Current and future patterns of forest fire occurrence in China

Zhiwei Wu\textsuperscript{A,B,C,J}, Hong S. He\textsuperscript{D,E}, Robert E. Keane\textsuperscript{F}, Zhiliang Zhu\textsuperscript{G}, Yeqiao Wang\textsuperscript{H} and Yanlong Shan\textsuperscript{I,J}

\textsuperscript{A}Key Laboratory of Poyang Lake Wetland and Watershed Research, Ministry of Education, Jiangxi Normal University, Nanchang 330022, China.
\textsuperscript{B}School of Geography and Environment, Jiangxi Normal University, Nanchang 330022, China.
\textsuperscript{C}Jiangxi Provincial Key Laboratory of Poyang Lake Comprehensive Management and Resource Development, Jiangxi Normal University, Nanchang 330022, China.
\textsuperscript{D}School of Natural Resources, University of Missouri–Columbia, Columbia, MO 65211-7270, USA.
\textsuperscript{E}USDA Forest Service, Rocky Mountain Research Station, Missoula Fire Sciences Laboratory, Missoula, MT 59808, USA.
\textsuperscript{F}US Geological Survey, Reston, VA 20192, USA.
\textsuperscript{G}Department of Natural Resources Science, University of Rhode Island, Kingston, RI 02881, USA.
\textsuperscript{H}Forestry College, Beihua University, Jilin 132013, China.
\textsuperscript{I,J}Corresponding authors. Email: wuzhiwei@jxnu.edu.cn; shanyl@163.com

Abstract. Forest fire patterns are likely to be altered by climate change. We used boosted regression trees modelling and the MODIS Global Fire Atlas dataset (2003–15) to characterise relative influences of nine natural and human variables on fire patterns across five forest zones in China. The same modelling approach was used to project fire patterns for 2041–60 and 2061–80 based on two general circulation models for two representative concentration pathways scenarios. The results showed that, for the baseline period (2003–15) and across the five forest zones, climate variables explained $37.4–43.5\%$ of the variability in fire occurrence and human activities were responsible for explaining an additional $27.0–36.5\%$ of variability. The fire frequency was highest in the subtropical evergreen broadleaf forests zone in southern China, and lowest in the warm temperate deciduous broadleaved mixed-forests zone in northern China. Projection results showed an increasing trend in fire occurrence probability ranging from 43.3 to 99.9\% and 41.4 to 99.3\% across forest zones under the two climate models and two representative concentration pathways scenarios relative to the current climate (2003–15). Increased fire occurrence is projected to shift from southern to central-northern China for both 2041–60 and 2061–80.

Additional keywords: boosted regression trees, fire probability, MODIS, relative importance.

Introduction
Forest fire is a frequent disturbance that burned $\sim 67 \text{ M ha}$ of forests annually around the world from 2003 through 2012 (van Lierop et al. 2015). The effects of forest fires include short- and long-term changes in structure and function of forest ecosystems (Bond-Lamberty et al. 2007; Liu and Yang 2014). Fire frequency and burned area have substantially increased with prolonged growing seasons under a warming climate (Pitman et al. 2007; Malesky-Malevich et al. 2008; Flannigan et al. 2009). Increases in fire frequency and burned area will pose great challenges to forests and humans (Nitschke and Innes 2008; Carvalho et al. 2010). For example, the Canadian wildfires from 9 to 12 June 2015 produced extensive areas of forest loss and spread of smoke across most of North America (Dreessen et al. 2016).

Forest fires also occur frequently in China (Adams and Shen 2015). According to Chang et al. (2015), there were an annual number of 8182 fires between 1987 and 2007, with an average burned area of 398 197 ha per year. Historical changes in fire frequency and burned area in regions of China are highly influenced by climate variability (Zhao et al. 2009; Wang et al. 2010; Shirazi et al. 2017). China’s climate ranges from tropical to arctic and from wet to extremely dry, with corresponding forest ecosystems from subtropical to boreal ecotones (Piao et al. 2004) and resultant diverse fire regimes (Chen et al. 2017). Smaller fires occur most frequently in the humid and hot climate regions, while larger fires occur in the dry and semiarid regions.
subtropical evergreen broadleaf forests of southern China (Tian et al. 2013). In contrast, larger fires are found in the temperate-to-cool deciduous conifer forests (e.g. Larix) of north-eastern China and usually account for the majority of area burned in the country (Chang et al. 2015). The RCPs (Representative Concentration Pathways) climate scenarios released by Intergovernmental Panel on Climate Change (IPCC) show a trend of increasing temperature across China between 2011 and 2100 by 0.06°C for RCP 2.6 and 0.63°C for RCP 8.5 per decade on average from 11 General Circulation Models (GCMs) (Xu and Xu 2012). Given the projected warming trend, changes in fire provinces (Yang et al. in China has mostly been focused on specific regions or frequency are therefore anticipated throughout the country. Probability as a metric for characterising fire occurrence, which may help policy makers design strategies in response to climate change.

Several studies have suggested the use of fire occurrence probability as a metric for characterising fire occurrence (Pretzler et al. 2004; Catry et al. 2009; Chang et al. 2013; Guo et al. 2016). Fire probability, ranging from 0 to 1.0, is commonly defined as the probability for fires to occur in a given area (e.g. 1-km pixels) and over a defined time period (e.g. 2003–15) considering the effects of climatic and other environmental factors. In the present study, we applied the boosted regression trees (BRT hereafter) approach to model spatial patterns of fire probability under current and future climate scenarios (Liu and Wimberly 2015, 2016). The BRT modelling approach is flexible and does not rely on a priori assumptions of the shape of the response–predictor relationship; such an assumption is difficult for traditional linear regression models to reveal (Elith et al. 2008; Parisien and Moritz 2009).

In the present study, the term fire occurrence is defined as the probability of fire ignition at any location (1-km pixels) in forested lands within current (2003–15) and future (2041–60 and 2061–80) time periods. Our overall objective was to explore national-scale spatial and temporal patterns of fire occurrence under current and future climate change across China. We addressed the following specific issues: (1) how do current fire patterns vary spatially and temporally over China? (2) What are main drivers exerting the most important influence on variability in fire patterns? (3) How would China's forest fires likely respond to various scenarios of future climate?

**Materials and methods**

**Forest zones and fire seasons**

China can be generally divided into five large forest zones: (1) a cold temperate deciduous coniferous forest zone (boreal forests in the Great Xing’an Mountains of north-eastern China, with a low fire frequency but a high average burned area); (2) a temperate deciduous mixed broadleaf–conifer forest zone; (3) a warm temperate deciduous broadleaf–mixed forest zone (low forest coverage and low forest fire frequency); (4) a subtropical evergreen broadleaf forest zone (with a high forest coverage, high fire frequency but low average burned area); (5) a tropical rainforest zone (high forest coverage but low fire frequency) (Guo et al. 2017) (Fig. 1). Fire seasons in China vary by geographical region and forest zone. Generally, north-eastern China (i.e. cold temperate deciduous needle-leaf forest zone and temperate deciduous mixed broadleaf and needle-leaf forest zone) has a bimodal fire season that spans March to June in spring and September to November in autumn. Fire seasons span October to May in northern China (i.e. warm temperate deciduous broadleafed mixed forests zone). There is a long fire season from November to the end of May in the southern part of China (i.e. subtropical evergreen broadleaf forest zone and tropical rainforest zone).

**Data sources and data management**

**Fire occurrence (ignition) data**

We downloaded the Global Fire Atlas dataset for January 2003–December 2015 from the FTP server ftp://fusionftp.gsfc.nasa.gov/fire_atlas/ (accessed 30 December 2018). The fire atlas was produced with support from NASA’s Carbon Monitoring System program. The Global Fire Atlas was developed from the Moderate Resolution Imaging Spectroradiometer (MODIS) Collection 6 MCD64A1 burned area product (Giglio et al. 2018; Andela et al. 2019). In this dataset, clusters of burned area were subdivided into individual fires based on the spatial structure of estimated burn dates in the MCD64A1 burned area product. For each individual fire, the fire dataset contains information on the geographic location (latitude and longitude coordinates) of fire ignition (point), perimeter (polygon) and other information (fire size, duration, daily expansion, fire line length, speed and direction of spread). The MODIS-based fire data here are from a mixture of surface and crown fires. The methodology and validation of the dataset are presented in Andela et al. (2019), while details on the 500-m-resolution burned area product (MCD64A1 collection 6) are described in Giglio et al. (2018). We identified 25 729 forest fires (ignitions) between 2003 and 2015 across China with the MODIS-based Global Fire Atlas dataset, which was less than the number of forest fires recorded in the government’s fire statistics dataset (103 711). The mean fire size and total burned area derived from the MODIS-based Global Fire Atlas dataset were 79.1 and 15.5% higher than those derived from the Chinese government’s fire statistics dataset respectively, but the spatial distributions of forest fires from these two datasets were generally similar (Supplementary material).

The five forest zones covered 97.3% of total forest fires in China. The other 2.7% were sparsely distributed in non-forest dominated regions such as temperate steppes and deserts zones in north-western China. Specifically, the subtropical evergreen broadleaf forest zone in southern China had the highest number of fires (69.8% of the national total), followed by the tropical rainforest zone in southernmost China (15.3%), temperate deciduous mixed broadleaf and needle-leaf forest zone in central parts of north-eastern China (6.3%), then cold temperate deciduous needle-leaf forest zone north-eastern China (5.5%), and the warm temperate deciduous broadleafed forest zone in central-northern China had a low fire occurrence (0.5%).
Forest type data

Data of forest types were from the NASA MODIS Global Land Cover product (MCD12Q1) for the year 2003 at 500-m spatial resolution (https://modis.gsfc.nasa.gov/, accessed 18 January 2019). The forest types included evergreen conifer forests, evergreen broadleaf forests, deciduous conifer forests, deciduous broadleaf forests and mixed forests (Fig. 1). The overall accuracy of the MCD12Q1 product was 50.9–70.2% over China (Yang et al. 2017). Details about the accuracy of the MCD12Q1 product are in the Supplementary material.

Vegetation productivity (NDVI) data

The Normalized Difference Vegetation Index (NDVI) is an indicator of vegetation productivity (Hawbaker et al. 2013; Argañaraz et al. 2015). We therefore used the NDVI to represent spatial variability in vegetation productivity for the period January to December 2003–15. We derived the monthly NDVI data from the NASA MODIS Global MOD13A3 product at 1-km spatial resolution (https://modis.gsfc.nasa.gov/). We calculated the average of 13-year monthly spring (March–May), autumn (September–November) and annual mean (January–December) NDVI for the study period 2003–15.

Topography data

Elevation (m), slope (°) and aspect (Hawbaker et al. 2013) and a terrain-related index (e.g. topographic roughness index) (Stambaugh and Guyette 2008) have been identified as the major topographic variables related to fire occurrence. Given that slope and aspect are more properly expressed at scales less than 1 km and elevation may highly correlate with temperature
and precipitation (Parks et al. 2011), we therefore used a topographic roughness index to explain spatial patterns of forest fire occurrence (Stambaugh and Guyette 2008). We derived a 30-m spatial resolution grid of digital elevation model (DEM) data from the China Geospatial Data Cloud Platform (http://www.gscloud.cn/, accessed 16 June 2019). We created a surface of topographic roughness index with the DEM data based on the ‘tri’ function in the ‘spatiaEco’ package (Evans 2018) of R software (R Core Team 2017).

Infrastructure and population density (human-activity related data)

Anthropogenic factors such as roads, settlements and population density influence fire occurrence by increasing human-initiated ignition probability (Liu et al. 2012). We obtained a Global Roads Open Access Dataset (gROADSv1) covering the period 1980–2010 from the Center for International Earth Science Information Network of the Earth Institute at Columbia University (http://sedac.ciesin.columbia.edu, accessed 5 November 2018). We derived settlement points and population density data in the year 2000 from the Global Rural–Urban Mapping Project (GRUMPv1) (http://sedac.ciesin.columbia.edu). Surfaces of distance to nearest roads or settlements were created by calculating the Euclidean distance from each pixel to the nearest road or settlement with 1-km spatial resolution.

Climate data

Annual and seasonal temperature (°C) and precipitation (mm) are variables used to represent climate conditions while addressing fire characteristics (Zumbrunnen et al. 2009; Chang et al. 2015). We acquired temperature and precipitation data (January 2003–December 2015) with 0.1° × 0.1° spatial and 3-hourly temporal resolutions from the China Meteorological Forcing Dataset (http://westdc.westgis.ac.cn). The national climate dataset was derived from integration of data from 740 Chinese meteorological stations (Chen et al. 2011), and has been used in ecological research in various regions and ecosystems in China (Huang et al. 2016; Dai et al. 2018). We calculated the average of the 13-year spring (March–May), summer (June–August), autumn (September–November), winter (December–February) and annual mean (January–December) temperature and precipitation for the study period of 2003–15.

Future GCM data of annual temperature and precipitation with 5-min resolution of longitude and latitude degrees were obtained from the WorldClim version 1.4 (http://worldclim.org/) (Hijmans et al. 2005). Based on the GCMs outputs in the WorldClim-GLOBAL climate dataset, we selected GFDL-CM3 (Geophysical Fluid Dynamics Laboratory) and GISS-E2-R (NASA Goddard Institute for Space Studies) to capture uncertainties regarding future climate change. The GFDL-CM3 is relatively hot and wet, whereas GISS-E2-R projects a relatively cold and dry future. For each of the two GCMs, we used two scenarios of greenhouse gas concentrations (RCP 2.6 and 8.5) and two periods: 2050 (average for 2041–60) and 2070 (average for 2061–80). The RCP 2.6 scenario assumes that annual global greenhouse gas concentration peaks between 2010 and 2020, and declines substantially thereafter. In the RCP 8.5, concentrations would continue to rise throughout the 21st century (Meinshausen et al. 2011). The future climate projected by the two GCMs in WorldClim version 1.4 had been calibrated using historical climate layers (1960–90). Details on the historical climate layers can be found in Hijmans et al. (2005). We compared historical mean annual temperature and precipitation from WorldClim (1960–90) with observations from 613 benchmark weather stations (1960–90) across China (http://data.cma.cn/site/index.html). Results showed that historical annual temperature and precipitation values from these two datasets were close (Supplementary material), indicating validity of using the GCM data for future climate in the present study.

Mean annual temperature in China would increase significantly under both the RCPs, ranging from 2.2° to 8.5°C across all five forest zones by 2041–60 and 2061–80 (Fig. 2). Most of the forest zones had increasing trends projected in mean annual precipitation under GFDL-CM3 that ranged from 5.8 to 219.4 mm, but only the RCP 8.5 scenario of 2041–60 projected a decreasing trend in forest zones III (−24.5 mm) and IV (−57.7 mm). For the GISS-E2-R, forest zones II and IV had a decreasing trend in mean annual precipitation change ranging from −8.3 to −48.3 mm projected. Other forest zones generally had an increasing trend in mean annual precipitation change ranging from 0.17 to 64.8 mm projected, but only the RCP 8.5 scenario of 2041–60 projected a decreasing trend in forest zones I (−8.4 mm) and III (−24.5 mm), and RCP 2.6 in zone I (−57.7 mm) (Fig. 3).

Analysis

We modelled patterns of forest fires with 1-km spatial resolution. We resampled all the climate, vegetation, topography and human activity layers to 1-km spatial resolution in the ArcGIS environment. We resampled continuous variables (layers) using the bilinear interpolation method and categorical variables using the nearest-neighbour method. The focus of this study was to explore spatial patterns of forest fires. Therefore, the non-forested areas were excluded from data analyses based on the MCD12Q1 product.

We created a fire-point dataset based on the latitude and longitude coordinates of fire ignition location from the Global Fire Atlas dataset. We overlaid the fire-point dataset with the pixels of the 1-km gridded forest-type data (extracted from the MODIS MCD12Q1 product). The pixels of forest-type data with one or more fire occurrences (ignitions) observed were coded as ‘1’, representing fire occurred, and non-fire pixels were coded as ‘0’. The ‘0,1’ dataset was subsequently used as the response (y) variable in BRT models. We randomly sampled the ‘0,1’-coded pixels for each forest zone to construct and validate BRT models. The number of sampling pixels for each forest zone was determined using the equation according to Peduzzi et al. (1996):

$$N = \frac{10 \times k}{p}$$

where $N$ is the number of sampling pixels; $k$ is the number of explanatory variables; $p$ is the proportion of fire pixels in each forest zone.

There were 8713, 11 675, 54 609, 5283 and 3191 sampling pixels for the five forest zones I–V. We randomly separated the
sample-pixel dataset into a training dataset (70%) and a validation dataset (30%) for each forest zone. This subsampling procedure was performed five times for each forest zone, resulting in five random subsamples of datasets. We extracted values of all 22 explanatory variables (Table 1) for these sample pixels for the five forest zones. We used the training dataset to construct BRT models and the validation dataset to evaluate the performance of models for each forest zone.

High correlation between explanatory variables may make variable redundant in a model. It is also important to reduce collinearity to strengthen the interpretation of BRT model outputs (e.g. the relative importance of variables). We used the Generalised Variance Inflation Factor (GVIF) approach to test for possible multicollinearity among the 22 explanatory variables (Fox and Monette 1992). We used the rule of thumb that when the GVIF index >5, then the collinearity of given explanatory variables would be problematic and they should be removed in the iterative process. As the result, we only included nine explanatory variables in subsequent fire occurrence model fitting and data analyses (Table 1).

**Converting fire occurrence to fire probability (BRT model construction)**

We applied the boosted logistic regression trees (BRT) approach to model spatial patterns of fire probability for the entire 13 years (2003–15) based on the training data from the Global Fire Atlas dataset. The BRT model is a form of logistic regression that models the probability of a fire occurring, $y = 1$, at a location with explanatory variables (covariates) $X$, $P(y = 1|X)$. The fire probability is modelled via a logit function: $\text{logit} P(y = 1|X) = f(X)$. To minimise predictive error, we tested several combinations of key fitting parameters of the BRT

---

**Fig. 2.** Changes (mean ± s.d.) in annual temperature (°C) (future minus current) for future RCPs (Representative Concentration Pathways) scenarios (RCP 2.6 and RCP 8.5) under two Global Circulation Models (GFDL-CM3 and GISS-E2-R) in years 2041–60 and 2061–80 by forest zone. Forest zones refer to Fig. 1.
model (learning rate, tree complexity and number of trees) as recommended by Elith et al. (2008) (Table 2). We ran the BRT models five times using the randomly sampled training dataset (70%) for each forest zone. We used the ‘gbm’ package (Greenwell et al. 2018) in R (R Core Team 2017) to construct BRT models referencing the R scripts developed by Elith et al. (2008).

We evaluated the performance of the BRT models using the area under curve (AUC) of a receiver operating characteristic curve (ROC) plot with the randomly sampled validation dataset (30%) for each forest zone. The AUC was used to measure the probability of correctly classifying a random pair of fire and non-fire observations. AUC values varied from 0.5 (random discrimination) to 1.0 (perfect discrimination), and values above 0.8 indicated excellent performance of a model in discrimination (Vilar del Hoyo et al. 2011; Guo et al. 2016). We calculated the AUC values with the ‘ROCR’ package (Sing et al. 2005) in R (R Core Team 2017).

Quantifying relative importance of explanatory variables

We used the BRT models to quantify the relative importance of explanatory variables on fire probability for each forest zone. The BRT model calculated the relative importance of explanatory variables using the formula developed by Friedman (2001). Calculations of a variable’s relative importance in the BRT model were based on how often a variable was selected in splitting a tree (tree node), weighted by the squared improvements to the model as a result of each split, and averaged over all trees (De’ath 2007; Elith et al. 2008). This gave a relative
Table 1. Summary of vegetation, topography, human and climate variables used in explaining and predicting fire occurrence

GVIF (Generalised Variance Inflation Factor) was used to measure the amount of multicollinearity in the explanatory variables. This study used the rule of thumb that when GVIF > 5, then collinearity in the explanatory variable exists and is excluded in the boosted regression tree model construction. NDVI, Normalized Difference Vegetation Index.

<table>
<thead>
<tr>
<th>Variable group</th>
<th>Variable name</th>
<th>Unit</th>
<th>GVIF value</th>
<th>Source and original spatial resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetation</td>
<td>Forest type</td>
<td>Class 1–5</td>
<td>3.3</td>
<td>NASA MODIS Global MCD12Q1 Product (<a href="https://modis.gsfc.nasa.gov/">https://modis.gsfc.nasa.gov/</a>), 500 m</td>
</tr>
<tr>
<td></td>
<td>Mean spring NDVI (March–May)</td>
<td>Range: −1 to 1</td>
<td>3.1</td>
<td>NASA MODIS Global MOD13A3 Product (<a href="https://modis.gsfc.nasa.gov/">https://modis.gsfc.nasa.gov/</a>), 1 km</td>
</tr>
<tr>
<td></td>
<td>Autumn NDVI (September–November)</td>
<td>Range: −1 to 1</td>
<td>&gt;5.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Annual NDVI (January–December)</td>
<td>Range: −1 to 1</td>
<td>&gt;5.0</td>
<td></td>
</tr>
<tr>
<td>Topography</td>
<td>Topographic roughness index</td>
<td>Dimensionless</td>
<td>1.3</td>
<td>China Geospatial Data Cloud Platform (<a href="http://www.gscloud.cn/">http://www.gscloud.cn/</a>), 30 m</td>
</tr>
<tr>
<td>Human</td>
<td>Distance to nearest road</td>
<td>km</td>
<td>1.5</td>
<td>Center for International Earth Science Information Network (<a href="http://sedac.ciesin.columbia.edu">http://sedac.ciesin.columbia.edu</a>), shape files</td>
</tr>
<tr>
<td></td>
<td>Distance to nearest settlement</td>
<td>km</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Population density</td>
<td>No. people km(^{-2})</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>Climate</td>
<td>Max. spring temperature (March–May)</td>
<td>°C</td>
<td>&gt;5.0</td>
<td>China Meteorological Forcing Dataset (<a href="http://westdc.westgis.ac.cn">http://westdc.westgis.ac.cn</a>), 0.1° × 0.1°</td>
</tr>
<tr>
<td></td>
<td>Min. spring temperature (March–May)</td>
<td>°C</td>
<td>&gt;5.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Max. summer temperature (June–August)</td>
<td>°C</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Min. summer temperature (June–August)</td>
<td>°C</td>
<td>&gt;5.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Max. autumn temperature (September–November)</td>
<td>°C</td>
<td>&gt;5.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Min. autumn temperature (September–November)</td>
<td>°C</td>
<td>&gt;5.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Max. winter temperature (December–February)</td>
<td>°C</td>
<td>&gt;5.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Min. winter temperature (December–February)</td>
<td>°C</td>
<td>&gt;5.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Annual temperature (January–December)</td>
<td>°C</td>
<td>&gt;5.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean spring precipitation (March–May)</td>
<td>mm</td>
<td>&gt;5.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean summer precipitation (June–August)</td>
<td>mm</td>
<td>&gt;5.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean autumn precipitation (September–November)</td>
<td>mm</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean winter precipitation (December–February)</td>
<td>mm</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean annual precipitation (January–December)</td>
<td>mm</td>
<td>&gt;5.0</td>
<td></td>
</tr>
</tbody>
</table>

measure of variable importance in the BRT model. The relative importance of each variable was scaled so that the sum added to 100; a higher number indicated stronger influence on fire occurrence (Elith et al. 2008).

Projection of future fire changes

Future fire probabilities were projected using the BRT model for 2041–60 and 2061–80 under the two GCM models and two RCP scenarios and assuming stable vegetation and other environmental and human variables. Relative changes of fire occurrence probability for each 1-km pixel between the baseline (current) and future years were calculated as:

$$\Delta P_{\text{change}} = \frac{P_{\text{future}} - P_{\text{current}}}{P_{\text{current}}} \times 100\%$$

where $P_{\text{future}}$ and $P_{\text{current}}$ represent fire probability for future years (2041–60 and 2061–80) and current baseline (2003–15) respectively; $\Delta P_{\text{change}}$ represents changes of fire probability between current and future climate scenarios.

Results

Model valuation

The baseline fire models (BRTs) were evaluated against the AUC statistic, and the evaluation showed satisfactory results in discriminating presence or absence of fire on a pixel basis. The AUC values ranged between 0.831, 0.887, 0.820, 0.702 and 0.761 for forest zones I, II, III, IV and V respectively (Fig. 4). As noted before, an AUC value of 0.5 represents random assignment whereas 1.0 represents perfect discrimination.

Relative importance of explanatory variables in the models

We ranked explanatory variable groups by forest zones according to their total relative contribution (relative importance) (Table 3). Specifically, according to the relative importance values (%) calculated with the BRT models for explanatory variables by forest zones, the set of climate variables ranked first in forest zones I (41.8%), II (37.1%), III (40.5%), IV (43.1%) and V (43.5%). The human variable group was another important contributor and ranked second in forest zones I (33.6%), II (32.8%), III (27.0%), IV (31.9%) and V (36.5%). Vegetation (10.4–23.9%) and topography (9.3–10.3%) variable groups generally ranked low in terms of their contributions to fire occurrence. For individual variables, we found that temperature, precipitation, population density, distance to settlement and vegetation productivity (NDVI) were commonly ranked as the most important variables, whereas variables of forest type and topographic roughness index were less important in most forest zones in China (Fig. 5).

Spatial patterns under current climate

The fire probability, as shown on the Fig 6b, developed using the BRT model ranged from 0.0006 to 0.7726 with a median value of 0.0118 across the entire forested land of China between 2003
These results indicate that within a 1-km pixel, there would be as low as 0.06% and as high as 77.3% chance of a fire occurring over the 13 years of study. We found considerable variability in the estimated probability of fire occurrence among and within the five forest zones. Predicted high probability was generally distributed in north-eastern and southern China. Central-northern China was identified as having lower probability of fire occurrence throughout (Fig. 6b). Moreover, the BRT model-predicted fire patterns generally agreed with the MODIS observed fires (Global Fire Atlas dataset) (Fig. 6a).

### Spatial patterns under future climate

In general, future fires projected under climate change scenarios had similar distributions (fires mainly located in southern and north-eastern China) compared with the baseline, but with increased values of fire occurrence probability (Fig. 7). In the results, areas with projected high probability of fire appear to have shifted from southern to central-northern China, which was more significant under the GFDL-CM3 climate scenarios (Fig. 8).

Specifically, the percentage of pixels with an increasing trend in fire probability ranged from 43.3 to 99.9% under the GFDL-CM3 scenarios, and 41.4 to 99.3% under the GISS-E2-R scenarios compared with the current climate. Within

<table>
<thead>
<tr>
<th>Forest zones</th>
<th>Samples</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Family</td>
</tr>
<tr>
<td>Zone I</td>
<td>Sample 1</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 2</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 3</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 4</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 5</td>
<td>Bernoulli</td>
</tr>
<tr>
<td>Zone II</td>
<td>Sample 1</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 2</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 3</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 4</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 5</td>
<td>Bernoulli</td>
</tr>
<tr>
<td>Zone III</td>
<td>Sample 1</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 2</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 3</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 4</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 5</td>
<td>Bernoulli</td>
</tr>
<tr>
<td>Zone IV</td>
<td>Sample 1</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 2</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 3</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 4</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 5</td>
<td>Bernoulli</td>
</tr>
<tr>
<td>Zone V</td>
<td>Sample 1</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 2</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 3</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 4</td>
<td>Bernoulli</td>
</tr>
<tr>
<td></td>
<td>Sample 5</td>
<td>Bernoulli</td>
</tr>
</tbody>
</table>

Fig. 4. Area under curve of the receiver operating characteristics curve for boosted regression tree models with the 30% validation dataset by forest zone. Forest zones refer to Fig. 1.
GF DL-CM3, the highest fire probability increase (by 99.9%) was in zone I (RCP 8.5 in 2061–80) and lowest (by 43.3%) in zone V (RCP 2.6 in 2061–80). For GISS-E2-R, the highest increase (99.3%) was seen in forest zone I (RCP 8.5 in 2061–80) and lowest (41.4%) in zone V (RCP 2.6 in 2041–60) (Table 4).

For each RCP scenario by forest zone and GCM model, we divided the percentage of pixels with an increasing trend of fire occurrence probability by forest zones. Forest zones refer to Fig. 1. NVDI, Normalized Difference Vegetation Index.
occurrence probability in 2061–80 by percentage of pixels with an increasing trend of fire occurrence probability in 2041–60 (Table 4). We found that there were no remarkable differences in changes of fire probability between 2041–60 and 2061–80 for both GFDL-CM3 and GISS-E2-R scenarios. Specifically, for both the GCM models and both RCP scenarios, the percentage of pixels with an increasing trend of fire occurrence probability in 2061–80 was 0.87–1.15 times that in 2041–60 (Table 4).

Discussion
Effects of explanatory variables
Forest fires in China were not randomly distributed but showed an aggregative distribution pattern, mainly located in the south and north-east parts of China, demonstrating that fire occurrence is not a random process, but exhibits a high degree of clustering on landscapes and regionally, driven by climate (Tian et al. 2013; Chang et al. 2015). This situation has been observed in other countries such as the United States (Hawbaker et al. 2013) and Australia (Russell-Smith et al. 2007).

Spatial patterns of forest fires are a function of numerous influencing factors such as climate, topography, vegetation and human activity. Among the factors considered in this study in China, we found that climatic factors have the greatest influence on patterns of fire occurrence, which is in agreement with results from an extensive body of fire-science studies (Pitman et al. 2007; Flannigan et al. 2009; Liu et al. 2012), such as wildfires of United States (Hawbaker et al. 2013) and Australia (Russell-Smith et al. 2007). In the present study, it is evident that higher fire probabilities are correlated strongly with increased precipitation (rainfall), which favours vegetation growth (i.e. increasing biomass production) if there are also related high temperature (drier fuels) conditions in the region (O’Donnell et al. 2011; Zhang and Lim 2019). However, there are some other underlying causes that shape fire patterns in China. Previous studies have suggested that fire patterns may be associated with weather events such as thunderstorm activity that are related to precipitation and warm temperatures, and could provide a mechanism for starting fires (e.g. lightning fires). For example, Liu et al. (2012) suggested that areas with the greatest chance of lightning fires were distributed in the northern part of forest zone I (cold temperate deciduous coniferous forests in north-eastern China), which coincides with the highest predicted increases of annual temperature and precipitation.

We found that human activity variables (e.g. distance to nearest settlements and roads) were also important in shaping fire patterns. These findings reinforce claims that, despite the strong influence of climate, effects of human activities cannot be ignored (Syphard et al. 2007; Achard et al. 2008; Ganteaume et al. 2013). For example, Russell-Smith et al. (2007) found that most fires in Australia appears to be anthropogenic, especially in the northern wet–dry tropics and arid Australia. In the United States, human-related variables were ranked highest in explaining fire occurrence patterns in the Central Plains, the Mixed-wood Plains, and the Ozark Ouachita Appalachian forests (Hawbaker et al. 2013). Some studies have suggested that fire occurrence was high when the distance to settlements or roads is low (Romero-Ruiz et al. 2010). For example, Catry et al. (2009) reported that ~98% of fires occurred less than 2 km from the nearest roads in Portugal. In China, for example, a large number of people reside in low-relief plains in forest zone II (the temperate deciduous mixed broadleaf–conifer forests) and have a significant influence on regional fire regimes, especially from their farming activities, as they tend to burn crop residues before planting (Zhang et al. 2015). Forest zone III is characterised by low forest coverage and low forest fire frequency, but these areas may burn if given a chance, particularly under higher human activity and warmer climate conditions. It is worth noting that in China, fires in areas with high population density are easier to detect and suppress in time once they occur, and consequently the burnt area is usually smaller than in areas with low population density, such as in forest zone I (cold temperate deciduous coniferous forests zone).

Under current conditions, fires were most frequent in the evergreen broadleaf forest zone (zone IV) located in the southern part of China. In this region, fires occur in coniferous forests,
such as Masson pine (*Pinus massoniana*) and Chinese fir (*Cunninghamia lanceolata*) forests (Pan et al. 2013). The high frequency of fires in this zone is related to the strong effects of climate and human activity. In southern China, strong and dry valley winds occur frequently in early spring each year, which thus leads to favourable conditions for fire ignitions (Chang et al. 2015). Moreover, road networks are well developed and population density is high in this region. Agricultural cultivation is also near settlements and roads in this region. Fire is a popular tool for agricultural activities (e.g. burning grasses), and
agricultural fires have escaped and burned adjacent forests frequently in this region (Tian et al. 2013).

Compared with climate and human activities, forest types in China have a relatively small role in our model in explaining fire patterns across all five forest zones. With strong effects from climate and human activities, it is possible that effects by forest types are already explained in effects of climate (e.g. temperature and precipitation) and human activities. Moreover, we...
modelled fire occurrence patterns using forest type data at 1-km spatial resolution, which may not capture spatial variation of forest type well.

Effects of climate change on future fire patterns

The GFDL-CM3 projections are relatively warmer (higher temperature) and wetter (higher precipitation) than those of GISS-E2-R (Figs 2 and 3). Normally, wetter climate projections should lead to decreased fire activities. However, our results showed that the GFDL-CM3 scenarios projected greater increases in fire occurrence than the GISS-E2-R scenarios (Table 4). This supports the claim that an increase in temperature would greatly offset increases in precipitation at a broad scale (Boulanger et al. 2013), particularly in forest zone I (cold temperate coniferous forests in north parts of north-eastern China). In a forest landscape, higher temperatures can contribute to increasing transpiration and thus decreasing moisture content of live fuels, leading to an increase in the probability of fire occurrence. Moreover, previous studies have shown that warmer climates could extend vegetation growing seasons and increase plant production and dry fuel availability in late autumn, and consequently fire occurrence would increase substantially. For example, warmer climates and longer vegetation growing seasons in recent decades explain much of the large fire occurrence patterns in the western United States (Westerling et al. 2006; Westerling 2016).

Our results showed that some regions in China such as forest zones I (north parts of north-eastern China) and III (central-northern China) would experience significant increases in fire probability in future years compared with the current period, whereas other areas (e.g. most areas of forest zone V) would remain stable. The regional variability is generally consistent with the relative importance of the explanatory variables under baseline conditions (2003–15). For example, forest zone I (i.e. boreal forests of north-eastern China) is projected to have future higher temperatures (Fig. 3) compared with current conditions; at the same time, 78.9–99.3% of the zone is projected to have increased fire probabilities in 2041–60 and 2061–80. This is similar to recent findings in the literature such as Liu et al. (2012) where fire density (number of fires per 1000 km² per year) in the boreal forests of China is reported to increase 30–230% in 2081–2100. But we also found that the effect of climate change on fire occurrence was constrained by human activity. For example, although temperature would increase across all scenarios in both 2041–60 and 2061–80 in forest zone V (southern China), only 41.4–59.6% of pixels would show an increasing trend. In southern China, agricultural cultivation is mixed with developed lands and fire is a popular tool for agricultural activities (e.g. burning grasses on wasteland). Agricultural fire use has become a major source of forest fires. We therefore assume that the temperature effect was constrained by human activity effects, which explained 36.5% of the spatial variation of fire occurrence in forest zone V.

Limitations and uncertainties

Long-term and large-scale data are important for projecting responses of wildland fires to climate change. However, such historical data are usually unavailable or inconsistent, which may lead to limitations and uncertainties in our study. Fire occurrences were modelled together regardless of fire types (e.g. lightning-caused fires and human-caused fires) in this study. Lightning-caused fires occur more often in isolated and high-elevation areas, whereas human fires show different distribution patterns as they are most frequent at lower elevations with high population density. Therefore, the relative importance of the explanatory variables (e.g. vegetation and topographic parameters in this study) may be different if lightning-caused fires and human-caused fires were modelled separately.

We used the 13-year temperature and precipitation data to build the fire occurrence model. Short-term climate data may not capture long-term variability of the fire–climate relationship (Hawbaker et al. 2013), which may affect the comparability of fire occurrence under current (2003–15) and future climates (2041–60 and 2061–80). Short-term studies should be carefully examined when extrapolating the results to long-term fire–climate relationships. Furthermore, given that fires tend to occur following drought events, it would be valuable in such a study to include antecedent climate, e.g. drought conditions for half a month before fire ignition.

We assumed human activity and vegetation during 2041–60 and 2061–80 would be similar to the present (2003–15), but fire patterns could be a function of changes in vegetation composition, structure and distribution because of climate change (Mitchell et al. 2014; Keane et al. 2018). The changed vegetation structures and patterns could affect fire regimes in future (Pausas and Bradstock 2007) in different relationships, which would in turn alter the composition and structure of forests. The next step therefore would be to incorporate vegetation dynamics models to address feedback effects of vegetation to fire caused by a changing climate (Flannigan et al. 2005).

Conclusions

Forest fires primarily occurred in southern, south-western and north-eastern China during the years between 2003 and 2015. Climate variables had primary effects on spatial variability of
fire occurrence; human activities were secondary. The modeling results showed that, under future climate scenarios, the percentage of pixels with an increasing trend in fire probability ranged from 41.4 to 99.9%. High-fire-occurrence regions would shift from south to central-north parts of China for both 2041–60 and 2061–80. This research will aid in providing a national-scale understanding of future potential fire patterns in China and help policymakers to design fire management strategies to mitigate potential risks.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

This research was funded by the National Natural Science Foundation of China (grant no. 31570462) and the Scientific and Technological Research Project of Jiangxi Provincial Department of Education (grant no. GJJ160275).

References


Evans JS, Ram Karthik (2018) spatialEco. R package version 0.1.1–1. Available at R Core Team https://CRAN.R-project.org/package=spatialEco [Verified 20 June 2019]


Huang Y, Salama MS, Su ZB, van der Velde R, Zheng DH, Krol MS, Hoekstra AY, Zhou YX (2016) Effects of roughness length parameterizations on regional-scale land surface modeling of alpine grasslands in...
Forest fire patterns in China under climate change

Int. J. Wildland Fire


