to large wildfires that developed dynamically over time, the same fire was recorded repeatedly, so an elimination was conducted manually. Finally, 2374 points were confirmed as wildfire ignition points (Fig. 2), of which 2017 over the previous 6 years were training points and the rest served as validation points.

2.3. Independent variables

According to prior work, the distances to roads and railways, and population density, tourism sites, and farmland are usually considered as human variables that contribute to wildfire ignition (Curt et al., 2016; Fusco et al., 2016; Maingi & Henry, 2007; Zhang et al., 2016). Location layers were collected that included the settlements, transportation, and infrastructure from databases of the *Yunnan Forestry Disaster Provincial Science and Technology Innovative Team*. Considering that different kinds of transportation and settlements lead to distinctive human mobility ranges, the transportation data was categorised into two line layers of *main roads* (national road, highway, provincial road, and railway) and *secondary roads* (township roads and paths), while settlements were divided into *urban settlement* points (city residents, county residents, township residents) and *village* points. Additionally, *water* was considered, usually involving hydropower stations, also tourism sites such as temples leading to high human density. *Farmland* areas were included in which farming results in more human activity, and so fires.

In order to prepare the thematic evidence, proximity maps were generated for 6 layers by a cost distance instead of Euclidean distance, which more closely approximates the concept of the human activity function (DeMers, 2002). Since the human travel cost of accessibility is essentially limitless, we constrained the extent with a corridor, for instance, there is less probability for humans to access a cliff than a plain. Thus, we generated a cost raster to define the impedance based on slope and elevation influence in terms of empirical knowledge, as below.

$$CST = (\sin(SP) \times H^{0.3} + 0.1)^{0.3} \tag{1}$$

where CST is the cost raster, SP is slope, and H is height, based on the NASA Shuttle Radar Topographic Mission 90 m digital elevation model.

2.4. Methodology

The method of WofE modeling is based on Bayes' theorem and on the concepts of prior and posterior probability, and was developed originally for medical diagnosis, but has been applied subsequently to mapping mineral potential (Agterberg & Cheng, 2002). The technique was later extended to determining animal habitats, archaeology, environmental impact evaluation, risk assessment and also used in wildfire prediction (Dickson et al., 2006; Ford, Clarke, & Raines, 2009; Ilia, Tsangaratos, Koumantakis, & Rozos, 2017; Lv, Zheng, Zhao, & Hu, 2013; Raines & Bonham-Carter, 2007; Romero-Calcerrada, Novillo, Millington, & Gomez-Jimenez, 2008, 2010; Tahmassebipoor, Rahmati, Noormohamadi, & Lee, 2016; Weed, 2005). The method assumes that an event occurs at a set of known presence points and uses evidence from multiple evidential themes that might cause the event expressed in

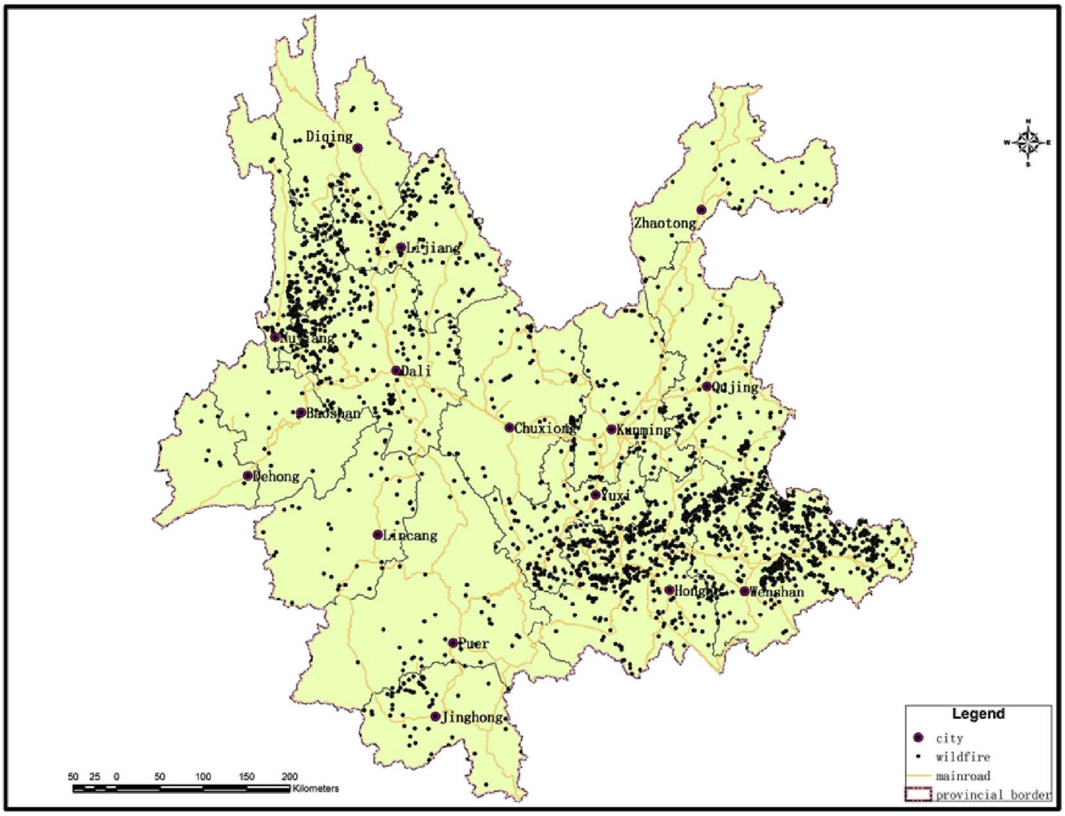

Fig. 2. Map of known wildfire ignitions in Yunnan province, China, during 2007–2013.

binary or in categorical variables. The causative variables are represented in a log-linear model based on an assumption of conditional independence among them and used to compute a probability map. WofE has the advantage of analyzing the relationship between variables and a distribution using a geostatistical method in an objective way instead of using prior knowledge. It has been widely applied and integrated into ESRI's ArcGIS as an extension called ArcSDM (Sawatzky, Raines, Bonham-Carter, & Looney, 2009). The main algorithm is described by Bonham-Carter (Bonham-Carter, Agterberg, & Wright, 1989) as below.

Supposing the total area is partitioned into equal cells, D represents the cells where the wildfire has occurred, then the weight for every evidence layer B can be expressed by W^+ and W^-:

$$W = \begin{cases} W^+ = \ln \frac{P(B \mid D)}{P(B \mid \bar{D})} & \text{Evidence is present} \\ W^- = \ln \frac{P(\bar{B} \mid D)}{P(\bar{B} \mid \bar{D})} & \text{Evidence is absent} \\ 0 & \text{Data is missing} \end{cases} \tag{2}$$

$$C = W^+ - W^- \tag{3}$$

where, $W^+$and W^- are the measurement of positive or negative correlation between event location and evidence. If W^+ is positive and W^- is negative, that indicates more events occur due to an evidential theme than would be expected by chance. Conversely, when W^+ is negative and W^- is positive, implies there are fewer events due to an evidential theme than would be expected by chance (Bonham-Carter et al., 1989). The difference between the two weights gives the contrast C, which can be used to define the optimal thresholds to find the best pattern of evidence. When C equals zero, it indicates that the conditional factor under consideration is not significant for the analysis, i.e. is not significantly different from a random distribution. A negative contrast means a negative correlation (the factor reduces wildfire ignition), and conversely, a positive contrast represents a positive relation or direct causation (Corsini, Cervi, & Ronchetti, 2009; Tahmassebipoor et al., 2016). Since C usually shows uncertainty (Bonham-Carter, 1994), C_S, the Studentized value of C, is more often used than C in the selection of evidence variables and in the breakpoints in their classes (Ilia et al., 2017; Romero-Calcerrada et al., 2010).

Given n evidence factors, the potential distribution of probability for every spatial cell, i.e. the posterior odds, can be depicted as:

$$Oposterior = \exp\left\{ \sum_{j}^{n} W_j^k + \ln(O_{prior}) \right\} \tag{4}$$

Thus the posterior probability is:

$$P_{posterior} = \frac{O_{posterior}}{1 + O_{posterior}} \tag{5}$$

Several steps are involved in the weights-of-evidence calculation:

(1) Select known present points of events as the training samples, such as the wildfire locations of occurrence (points), then reduce the points by thinning to one per cell to satisfy the assumption.
(2) Calculate the weights of evidential variables and generate thematic evidence.

Thematic maps of suspected independent factors to point distribution are assembled and made binary or categorical, with as few classes as possible. Then the weights can be calculated using the training points, and outputs are stored in a cumulative categorical table of the correlation indicators of W^+, W^- and C. These variables are clues reflecting whether the variables and classes are acceptable evidence or not, and also can instruct the reclassification of the themes in the most predictive manner. This procedure is a trial and error process.

(3) Check conditional independence.

WofE is established on the assumption that map layers are approximately conditionally independent of each other, and that conditional independence (*CI*) can be tested by statistical validation between layers in a contingency table and an overall or "omnibus" test (Agterberg & Cheng, 2002; Lv et al., 2013) applied. An overall *CI* ratio (n/T, n is the number of observed training points, and T is expected number of training points), varying from 0 to 1 is usually used. When the value is greater than 0.85 the layer satisfies conditional independence, then the evidence can be accepted and the prediction is valid, otherwise, the evidence violates the assumption and the prediction is invalid (Bonham-Carter, 1994). Another one-tailed significance test uses a null hypothesis, expressed as the Agterberg & Cheng Conditional Independence test in ArcSDM. The test statistic of (T-n)/standard deviation of T, being greater than probability values at the 95% or 99% level of significance indicates that the hypothesis of *CI* should be rejected (Agterberg & Cheng, 2002; Sawatzky et al., 2009).

(4) Calculate posterior probability and generate risk map.

According to the weighted layers and the evidence, the posterior probability map of wildfire risk induced by human activity can be calculated, and then reclassified into a thematic map with 4 risk categories {very high, high, medium, low } based on the appropriate threshold of indicator R (the ratio of posterior to prior) (Carranza & Hale, 2000; Romero-Calcerrada et al., 2010). The method can also produce a binary predictive map with presence and absence zones.

$$R = \begin{cases} 0.5 & \text{Low predictive risk} \\ 0.5{-}0.8 & \text{Medium predictive risk} \\ 0.8{-}1.5 & \text{High predictive risk} \\ >1.5 & \text{Very High predictive risk} \end{cases} \tag{6}$$

(5) Validation of the model

The precision of model can be validated by using wildfire location data with both absence and presence on a binary map of occurrence. We applied with- and out-of-sample validation on training and testing observations, which are 2374 and 357 wildfire presence accounting for 85% and 15% respectively. A model is considered valid when it identifies at least 70% of the occurrences used in developing the model or at least of 50% of "undiscovered" occurrences whose samples are withheld (Carranza & Hale, 2000).

3. Results

3.1. Evidential themes

All explanatory variables were transformed into proximity raster layers as the evidential layers by cost distance computation in 100 m (Fig. 3) and 1000 m increments. From the pattern of human proximity, human activity more often occurs close to settlements, roads, farmland and water. However, different human variables are quite heterogeneous spatially. Owing to Yunnan's topography characteristics with higher in the north and lower in the south, decreasing in steps (Yuming et al., 2004), human accessibility also varies significantly spatially.

As to the proximity to urban settlements, human activity has a quite high accessibility of 5 km in the east and west rather than the north and south, while human activity from villages can access almost all locations within 3 km except in the northwestern canyons. Similar patterns were obtained from transportation proximity, with high accessibility in the center extending to the east due to two major highways of Hukun and Shankun connecting to the neighboring province. Using main roads, humans can travel to each direction of the province, and also easily reach other places within 3 km using secondary roads except for areas of poor accessibility in the east and northwest. For proximity to water, humans move within 5 km near the water across the region along its great rivers and waters. In terms of proximity to farmland, humans can extend to most places within 3 km which reflects a landscape mosaic of agriculture and forest.

3.2. Weight calculation and evidence generation

To strengthen the exploration of the association between wildfire and human activity, all the weight statistics for every variable were calculated significantly in binary levels, as categories and at the different scales (resolutions) of 100 m and 1000 m, shown in Tables 1 To 3. As a result, The 2017 training points led to a prior annual probability of 0.000 052 in every 100 m cell and 0.005 016 in 1000 m. The cutoffs of each binary evidential theme were determined by test reclassification and the highest C_S weight statistics in accumulative proximity chosen (Fig. 4). Fig. 4 shows the variation of studentized value, indicating that the relation alters as proximity increases. Then the weights for the binary evidential themes were calculated (Table 1). From Table 1, all the evidential layers have positive W^+ and negative W^- weights with positive contrasts, which means more wildfire will be predicted considering all these variables, and all these variables were positively associated with wildfire ignition. However, different layers had distinct contrasts, showing various strengths of the causative relationships. Proximity to villages ranks at the top with the greatest contrast of 0.9926 followed by proximity to secondary roads of 0.8197, then proximity to farmland of 0.5720 and proximity to urban settlements of 0.4389 while the proximity to water and main roads have relatively small value contrasts of 0.2825 and 0.2299 respectively. These binary weight statistics show that wildfire ignition is strongly associated with proximity to villages, secondary roads, and farmland.

When calculating the weight for categories, a detailed relation was obtained between specific patterns of human variables with wildfire ignition. From Table 2, studentized contrast (C_S) of each variable fluctuates over space from negative to positive, then negative in diverse classes rather than simply negative or positive trend, which indicates more complication exists and wildfire ignitions more often occur at a particular distance from settlements, roads, and other infrastructure. Despite the variables being proximities, the patterns of the influence on wildfire ignition is quite different. In terms of proximity to urban settlements, distances less than 6 km and more than 20 km shows a negative correlation with wildfire ignition, the distance between them indicates a negative influence on wildfire with less wildfire than there would be expected than by chance accounting for these classes. Wildfire ignition most often occurs at distances from urban settlements between 12 km and 20 km with the highest studentized contrast (C_S) of 7.9270. This is similar for proximity to villages, wildfire ignition favors a distance between 5 km and 8 km with the greatest C_S at 7.7314.

For proximity to main roads, distances between 5 km and 20 km are positively associated with wildfire ignition, showing the highest C_S 5.0645 at 10 km. For proximity to secondary roads, wildfire ignition favors distance from 1 km to 8 km, especially below 2 km with the highest level of C_S at 4.1674. As far as proximity to farmland, positive C_S appears at distances between 1 km and 15 km with the largest at 14.4813 between 3 km and 8 km. For proximity to water, positive association appears at the distance of 3 km and 15 km and the maximum C_S is 6.0506 at a distance less than 8 km.

When changing the scale of each variable to 1000 m, almost the same patterns of weight statistics were obtained compared with 100 m. From Table 3, the positive relationship between wildfire and human activities exists in those areas neither too close nor too far for human proximity, where the region at a medium distance shows a high probability of wildfire ignition. As to the detailed impact of each class, the proximity to urban settlements shows a different pattern with the highest C_S of 5.6711 in the distance range of 6 km–11 km, much closer

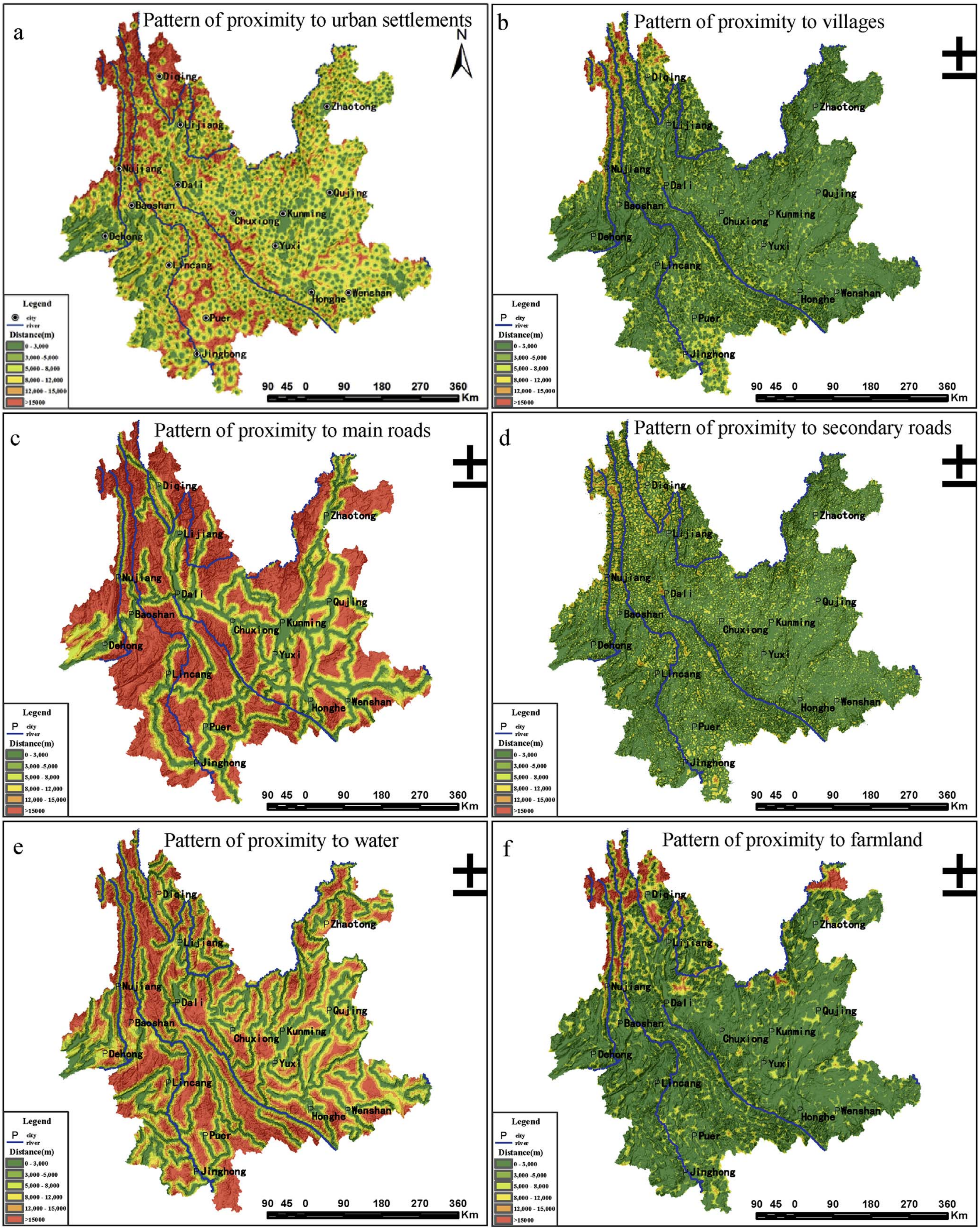

Fig. 3. Human proximity maps: (a)proximity to urban settlements; (b)proximity to villages; (c)proximity to main roads; (d)proximity to secondary roads; (e)proximity to farmland; (f) proximity to water.

Table 1
Summary of weights for evidential themes in binary.

Variables	Proximity (km)	N	W^+	$s(W^+)$	W^-	$s(W^-)$	C	$s(C)$	Cs
Proximity to villages	8	1991	0.0217	0.0224	−0.9709	0.1961	0.9926	0.1974	5.0283
Proximity to urban settlements	20	1952	0.0176	0.0226	−0.4213	0.1240	0.4389	0.1261	3.4808
Proximity to secondary roads	6	2012	0.0031	0.0223	−0.8165	0.4472	0.8197	0.4478	1.8306
Proximity to main roads	19	1371	0.0796	0.0270	−0.1504	0.0393	0.2299	0.0477	4.8183
Proximity to water	19	1942	0.0121	0.0227	−0.2704	0.1155	0.2825	0.1177	2.4002
Proximity to farmland	12	1876	0.0526	0.0231	−0.5195	0.0842	0.5720	0.0873	6.5508

Note: W^+ is positive weights; $s(W^+)$ is stand deviation of W^+; W^- is negative weights; $s(W^-)$ is stand deviation of W^-; The contrast C is the difference between W^+ and W^-; $s(C)$ is stand deviaton of C; C_S is the studentized C.

than those at the 100 m scale. The pattern of the proximity to secondary roads is the same as at 100 m with most possibility of wildfire ignition between 1 km and 8 km with a greatest C_S of 4.9422. This similarity appears in other layers except for small differences of maximum C_S at various distances. According to the weight statistics, different evidence layers were generated for the following WofE models.

3.3. Conditional independence test

Conditional independence was examined for models of different evidence using ArcSDM. Since 6 layers did not satisfy the assumption of conditional independence, we tested several most probable combinations of evidence layers (Table 4). From Table 4, we can see a 3 evidence layer model at 100 m of the proximity to villages, secondary roads and farmland satisfies the conditional independence test. The observed number of training points n is 2017 while the expected number of training points T is 2066.9. The difference is 49.9 and the standard deviation of T is 26.013. The Conditional Independence Ratio is 0.98, which being greater than 0.85, indicates no conditional independence among the causative factors, implying that the evidence can be accepted and the prediction is valid. For the one-tailed significance test, the test statistic is 1.9168 with a probability of 0.972, which is significant given a 99% confidence level and also indicates that the assumption of conditional independence can be accepted. Otherwise, a 4 layer model of proximity to villages, main roads, secondary roads, and farmland is accepted given a 95% confidence level. Also, a 6 layers model in binary meets the conditional independence with the total conditional ratio of 0.99 and the one-tailed test was accepted at the 95% confidence level.

3.4. Wildfire ignition risk

With ArcSDM, we calculated three responses from the binary and categorical evidence themes at 100 m and 1000 m respectively, representing the wildfire ignition risk by the ratio of posterior to prior (Fig. 5), the legends present the probabilities of human wildfire ignition. In addition, a forward stepwise logistic regression was made to compare with the WofE, a three variables model of the proximity to secondary roads, villages and farmland was accepted at the 0.05 significance level.

$$P = \frac{\exp(-1.09 + 0.571a_1 + 0.417a_2 - 0.5141a_3)}{1 + \exp(-1.09 + 0.571a_1 + 0.417a_2 - 0.5141a_3)} \tag{7}$$

In equation (7) P, a_1, a_2, a_3 are predictive dependence, proximity to secondary roads, villages and farmland respectively.

In Fig. 5, the patterns of human wildfire ignition risk show variation across space. The predictive maps at 100 m and 1000 m appear quite similar and different risks scatter spatially, while the binary and logistic regressions show much more area in the very high ignition risk level. The percentage of every risk map differs, low risks are 25.12% (1000m), 28.08% (100m), 14.61% (binary) and 7.54% (logistic) sequentially. Medium risks take the proportion of 215.7%, 21.84%, 26.22% and 28.56% in turn. High risks are 30.57%, 22.34%, 28.48% and 14.07% accordingly, while the very high risk accounts 18.09%, 27.74%, 53.61% and 51.70% respectively. In the result for categorical WofE, there is an obvious high risk in the southwestern and southern underdeveloped region, especially in Lijiang, then scattered high risk in the south of Chuxiong prefecture and Zhaotong. Medium risk regions are mostly distributed in the west and east, while low risk areas exist in the east of Kunming to Qujing, northeast of Zhaotong, west of Dehong and Baoshan. The wildfire risk is obvious aggregated in the WofE binary model, with most areas joining the high risk, especially in Kunming city and Dianchi Lake nearby, while the very high risk for the logistic model is focused in the northwest, even in alpine regions.

3.5. Model validation

An additional non-overlapping sample of 2017 wildfire ignition points was used to validate the models, the accuracy of each model was greater than 70% with the Logistic regression ranked highest at 95.04%. When evaluated with 357 presence and 457 absence additional wildfire ignition points to test the validation, the WofE categorical model had a higher accuracy of wildfire absence but a lower precision of wildfire presence compared to the WofE binary and logistic regression models (Table 5). Note that the logistic regression model has very high precision in both predictive and unknown wildfire presence, but when used with unknown absence, the accuracy dramatically decreased to 30.85%, which consequently impacted the total accuracy. Overall, the WofE model at 1000 m achieved the highest total accuracy of 69.38%, then the WofE model at 100 m at 69.38%, logistic and the WofE in binary with 59.28% and 48.28% respectively. From Table 5, categorical WofE was more useful in identifying at least 70% of occurrence in developing the model and 50% of the "undiscovered" (Carranza & Hale, 2000), which indicates the categorical WofE model at 1000 m categories is most robust in predicting human wildfire ignition risk for the ecoregion.

When overlaying the wildfire ignition risk map with the main vegetation ecosystems, results show that wildfire ignitions in Yunnan mainly occur in four ecosystems: (I) semi-warm wet evergreen broadleaf forest, due to strong human disturbance in the long term, the main vegetation-fuel in this ecosystem is secondary *Pinus yunnanensis*, which mostly exists on the plateaus in middle Yunnan; (II) semi-savanna in the dry-hot valleys, where the main fuel consists of *Quercus acutissima* shrubs and dry-hot herbs located in the valley of the Mekong River, Salween River, Dulong River, Yangtze River and Yuanjiang Rivers; (III) cold temperate alpine dark coniferous forest, where *Abies forresti* is the main fuel scattered in the high mountains of the northwest; and (IV) montane humid evergreen broad leaved forest, whose main fuel is *Pinus kesiya* plantations which developed after the initial vegetation was cleared and mostly exists in the regions of Puer, Lingchang and Jinghong. These locations are consistent with the known wildfires statistics among the training points.

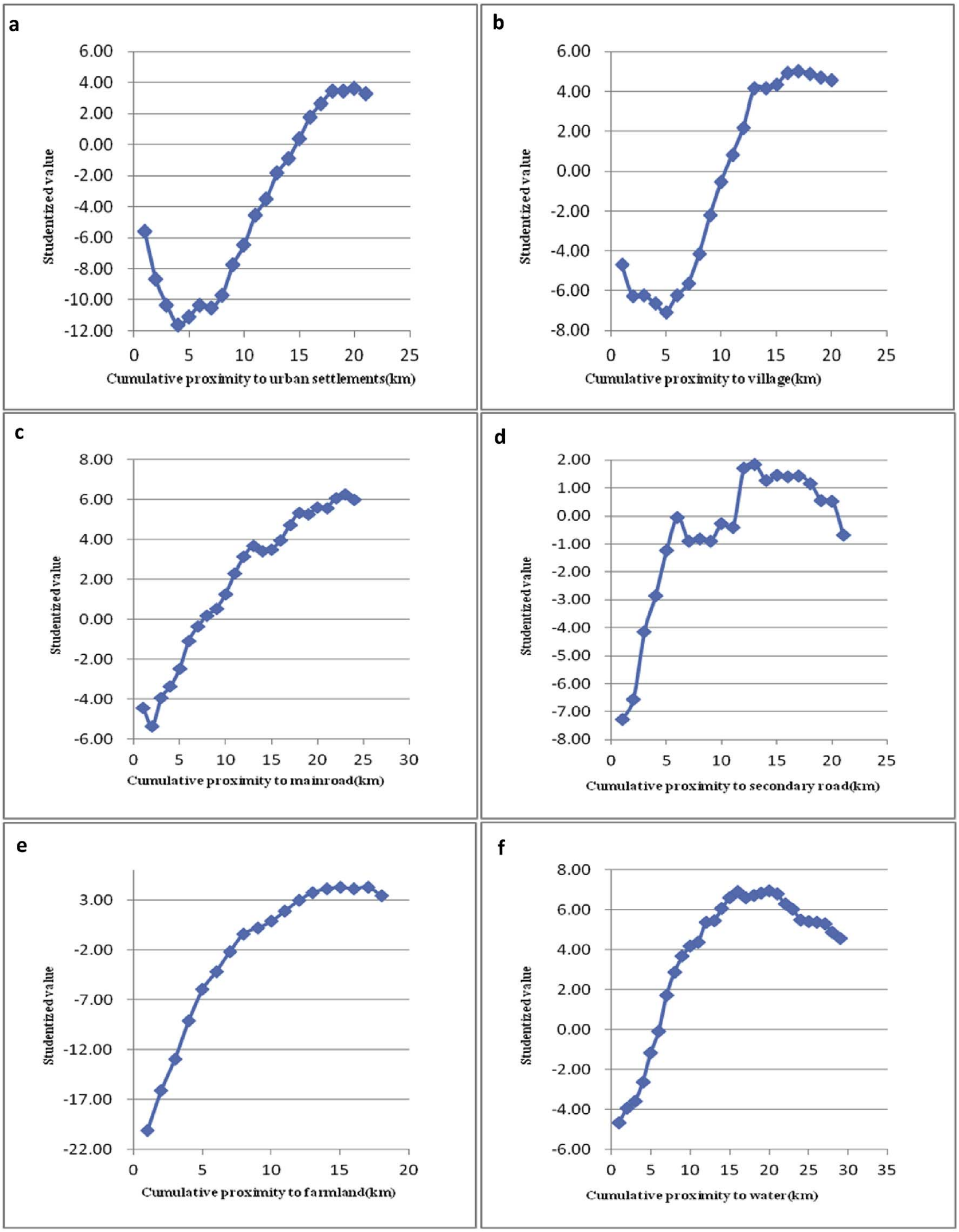

Fig. 4. Variations of studentized value for cumulative proximity to (a)urban settlements; (b)villages; (c)main roads; (d)secondary roads; (e)farmland; (f)water.

4. Discussion

4.1. Determinants of human wildfire ignition

Starting wildfires is one among the multitude of human influences on wildfire regimes with diverse motivations such as accident, arson and farming. Wildfires are mostly ignited by human activity in China, especially in southwest China (Chen et al., 2014, 2017; Zhong et al., 2003). We quantified human accessibility to explore the influence of human activity on wildfire ignition instead of considering more specific measures of human activity, such as population density, which is generally used in wildfire study (Catry et al., 2010; Chang et al., 2013; Guo et al., 2016a; Kwak et al., 2012), but has proven a poor predictor for ecoregions (Fusco et al., 2016). From Table 4, not all of our previously suspected human factors were good predictors of wildfire ignition, our study revealed that proximity to roads, villages and farmland are strong predictors of wildfire ignition, while urban settlements and access to water show poor predictability, which is consistent with other findings (Catry et al., 2010; Dlamini, 2010; Fusco et al., 2016; Gralewicz, Nelson, & Wulder, 2012b; Maingi & Henry, 2007; Vilar, Woolford,

Table 2
Summary of weights for categories of evidential themes in 100 m.

Variables	Proximity (km)	N	W^+	$s(W^+)$	W^-	$s(W^-)$	C	s(C)	Cs
Proximity of urban settlements	0–6	504	−0.3752	0.0445	0.1645	0.0257	−0.5397	0.0514	−10.4946
	6–12	945	0.1464	0.0325	−0.1134	0.0305	0.2597	0.0446	5.8211
	12–20	499	0.3224	0.0448	−0.0867	0.0257	0.4091	0.0516	7.9270
	20–35	68	−0.2989	0.1213	0.0122	0.0227	−0.3111	0.1234	−2.5219
	> 35	1	−2.6391	1.0000	0.0065	0.0223	−2.6456	1.0002	−2.6449
Proximity of villages	0–1	192	−0.4182	0.0722	0.0562	0.0234	−0.4743	0.0759	−6.2519
	1–3	726	−0.0784	0.0371	0.0469	0.0278	−0.1253	0.0464	−2.7006
	3–5	608	0.1471	0.0406	−0.0573	0.0266	0.2044	0.0485	4.2127
	5–8	407	0.356	0.0496	−0.073	0.0249	0.4290	0.0555	7.7314
	8–15	82	−0.3538	0.1104	0.0181	0.0227	−0.3719	0.1127	−3.2985
	> 15	2	−1.9016	0.7071	0.0057	0.0223	−1.9073	0.7075	−2.6960
Proximity of main roads	0–5	392	−0.1553	0.0505	0.0414	0.0248	−0.1966	0.0563	−3.4946
	5–10	428	0.2233	0.0483	−0.0525	0.0251	0.2758	0.0545	5.0645
	10–20	605	0.1752	0.0407	−0.0666	0.0266	0.2418	0.0486	4.9761
	20–50	561	−0.1023	0.0422	0.0424	0.0262	−0.1447	0.0497	−2.9127
	> 50	31	−1.0387	0.1796	0.0289	0.0224	−1.0676	0.181	−5.8982
Proximity of secondary roads	0–1	885	−0.1662	0.0336	0.1523	0.0297	−0.3185	0.0449	−7.0981
	1–2	583	0.1497	0.0414	−0.055	0.0264	0.2047	0.0491	4.1674
	2–4	472	0.1641	0.0460	−0.0452	0.0254	0.2093	0.0526	3.9793
	4–8	75	0.1571	0.1155	−0.0056	0.0227	0.1627	0.1177	1.3828
	> 8	2	−0.9994	0.7071	0.0017	0.0223	−1.0011	0.7075	−1.415
Proximity of farmland	0–1	596	−0.5353	0.0410	0.3523	0.0265	−0.8876	0.0488	−18.1866
	1–3	515	0.3252	0.0441	−0.0909	0.0258	0.4161	0.0511	8.149
	3–8	734	0.4745	0.0369	−0.1957	0.0279	0.6702	0.0463	14.4813
	8–15	145	0.2788	0.0830	−0.0187	0.0231	0.2975	0.0862	3.4507
	> 15	27	−0.8094	0.1925	0.0171	0.0224	−0.8264	0.1938	−4.2654
Proximity of water	0–3	345	−0.1335	0.0538	0.0299	0.0245	−0.1634	0.0591	−2.7604
	3–8	701	0.1934	0.0378	−0.0895	0.0276	0.2829	0.0468	6.0506
	8–15	652	0.0719	0.0392	−0.0326	0.0271	0.1045	0.0476	2.1941
	15–25	284	−0.2063	0.0593	0.0383	0.024	−0.2446	0.0640	−3.8205
	> 25	35	−0.9346	0.1690	0.0277	0.0225	−0.9623	0.1705	−5.6435

Note: W^+ is positive weights; $s(W^+)$ is stand deviation of W^+; W^- is negative weights; $s(W^-)$ is stand deviation of W^-; The contrast C is the difference between W^+ and W^-; $s(C)$ is stand deviaton of C; C_S is the studentized C.

Table 3
Summary of weights for categories of evidential themes in 1000 m resolution.

Variables	Proximity (km)	N	W^+	$s(W^+)$	W^-	$s(W^-)$	C	s(C)	Cs
Proximity to urban settlements	0–2	52	−1.0573	0.1388	0.0534	0.0231	−1.1108	0.1407	−7.8938
	2–6	456	−0.2357	0.0469	0.0859	0.0261	−0.3216	0.0537	−5.9887
	6–11	809	0.1606	0.0353	−0.1018	0.0299	0.2624	0.0463	5.6711
	11–14	305	0.2801	0.0575	−0.0449	0.0249	0.3250	0.0626	5.1908
	14–21	262	0.2592	0.0620	−0.0353	0.0246	0.2944	0.0667	4.4164
	> 21	46	−0.5806	0.1493	0.0190	0.0231	−0.5996	0.1511	−3.9695
Proximity to villages	0–1	188	−0.3934	0.0731	0.0534	0.0240	−0.4468	0.0769	−5.8099
	1–3	442	−0.1357	0.0477	0.0441	0.0260	−0.1798	0.0543	−3.3115
	3–7	1237	0.1731	0.0285	−0.2501	0.0381	0.4232	0.0476	8.8950
	7–12	56	−0.3166	0.1339	0.0112	0.0232	−0.3278	0.1359	−2.4123
	> 12	7	−1.5025	0.4085	0.0110	0.0229	−1.5135	0.4091	−3.6995
Proximity to main roads	0–2	135	−0.4336	0.0862	0.0417	0.0237	−0.4753	0.0894	−5.3172
	2–15	975	0.168	0.0321	−0.1467	0.0324	0.3148	0.0457	6.8939
	15–25	425	0.0865	0.0486	−0.0231	0.0258	0.1096	0.0551	1.9902
	25–50	360	−0.1612	0.0528	0.0409	0.0253	−0.2021	0.0586	−3.4506
	> 50	35	−0.9016	0.1717	0.0266	0.023	−0.9282	0.1732	−5.3589
Proximity to secondary roads	0–1	1028	−0.1331	0.0313	0.1755	0.0333	−0.3086	0.0458	−6.7446
	1–2	593	0.1755	0.0412	−0.0691	0.0274	0.2446	0.0495	4.9422
	2–5	283	0.1761	0.0596	−0.0274	0.0247	0.2035	0.0645	3.1524
	5–8	24	0.3356	0.1931	−0.004	0.023	0.3397	0.1945	1.7465
	> 8	2	−0.9533	0.7078	0.0017	0.0228	−0.9549	0.7082	−1.3485
Proximity to farmland	0–1	384	−0.7688	0.0511	0.3394	0.0255	−1.1082	0.0571	−19.4014
	1–5	1106	0.3513	0.0302	−0.3342	0.0349	0.6855	0.0462	14.8524
	5–8	261	0.4909	0.0622	−0.0590	0.0246	0.5499	0.0668	8.229
	8–14	141	0.3786	0.0845	−0.0246	0.0237	0.4032	0.0878	4.5927
	14–20	30	−0.1727	0.1830	0.0030	0.0230	−0.1756	0.1844	−0.9525
Proximity to water	> 20	8	−1.3453	0.3538	0.0119	0.0229	−1.3572	0.3545	−3.8282
	0–3	87	−0.4828	0.1074	0.0298	0.0234	−0.5126	0.1099	−4.6644
	3–7	773	0.1252	0.0361	−0.0758	0.0295	0.201	0.0466	4.3138
	7–20	970	0.0604	0.0322	−0.0575	0.0324	0.1179	0.0457	2.5828
	> 20	100	−0.6616	0.1001	0.0526	0.0234	−0.7142	0.1028	−6.9452

Note: W^+ is positive weights; $s(W^+)$ is stand deviation of W^+; W^- is negative weights; $s(W^-)$ is stand deviation of W^-; The contrast C is the difference between W^+ and W^-; $s(C)$ is stand deviaton of C; C_S is the studentized C.

Table 4
Summary of conditional independent test of different variables composition

Model	*CI*	*T*	*n*	*T-n*	*Std(T)*	*(T-n)/Std(T)*
6 layers (urban settlements, villages, main roads, secondary roads, water, farmland)	0.92	2182.3	2017	165.3	6.987	23.656
5 layers (urban settlements, villages, main roads, water, farmland)	0.94	2153.2	2017	136.2	10.873	12.528
5 layers (villages, main roads, secondary roads, water, farmland)	0.97	2082.6	2017	65.6	9.437	6.950
5 layers (urban settlements, villages, main roads, secondary roads, farmland)	0.93	2178.4	2017	161.4	11.571	13.946
4 layers (urban settlements, villages, main roads, farmland)	0.94	2147	2017	130.8	17.748	7.369
4 layers (villages, main roads, secondary roads, farmland)	0.97	2074.8	2017	57.8	15.655	3.692
3 layers in 100m (villages, secondary roads, farmland)	0.98	2066.9	2017	49.9	26.013	1.9168

Note: *CI* is Conditional Independence Ratio; *T* is Expected number of training points; *n* is Observed number of training points;*T-n* is Difference of *T* and *n*; *std(T)* is Standard Deviation of *T*; *(T-n)/std(T)* is one-tailed test statistics.

Martell, & Martín, 2010; Ye et al., 2017; Zhang, Zhang, & Zhou, 2010).

Among these determinants, proximity to villages is most highly linked with wildfire ignition. High frequency of ignition and intensity of fires in rural areas occur when residents interact frequently with wild areas for farming and other reasons. Urban areas had fewer ignitions due to lower availability of forest and fuel to start a wildfire, and more organized firefighting (Gonzalez-Olabarria, Mola-Yudego, Pukkala, & Palahi, 2011). With regard to the influence of roads, our results indicate that humans travel further by main roads, then easily access forest areas by secondary roads, thereby increasing the interaction between humans and the wild area surroundings to wildfires. This finding confirms previous studies of the influence of roads on wildfires occurrence (Catry et al., 2010; Dlamini, 2010; Rodrigues et al., 2014). As to proximity to farmland, more wildfire ignitions exist in the area adjacent to farmland, which reflects the local reality of slash-and-burn cultivation, the traditional agricultural practice of crop residue burning and logging to precede the agriculture frontier (Bowman et al., 2011; Chen et al., 2017; Tian et al., 2013). This result stressed others conclusions that

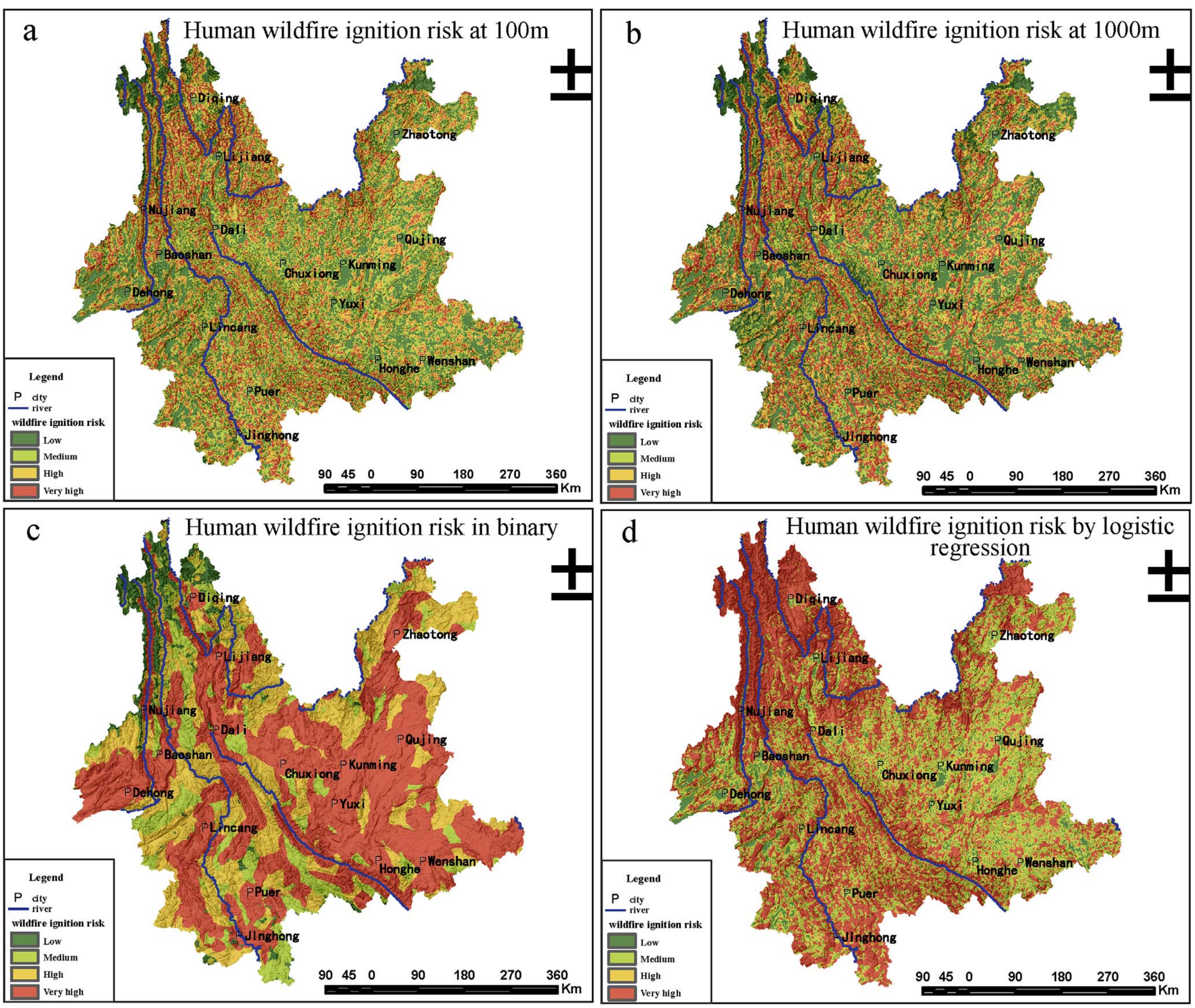

Fig. 5. Map of human wildfire ignition risk:(a)categories WofE at 100m; (b)categories WofE at 1000m; (c)binary WofE; (d)logistic regression.

Table 5
Validation of models.

model	Expected (2017)	Accuracy (%)	Presence (358)	Accuracy (%)	Absence (457)	Accuracy (%)	Total accuracy (%)
WofE (100m)	1544	76.55	214	59.78	333	72.87	66.32
WofE (1000m)	1579	78.28	221	61.73	352	77.02	69.38
WofE (Binary)	1790	88.75	301	84.08	57	12.47	48.28
Logistic	1917	95.04	314	87.71	141	30.85	59.28

wildfire ignitions in these areas are mainly caused by human agricultural practices (Benali et al., 2017; Chen et al., 2017; Tian et al., 2013).

A mosaic distribution of farmland with forest and the dependence on agricultural activities by ethnic minorities make this region a high wildfire risk, continuously troubling to authorities in wildfire management. However, we did not find wildfire associated with accessibility to water to support our initial assumption that more infrastructure was located along water. This may suggest that the relationship is more complicated, on one hand, the proximity to water increases human activities, but on the other the closer to water, the more humidity in soil and thus higher moisture in the fuel load, thereby less likelihood to ignite a wildfire (Dlamini, 2010; Sturtevant & Cleland, 2007; Zhang et al., 2010; Zumbrunnen et al., 2011). So, our findings suggest that a strict control of cropland fire should be adopted to improve wildfire prevention in the rural areas.

4.2. Human wildfire ignition patterns

Our study strengthened the previous finding that the relationship between human activity and wildfire occurrence is not linear (Bowman et al., 2011), as can be seen in Fig. 4. Considering model significance, we note that in Table 2 wildfire ignition patterns vary with distance, implying wildfire-prone regions mostly lie in the area of intermediate proximities to human activities. The results indicate that wildfire occurs more often at a particular distance from infrastructure, and has a low probability near high density human activity where there is a strong capacity for wildfire suppression. This finding disagrees with the common conclusion that wildfire often exists in near to infrastructure (Calef, McGuire, & Chapin, 2008; Ganteaume et al., 2013; Zhang et al., 2016). However, it coincides with previous research showing that the highest fire density is at intermediate levels of population density (Archibald, Roy, WilgenBrian, & Scholes, 2009; Syphard et al., 2007). With regard to proximity to villages, area between 3 and 7 km have high human wildfire ignition probability as distances increase, while when the distance is smaller than 3 km or greater than 7 km, the likelihood decreases as the distance increases. Similarly, wildfire is easily ignited by humans in the proximity of 2 km and 15 km to main roads, 1 km and 2 km of secondary roads and 1 km and 5 km of farmland. This finding is likely due to increasing levels of fire detection and suppression when the proximity of wildfire poses a threat to humans and infrastructure (Fusco et al., 2016), especially with early wildfire detection, consequently large wildfires are rare in these areas. At greater distances, far from villages (> 7 km), main roads (> 25 km), secondary roads (> 8 km), and farmland (> 14 km) the wildfire probability appears pretty low, especially as the distance increases. This may be explained by the reality that the most distant areas are high rugged topography and not accessible to humans. Therefore, it is not surprising that the low wildfire risk area in Fig. 5 (categories WofE in 1000m) is mainly in the southwestern and northeastern high mountains, plus the relatively flat areas, especially Dianchi Lake, also inaccessible to humans. The results also show how wildfire management and fire prevention patrols should proceed. Attention should be paid to the areas of medium proximity to infrastructure instead of directly adjacent to infrastructure where there is high human activity. Distant zones should receive less attention.

4.3. Model predictability

The WofE is a data driven and learning based method, its log algorithm is comparable to conventional logistic regression (Chang et al., 2013; Vilar et al., 2010; Zhang et al., 2010). By comparison, we have found that despite the same three variables being used in the categorical WofE model at 100 m being acceptable, logistic regression overestimated areas with high human wildfire ignition risk, even in the highest snow-covered mountains in the northwestern and northeastern, which contradicts common sense. This finding reflects a low prediction for absence of wildfire in the validation data but high accuracy in presence data, which is in accordance with the demonstration of the regression's predictability but not particularly good (Chang et al., 2013). Categorical WofE is superior to logistic regression with the total accuracy improved by 10%, and WofE also proved to be a useful approach for explicitly considering the spatial association between ignition occurrence and the evidential maps, plus it is relatively straightforward to implement and interpret (Romero-Calcerrada et al., 2008, 2010). We also affirm the conclusion that logistic regression may be subject to problems associated with a lack of conditional independence (Agterberg & Cheng, 2002). As to the independence of WofE, we have noticed the category pattern performed better due to it representing the relationship between dependent and independent variables in a more explicit way. With regard to spatial scale, we assumed that the high resolution would represent the variables with more accuracy. Actually, at 100 m scale, the prior and posterior probabilities were very low, indicating that wildfire ignition occurs in a small number of spatial cells at any specific scale. At detailed scales, probability variation among neighboring cells shows no difference when they cluster at a coarse scale. Our results have demonstrated that 1000 m resolution is more suitable for ecoregion-wide wildfire study (Table 5).

4.4. Limitations

The application of the WofE model has generated a posterior probability map of human wildfire ignition risk, especially the probability of fire caused by human accessibility to potentially flammable areas. Despite human activity not being a sufficient variable, we also note that our results may cause omitted variables bias, but it is impossible to include all the relevant variables in a regression equation and omitted variable bias is therefore unavoidable (Clarke, 2005), to fully study wildfire ignited by humans, it would be necessary to include vegetation-fuel, meteorological, and topographical factors. Additionally, we addressed the cost function to model human accessibility, including some potential variables related to wildfire occurrence, such as slope and elevation. The results may show some bias, and some reasons contribute to this. (1) Geometric error in the known wildfire location. As the wildfire data were extracted from satellite imagery at a coarse resolution of 1 km at nadir, geometric errors exist everywhere in the image and error is expected to be about 2–4 cells of 2383.36–3107.85m (Long, Zhao, Ding, Yang, & Zhou, 2016; Yang et al., 2015). (2)The known wildfires just represent large-scale fires rather than all small fires, or exact locations of fire starts. According to the mechanism of satellite monitoring, hotspots usually are fires with greater area, and those of a small initial fire or with short duration are hard to detect due to their lower heat radiation. Consequently, the map shows the pattern of scale for the wildfire that may be more suited to regional or province-wide management. (3) Samples also show wildfire occurrence rather than frequency, which made the results less capable of showing the spatio-temporal frequency pattern of wildfire.

For temporal scale, we emphasize the result of wildfire is time dependent due to the use of known wildfire over seven years and the evidential themes varying over time. Especially the farmland and roads are continuously changing with the development of the economy and changes in land use. Thus, we highlight that our result represents a static wildfire pattern over time, and snap-shot maps of causative factors.

5. Conclusion

As former studies show, human activity is a key factor in accounting for most wildfire ignition in the southwest of China (Chen et al., 2014, 2017; Zhong et al., 2003). This study analyzed the pattern of human wildfire ignition risk considering human accessibility independently of other fire factors. By applying weights-of-evidence analysis, the relationship between wildfire ignition and the evidential themes was clearly illustrated and the posterior prediction of wildfire shows high accuracy, statistical significance and validity. The reclassification criteria for the evidence layers is also a good reference for other qualified wildfire ignition prediction based on proximity variables. In summary, human wildfire ignition is strongly associated with proximity to villages, roads and farmland, prone areas are those at intermediate distances from high-density human activity rather than too close or far away. The result also reflects the management capacity of wildfire spatially. The method of weights-of-evidence has the advantages of objectivity, simplicity and predictive power compared to other methods (Agterberg & Cheng, 2002). Our results are consistent with what is known about wildfire ignition, especially in the high risk areas of certain ecosystems. There is also some bias due to the training points geometric errors and representing large scale fires rather than small fires due to the use of coarse satellite imagery. Moreover, the results are affected by choice of binary versus categorical evidence and appear to be influenced by the scale of analysis. The results show potential for the support of local government decision making in controlling human induced wildfire ignition and wildfire resources allocation.

Obviously, our results are assumed static and ignition densities are averaged. Nevertheless, the method can represent a baseline of the ignition pattern, with which an integrated, dynamic prediction of wildfire occurrence can be carried out by overlapping temporal variables and other fire factors. Based on the risk evaluation of single variables, analyzing posterior probabilities of every fire factor in order to support comprehensive wildfire prediction should be a goal of future efforts. Furthermore, combining posterior prediction with computer simulation, it is feasible to conduct multi-scale ignition simulation, visualization and wildfire prediction, which will strengthen the practical application of the model.

Acknowledgements

This research was supported by Special Task of Forestry Public Welfare Industry Research (Grant No. 201404402-2), Yunnan Province Innovative Team Project (Grant No. 2014HC014) and Key Founds of Yunnan Educational Administration (Grant No. 2014Z111). Research was conducted while the first author was on sabbatical in the Department of Geography at the University of California, Santa Barbara. We also thank Dr. Mengjun Kang and Dr. Shaohua Wang for their technical supports in ArcSDM operation.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.apgeog.2017.09.012.

References

Agterberg, F. P., & Cheng, Q. (2002). Conditional independence test for weights-of-evidence modeling. *Natural Resources Research, 11*, 249–255.

Archibald, S., Roy, D. P., Wilgen, V., Brian, W., & Scholes, R. J. (2009). What limits fire? An examination of drivers of burnt area in Southern Africa. *Global Change Biology, 15*, 613–630.

Benali, A., Mota, B., Carvalhais, N., Oom, D., Miller, L. M., Campagnolo, M. L., et al. (2017). Bimodal fire regimes unveil a global-scale anthropogenic fingerprint. *Global Ecology and Biogeography, 26*, 799–811.

Bond, W. J., & Keeley, J. E. (2005). Fire as a global 'herbivore': The ecology and evolution of flammable ecosystems. *Trends in Ecology & Evolution, 20*, 387–394.

Bond, W. J., Woodward, F. I., & Midgley, G. F. (2005). The global distribution of ecosystems in a world without fire. *New Phytologist, 165*, 525–538.

Bonham-Carter, G. F. (1994). Geographic information systems for geoscientists-modeling with GIS. *Computer Methods in the Geoscientists, 13*, 398.

Bonham-Carter, G. F., Agterberg, F. P., & Wright, D. F. (1989). Weights of evidence modelling: A new approach to mapping mineral potential. *Statistical Applications in the Earth Sciences, 89*, 171–183.

Bowman, D. M., Balch, J., Artaxo, P., Bond, W. J., Cochrane, M. A., D'antonio, C. M., et al. (2011). The human dimension of fire regimes on Earth. *Journal of Biogeography, 38*, 2223–2236.

Calef, M., McGuire, A., & Chapin, F., III (2008). Human influences on wildfire in Alaska from 1988 through 2005: An analysis of the spatial patterns of human impacts. *Earth Interactions, 12*, 1–17.

Cameron, P. A., Mitra, B., Fitzgerald, M., Scheinkestel, C. D., Stripp, A., Batey, C., et al. (2009). Black saturday: The immediate impact of the february 2009 bushfires in victoria, Australia. *The Medical Journal of Australia, 191*, 11–16.

Cao, Y., Wang, M., & Liu, K. (2017). Wildfire susceptibility assessment in southern China: A comparison of multiple methods. *International Journal of Disaster Risk Science, 8*, 164.

Cardille, J. A., Ventura, S. J., & Turner, M. G. (2001). Environmental and social factors influencing wildfires in the Upper Midwest, United States. *Ecological Applications, 11*, 111–127.

Carranza, E. J. M., & Hale, M. (2000). Geologically constrained probabilistic mapping of gold potential, Baguio district. *Philippines. Natural Resources Research, 9*, 237–253.

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

Chang, Y., Zhu, Z., Bu, R., Chen, H., Feng, Y., Li, Y., et al. (2013). Predicting fire occurrence patterns with logistic regression in Heilongjiang Province, China. *Landscape Ecology, 28*, 1989.

Chas-Amil, M., Prestemon, J., McClean, C., & Touza, J. (2015). Human-ignited wildfire patterns and responses to policy shifts. *Applied Geography, 56*, 164–176.

Chen, F., Fan, Z., Niu, S., & Zheng, J. (2014). *The influence of precipitation and consecutive dry days on burned areas in Yunnan Province, Southwestern China.* Advances in Meteorology 2014.

Chen, D., Pereira, J. M., Masiero, A., & Pirotti, F. (2017). Mapping fire regimes in China using MODIS active fire and burned area data. *Applied Geography, 85*, 14–26.

Clarke, K. A. (2005). The phantom menace: Omitted variable bias in econometric research. *Conflict Management and Peace Science, 22*, 341–352.

Corsini, A., Cervi, F., & Ronchetti, F. (2009). Weight of evidence and artificial neural networks for potential groundwater spring mapping: An application to the Mt. Modino area (northern apennines, Italy). *Geomorphology, 111*, 79–87.

Crisp, M. D., Burrows, G. E., Cook, L. G., Thornhill, A. H., & Bowman, D. M. (2011). Flammable biomes dominated by eucalypts originated at the Cretaceous-Palaeogene boundary. *Nature Communications, 2*, 193.

Curt, T., Fréjaville, T., & Lahaye, S. (2016). Modelling the spatial patterns of ignition causes and fire regime features in southern France: Implications for fire prevention policy. *International Journal of Wildland Fire, 25*, 785–796.

DeMers, M. N. (2002). *GIS modeling in raster.* John Wiley & Sons Incorporated.

Dickson, B. G., Prather, J. W., Xu, Y., Hampton, H. M., Aumack, E. N., & Sisk, T. D. (2006). Mapping the probability of large fire occurrence in northern Arizona, USA. *Landscape Ecology, 21*, 747–761.

Dlamini, W. M. (2010). A Bayesian belief network analysis of factors influencing wildfire occurrence in Swaziland. *Environmental Modelling & Software, 25*, 199–208.

FAO (2007). *Fire management: Global assessment 2006.* FAO Forestry Paper 0258–6150. Food and Agriculture: 151.

Ford, A., Clarke, K. C., & Raines, G. (2009). Modeling settlement patterns of the late classic Maya civilization with bayesian methods and geographic information systems. *Annals of the Association of American Geographers, 99*, 496–520.

Fusco, E. J., Abatzoglou, J. T., Balch, J. K., Finn, J. T., & Bradley, B. A. (2016). Quantifying the human influence on fire ignition across the western USA. *Ecological Applications, 26*, 2390–2401.

Ganteaume, A., Camia, A., Jappiot, M., San-Miguel-Ayanz, J., Long-Fournel, M., & Lampin, C. (2013). A review of the main driving factors of forest fire ignition over Europe. *Environmental Management, 51*, 651–662.

Gonzalez-Olabarria, J. R., Mola-Yudego, B., Pukkala, T., & Palahi, M. (2011). Using multiscale spatial analysis to assess fire ignition density in Catalonia, Spain. *Annals of Forest Science, 68*, 861–871.

Gralewicz, N. J., Nelson, T. A., & Wulder, M. A. (2012a). Factors influencing national scale wildfire susceptibility in Canada. *Forest Ecology and Management, 265*, 20–29.

Gralewicz, N. J., Nelson, T. A., & Wulder, M. A. (2012b). Spatial and temporal patterns of wildfire ignitions in Canada from 1980 to 2006. *International Journal of Wildland Fire, 21*, 230–242.

Guo, F., Innes, J. L., Wang, G., Ma, X., Sun, L., Hu, H., et al. (2014). Historic distribution

and driving factors of human-caused fires in the Chinese boreal forest between 1972 and 2005. *Journal of Plant Ecology, 8*, 480–490.
Guo, F., Su, Z., Tigabu, M., Yang, X., Lin, F., Liang, H., et al. (2017). Spatial modelling of fire drivers in urban-forest ecosystems in China. *Forests, 8*, 180.
Guo, F., Su, Z., Wang, G., Sun, L., Lin, F., & Liu, A. (2016a). Wildfire ignition in the forests of southeast China: Identifying drivers and spatial distribution to predict wildfire likelihood. *Applied Geography, 66*, 12–21.
Guo, F., Wang, G., Su, Z., Liang, H., Wang, W., Lin, F., et al. (2016b). What drives forest fire in Fujian, China? Evidence from logistic regression and Random Forests. *International Journal of Wildland Fire, 25*, 505–519.
Guo, F., Zhang, L., Jin, S., Tigabu, M., Su, Z., & Wang, W. (2016c). Modeling anthropogenic fire occurrence in the boreal forest of China using logistic regression and random forests. *Forests, 7*, 250.
Harrison, S. P., Marlon, J. R., & Bartlein, P. J. (2010). *Fire in the earth system.* Dordrecht: Springer.
Hawbaker, T. J., Radeloff, V. C., Stewart, S. I., Hammer, R. B., Keuler, N. S., & Clayton, M. K. (2013). Human and biophysical influences on fire occurrence in the United States. *Ecological Applications, 23*, 565–582.
Ilia, I., Tsangaratos, P., Koumantakis, I., & Rozos, D. (2017). Application of a Bayesian approach in GIS based model for evaluating landslide susceptibility. Case study Kimi area, Euboea, Greece. *Bulletin of the Geological Society of Greece, 43*, 1590–1600.
Johnston, F. H. (2009). Bushfires and human health in a changing environment. *Australian Family Physician, 38*, 720.
Kwak, H., Lee, W.-K., Saborowski, J., Lee, S.-Y., Won, M.-S., Koo, K.-S., et al. (2012). Estimating the spatial pattern of human-caused forest fires using a generalized linear mixed model with spatial autocorrelation in South Korea. *International Journal of Geographical Information Science, 26*, 1589–1602.
Legendre, P., & Legendre, L. F. (2012). *Numerical ecology.* Elsevier.
Li, S., Hughes, A. C., Su, T., Anberrée, J. L., Oskolski, A. A., Sun, M., et al. (2017). Fire dynamics under monsoonal climate in Yunnan, SW China: Past, present and future. *Palaeogeography, Palaeoclimatology, Palaeoecology, 465*, 168–176.
Long, X.-m., Zhao, C.-m., Ding, K., Yang, K.-y., & Zhou, R.-l. (2016). A method of improving geometric accuracy for forest fire monitoring satellite image based on terrain features. *Remote Sensing Information, 31*, 89–94.
Lv, L.-n., Zheng, X.-q., Zhao, L., & Hu, Y.-c. (2013). GIS-based weight of evidence modeling of basic farmland protection planning for basic farmland suitability mapping. *Journal of Food Agriculture and Environment, 11*, 1087–1092.
Maingi, J. K., & Henry, M. C. (2007). Factors influencing wildfire occurrence and distribution in eastern Kentucky, USA. *International Journal of Wildland Fire, 16*, 23–33.
Malamud, B. D., Millington, J. D., & Perry, G. L. (2005). Characterizing wildfire regimes in the United States. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 4694–4699.
Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature, 403*, 853–858.
Narayanaraj, G., & Wimberly, M. C. (2012). Influences of forest roads on the spatial patterns of human-and lightning-caused wildfire ignitions. *Applied Geography, 32*, 878–888.
Plucinksi, M. (2011). *A review of wildfire occurrence research.* Australia: Bushfire Cooperative Research Centre.
Prestemon, J. P., & Butry, D. T. (2005). Time to burn: Modeling wildland arson as an autoregressive crime function. *American Journal of Agricultural Economics, 87*, 756–770.
Prestemon, J. P., Pye, J. M., Butry, D. T., Holmes, T. P., & Mercer, D. E. (2002). Understanding broadscale wildfire risks in a human-dominated landscape. *Forest Science, 48*, 685–693.
Raines, G. L., & Bonham-Carter, G. (2007). Introduction to special issue on spatial modeling in GIS. *Natural Resources Research, 16*, 81–84.
Rodrigues, M., de la Riva, J., & Fotheringham, S. (2014). Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. *Applied Geography, 48*, 52–63.
Román-Cuesta, R. M., Gracia, M., & Retana, J. (2003). Environmental and human factors influencing fire trends in ENSO and non-ENSO years in tropical Mexico. *Ecological Applications, 13*, 1177–1192.
Romero-Calcerrada, R., Barrio-Parra, F., Millington, J., & Novillo, C. (2010). Spatial modelling of socioeconomic data to understand patterns of human-caused wildfire ignition risk in the SW of Madrid (central Spain). *Ecological Modelling, 221*, 34–45.
Romero-Calcerrada, R., Novillo, C., Millington, J., & Gomez-Jimenez, I. (2008). GIS analysis of spatial patterns of human-caused wildfire ignition risk in the SW of Madrid (Central Spain). *Landscape Ecology, 23*, 341–354.
Salis, M., Ager, A. A., Arca, B., Finney, M. A., Bacciu, V., Duce, P., et al. (2013). Assessing exposure of human and ecological values to wildfire in Sardinia, Italy. *International Journal of Wildland Fire, 22*, 549–565.
Sawatzky, D., Raines, G., Bonham-Carter, G., & Looney, C. (2009). *Spatial data modeller (SDM): ArcMAP 9.2 geoprocessing tools for spatial data modelling using weights of evidence, logistic regression, fuzzy logic and neural networks*.
Simon, M. F., Grether, R., de Queiroz, L. P., Skema, C., Pennington, R. T., & Hughes, C. E. (2009). Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences, 106*, 20359–20364.
Song, W., Wang, J., Satoh, K., & Fan, W. (2006). Three types of power-law distribution of forest fires in Japan. *Ecological Modelling, 196*, 527–532.
Sturtevant, B. R., & Cleland, D. T. (2007). Human and biophysical factors influencing modern fire disturbance in northern Wisconsin. *International Journal of Wildland Fire, 16*, 398–413.
Syphard, A. D., Radeloff, V. C., Keeley, J. E., Hawbaker, T. J., Clayton, M. K., Stewart, S. I., et al. (2007). Human influence on California fire regimes. *Ecological Applications, 17*, 1388–1402.
Tahmassebipoor, N., Rahmati, O., Noormohamadi, F., & Lee, S. (2016). Spatial analysis of groundwater potential using weights-of-evidence and evidential belief function models and remote sensing. *Arabian Journal of Geosciences, 9*, 79.
Tian, X., Zhao, F., Shu, L., & Wang, M. (2013). Distribution characteristics and the influence factors of forest fires in China. *Forest Ecology and Management, 310*, 460–467.
Vilar, L., Woolford, D. G., Martell, D. L., & Martín, M. P. (2010). A model for predicting human-caused wildfire occurrence in the region of Madrid, Spain. *International Journal of Wildland Fire, 19*, 325–337.
Watts, J. M., Jr., & Hall, J. R., Jr. (2016). *Introduction to fire risk analysis. Pages 2817–2826 SFPE handbook of fire protection engineering.* Springer.
Weed, D. L. (2005). Weight of evidence: A review of concept and methods. *Risk Analysis, 25*, 1545–1557.
Xu, H., Yang, X.-p., Zhu, Y., & Yu, L.-x. (2007). Study on the forest fire danger forecast in Yunnan Province based on GIS and RS. *Journal of Fujian Forestry Science and Technology, 2* 020.
Xu, K., Zhang, X., Chen, Z., Wu, W., & Li, T. (2016). Risk assessment for wildfire occurrence in high-voltage power line corridors by using remote-sensing techniques: A case study in Hubei province, China. *International Journal of Remote Sensing, 37*, 4818–4837.
Yang, J., He, H. S., & Shifley, S. R. (2008). Spatial controls of occurrence and spread of wildfires in the Missouri Ozark Highlands. *Ecological Applications, 18*, 1212–1225.
Yang, K.-y., Zhao, C.-m., Ding, K., Xu, S.-j., Li, Q.-y., & Zhou, R.-l. (2015). Method of highly accurate geometric correction for MODIS imagine based on SRTM3. *Journal of Beijing Forestry University, 37*, 130–137.
Ye, T., Wang, Y., Guo, Z., & Li, Y. (2017). Factor contribution to fire occurrence, size, and burn probability in a subtropical coniferous forest in East China. *PLos One, 12*, e0172110.
Yuming, Y., Kun, T., Jiming, H., & Shengji, P. (2004). Yunnan biodiversity and biodiversity conservation. *Biodiversity & Conservation, 13*, 813–826.
Zhang, Y., Lim, S., & Sharples, J. J. (2016). Modelling spatial patterns of wildfire occurrence in South-Eastern Australia. Geomatics. *Natural Hazards Review, 7*, 1800–1815.
Zhang, Z., Zhang, H., & Zhou, D. (2010). Using GIS spatial analysis and logistic regression to predict the probabilities of human-caused grassland fires. *Journal of Arid Environments, 74*, 386–393.
Zhao, F., Shu, L., Tiao, X., & Wang, M.-y. (2009). Change trends of forest fire danger in Yunnan Province in 1957–2007. *Chinese Journal of Ecology, 28*, 2333–2338.
Zhong, M., Fan, W., Liu, T., & Li, P. (2003). Statistical analysis on current status of China forest fire safety. *Fire Safety Journal, 38*, 257–269.
Zumbrunnen, T., Pezzatti, G. B., Menéndez, P., Bugmann, H., Bürgi, M., & Conedera, M. (2011). Weather and human impacts on forest fires: 100 years of fire history in two climatic regions of Switzerland. *Forest Ecology and Management, 261*, 2188–2199.