CSIRO PUBLISHING

International Journal of Wildland Fire
http://dx.doi.org/10.1071/WF13136

Controls on the spatial pattern of wildfire ignitions in Southern California

Nicolas Faivre[A,C], *Yufang Jin*[B], *Michael L. Goulden*[A] *and James T. Randerson*[A]

[A]Department of Earth System Science, University of California, 2101 E Croul Hall, Irvine, CA 92697-3100, USA.
[B]Department of Land, Air and Water Resources, University of California, Davis, CA 95616-8627, USA.
[C]Corresponding author. Email: nfaivre@uci.edu

Abstract. Wildfire ignition requires a combination of an open spark, and suitable weather and fuel conditions. Models of fire occurrence and burned area provide a good understanding of the physical and climatic factors that constrain and promote fire spread and recurrence, but information on how humans influence ignition patterns is still lacking at a scale compatible with integrated fire management. We investigated the relative importance of the physical, climatic and human factors regulating ignition probability across Southern California's National Forests. A 30-year exploratory analysis of one-way relationships indicated that distance to a road, distance to housing and topographic slope were the major determinants of ignition frequency. We used logistic and Poisson regression analyses to model ignition occurrence and frequency as a function of the dominant covariates. The resulting models explained ~70% of the spatial variability in ignition likelihood and 45% of the variability in ignition frequency. In turn, predicted ignition probability contributed to some of the spatial variability in burned area, particularly for summer fires. These models may enable estimates of fire ignition risk for the broader domain of Southern California and how this risk may change with future population and housing development. Our spatially explicit predictions may also be useful for strategic fire management in the region.

Additional keywords: biophysical drivers, fire frequency, fire ignition, human influence, Mediterranean ecosystems, spatial regression model, wildland fire risk.

Received 21 August 2013, accepted 7 May 2014, published online 28 July 2014

Introduction

Wildland fire regimes in Southern California are influenced by climate, ecosystem properties, the rate of human-caused ignition and fire suppression. More than 90% of the fires in Southern California are human ignited and rapid response often extinguishes ignitions that could otherwise become large wildfires (Keeley 1982). Previous studies have reported that the frequency of small fires in coastal Southern California increased during the late 20th century (Keeley *et al.* 1999). Increasing population size and an expansion of housing into fire-prone wildland areas (Hammer *et al.* 2007) has increased ignition risk (Keeley and Fotheringham 2001), and when coupled with severe fire weather (e.g. Santa Ana winds), has resulted in several recent catastrophic fire episodes (Keeley *et al.* 2009). Most ignitions lead to small fires with relatively insignificant effects (Strauss *et al.* 1989). Approximately 90% of fire ignitions accounted for only 1% of the total area burned between 1980 and 2010 in Southern California, whereas only 3% of ignitions led to fires larger than 400 ha and accounted for 96% of area burned during this period (USDA Forest Service 2010). A combination of altered ignition frequencies, changing climate and expansion of housing near wildland areas has modified fire risk in Southern California and presents a serious threat to human lives and property (Syphard *et al.* 2012, 2013).

Landscape-scale disturbances such as fire contribute to the distribution of vegetation and habitats across California (Callaway and Davis 1993). Fire is a natural process whose occurrence and magnitude are regulated by environmental and ignition agents. A location's fire environment, as determined by fuel, weather and topography, affects the occurrence and spread of fire (Countryman 1972). Topographic factors such as elevation, slope and aspect influence fuel characteristics including moisture content, and thus indirectly control fire occurrence and behaviour (Agee 1993; Pyne *et al.* 1996). Under normal weather conditions, the propensity to burn is fuel dependent and is controlled by the amount, arrangement and physical characteristics of vegetation (Whelan 1995). Extreme weather, including strong winds or extended periods with low humidity, over-ride this fuel dependency and dramatically increase fire risk in Southern California (Keeley and Zedler 2009; Moritz *et al.* 2010; Jin *et al.* 2014). Weather and climate, especially precipitation and temperature, affect moisture content and thus the flammability of both live and dead plant material (Verdú *et al.* 2012). Wind speed, relative humidity and air temperature control fire spread rate and direction, and thus possible future changes in weather and climate have the potential to modify wildfire risk (Cayan *et al.* 2008). Climate models predict a hotter and drier climate throughout California by the mid to late 21st century (Cayan *et al.* 2010; Pan *et al.* 2011) and recent studies have concluded that warmer temperatures are likely to increase the duration and intensity of the wildfire season (Westerling 2006; Westerling and Bryant 2008).

Existing models have emphasised the hydro-climate and biophysical controls on fire (Bradstock *et al.* 1998; Preisler *et al.* 2004; Spracklen *et al.* 2009; Westerling *et al.* 2009), and there is a growing understanding of the relationship between humans and ignitions patterns (Syphard *et al.* 2008; Bar Massada *et al.* 2009; Narayanaraj and Wimberly 2012). Southern California has an extensive wildland–urban interface (WUI), which accounts for nearly 60% of the landscape (Hammer *et al.* 2007). Humans influence fire in several ways, including ignition, landscape fragmentation and fire suppression; this complexity complicates efforts to predict fire risk (Perry 1998). Information on the relative importance of human *v.* environmental factors is currently inadequate for Southern California's wildland fires (Pyne 2001; Haight *et al.* 2004). Studies have reached a variety of divergent conclusions (Heyerdahl *et al.* 2001; Keeley and Fotheringham 2001; Moritz 2003; Dickson *et al.* 2006), probably due in part to variation in fire characteristics at different temporal and spatial scales (Moritz *et al.* 2005; Parisien and Moritz 2009; Jin *et al.* 2014). Understanding the spatial distribution of ignition and the relative importance of human *v.* landscape controls, will become increasingly important as climate changes, the WUI expands and human influences on fire regimes increase (Sugihara *et al.* 2006; Westerling and Bryant 2008).

Recent county-level studies in California found that ignition frequency is significantly related to population density, with the highest number of fires observed at intermediate levels of population density and intermediate distances from the WUI (Syphard *et al.* 2007). Syphard *et al.* (2008) analysed the spatial patterns of fire ignitions in the Santa Monica Mountains and found that ignition occurrence is correlated with distance to human infrastructure (i.e. to roads, trails or housing development) and slope and vegetation type, whereas fire return interval is explained mainly by biophysical aspects related to climate and terrain (i.e. temperature, aspect, elevation and slope).

Further research is needed to quantify how human-related variables and biophysical drivers constrain wildland fire at fine spatial resolutions and over larger regional domains. We performed a spatial regression of a 30-year dataset of fire ignitions using human and biophysical explanatory variables. Our specific goals were to (i) assess the relative importance of possible drivers (e.g. distance to roads or population density) on ignition occurrence and frequency, (ii) develop statistical models of the spatial distribution of ignition occurrence and frequency, and (iii) quantify how much of the spatial pattern of fire return times can be explained by spatial variation in ignition. Strategic fuel management requires a better understanding of how landscape characteristics explain the likelihood of wildfires (Dellasala *et al.* 2004). Our results are intended to provide information that will help local fire management agencies identify and quantify ignition risk; this information may prove useful for optimising fire suppression resources or prevention planning. Our results also carry ecological implications for the management of natural resources and protection of wildland ecosystems (Haidinger and Keeley 1993).

Data and methods

Study area

Our main study domain covers 23 500 km^2 of wildland areas within USA National Forests in Santa Barbara, Ventura, Los Angeles, San Bernardino, Orange, Riverside and San Diego counties. The United States Forest Service (USFS) has used the administrative boundaries of National Forests as the spatial template for recording fire ignition locations. We therefore used this layout to develop our model, focusing on the Los Padres, Angeles, San Bernardino and Cleveland National Forests (Fig. 1). Southern California experiences a Mediterranean climate with a long, dry summer and a relatively short and mild rainy season (Bailey 1966). Contrasting patterns of temperature and rainfall lead to a diverse range of vegetation associations (Franklin 1998). Particularly widespread vegetation types include chaparral, open oak woodland, coastal sage scrub, valley grassland, oak woodland and coniferous forest (Di Castri *et al.* 1981; Arroyo *et al.* 1995; Davis and Richardson 1995). The region experiences intense human pressure: over 22 million people lived in Southern California in 2010 and an extensive road network connects numerous communities (source: US Census Bureau 2012). California has ~8500 miles (~13 700 km) of state and federal highways and the average road density within the National Forests in 2000 was ~1.3 km km^{-2} (US Census Bureau 2000).

Datasets: fire ignitions and fire perimeters

We extracted the ignition records for 1980–2009 from the USFS FIRESTAT database of individual fire incident reports (USDA Forest Service 2010). This period overlapped with the availability of information on human and biophysical factors and was chosen because earlier records were less reliable for ignition location and date. The location of fire origin was only specified to within ~0.8 km for some fires. This resulted in an artificial clustering of ignitions at ~1.6-km intervals in some areas. Given this uncertainty, we opted to carry out our analysis at a 3 × 3-km resolution. This resolution was chosen as a trade-off between higher resolution grids where ignition location error had a larger effect and lower resolution grids where the loss of spatial information reduced the usefulness of model predictions for management applications. Grids of varying sizes, from 1 to 5 km, were tested in preliminary sensitivity analyses: the 3-km resolution yielded a large continuous range of ignition frequencies, which aided model development. A Mantel test confirmed there was no evidence of spatial autocorrelation at 3-km resolution ($r = -0.065$, $P > 0.05$). We added a 5-km external buffer to account for ignitions near the National Forest boundaries. We further reduced the noise in the data by excluding ignitions that initiated fires less than 0.1 acres (~400 m^2). ArcGIS desktop computer software (ArcGIS Desktop, Environmental System Research Institute, Redlands, CA) was used for all digital map analyses.

In a final step we compared our ignition estimates with the observed spatial patterns of fire frequency to quantify the role of ignition frequency in explaining the local fire return interval. We computed a fire frequency map using historical perimeter data for fires larger than 100 acres (~0.4 km^2) compiled by the California Department of Forestry and Fire Protection's Fire

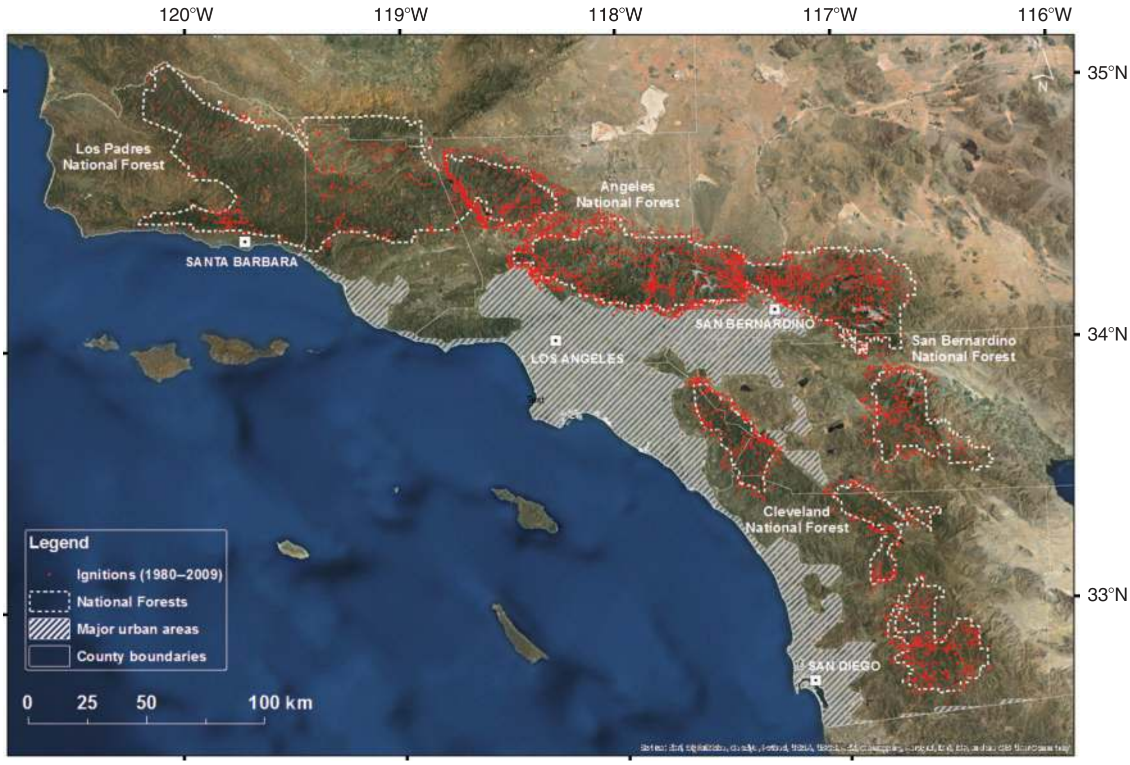

Fig. 1. Study area in Southern California, USA. The broader domain encompasses seven counties (white lines). All ignition incidents (red dots) recorded within National Forests boundaries (black dotted lines; major model domain) during 1980–2009 were overlaid on a 30-m-resolution Landsat satellite image mosaic (source: ESRI World Imagery database, November 2012).

Resource Assessment Program (FRAP 2010). We calculated fire frequency in each 3×3-km grid cell (i.e. number of fires during the 1980–2009 period per square kilometre), weighting each fire within a grid cell by its fractional burned area. We also classified these burns into Santa Ana and non-Santa Ana fires based on the fire start date reported in the FRAP (2010) database and the time series of Santa Ana days, following the approach described by Jin *et al.* (2014). Santa Ana days were identified when the north-easterly component of the daily mean wind speed was greater than $6\,\mathrm{m\,s^{-1}}$ at the exit of the largest gap across the Santa Monica Mountains (Hughes and Hall 2010).

Datasets: human factors

Recent studies identified numerous predictors of the occurrence of ignitions in densely populated areas (Syphard *et al.* 2008; Catry *et al.* 2009; Martínez *et al.* 2009). Fire ignitions recorded over the last three decades tend to be clustered around transportation networks and near urban areas. We therefore considered five variables to describe the human footprint: (1) distance to a major road, (2) distance to a minor road, (3) road density, (4) distance to low-density housing and (5) population density.

We used the US Census Bureau's TIGER road data (Topologically Integrated Geographic Encoding and Referencing; US Census Bureau 2000) to estimate the road density per grid cell and the distance from cell centroid to nearest road. Highways and state roads were classified as major roads (Fig. 2*a*), whereas streets and vehicle trails were classified as minor roads. This differentiation helped account for traffic volume as a controller of ignition.

We computed average housing density for 1980–2009 based on 1990 and 2000 US Census data (Hammer *et al.* 2004, 2007) (Fig. 2*b*). The WUI is often defined as areas with less than 50% vegetation and at least one house per 40 acres (6.2 houses $\mathrm{km^{-2}}$) that are located within 1.5 miles (~2.4 km) of an area over 500 ha that is more than 75% vegetated (Stewart *et al.* 2007). Assuming that ignitions most likely occur close to or within interface and intermix communities (Syphard *et al.* 2007, 2008), we used the distance to the nearest housing area with a density greater than 6.2 housing units $\mathrm{km^{-2}}$ as an indicator of the proximity to the WUI (Fig. 2*b*). In addition, we used the WUI vector maps created by the USFS (Radeloff *et al.* 2005) to calculate the number of ignitions within and outside the WUI areas.

Finally, we used the best available, fine-grained spatial database on population demographics from the US Census Bureau's block group data for 2000 (US Census Bureau 2001) to compute average population density per grid cell (Fig. 2*c*). This variable captured the direct influence of human presence within and outside the National Forest boundaries. All human variables were summarised at 3-km resolution.

Datasets: biophysical factors

We considered ten variables that were related to topography, land cover and climate: elevation, slope, south-westness, forest

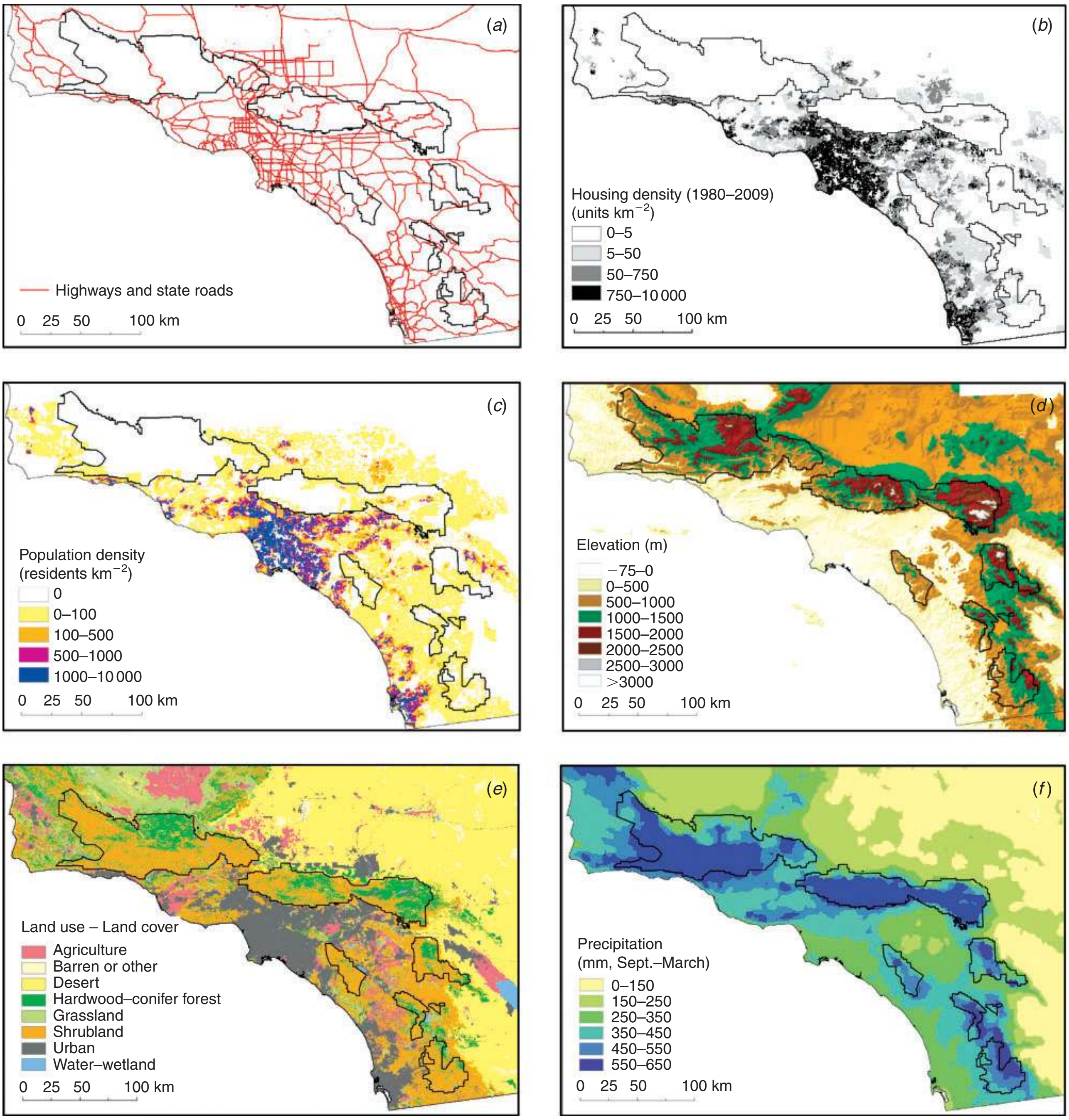

Fig. 2. Spatial maps of key human, topographic and biophysical drivers, including (*a*) major roads, (*b*) housing density, (*c*) population density, (*d*) elevation, (*e*) land cover and (*f*) mean annual winter (September–March) precipitation.

cover, shrubland cover, grassland cover, cover of other land cover types, annual average daily maximum temperature, annual average daily minimum temperature, and cumulative winter precipitation. We used the 3-arc-second digital elevation model from the US Geological Survey National Elevation Dataset (NED) to calculate the slope and aspect for each 3×3-km grid cell using ArcGIS software (Gesch *et al.* 2002) (Fig. 2*d*). Aspect was transformed trigonometrically to a south-facing index referred to as 'south-westness' (cos(aspect – 225°)) following Beers *et al.* (1966). This index provided a measure of sun exposure and dryness within each grid cell. Flat terrain with a slope of less than 5° was excluded from the aspect analyses. We assessed vegetation characteristics using the most recent and comprehensive land cover dataset at 100-m resolution from FRAP (2002). We classified the vegetation in the National Forests into three major types: 'shrubland' (60% of the area), 'forest/woodland' (22%) and 'grassland' (6%). The remaining non-vegetated land cover types were grouped as 'other', and

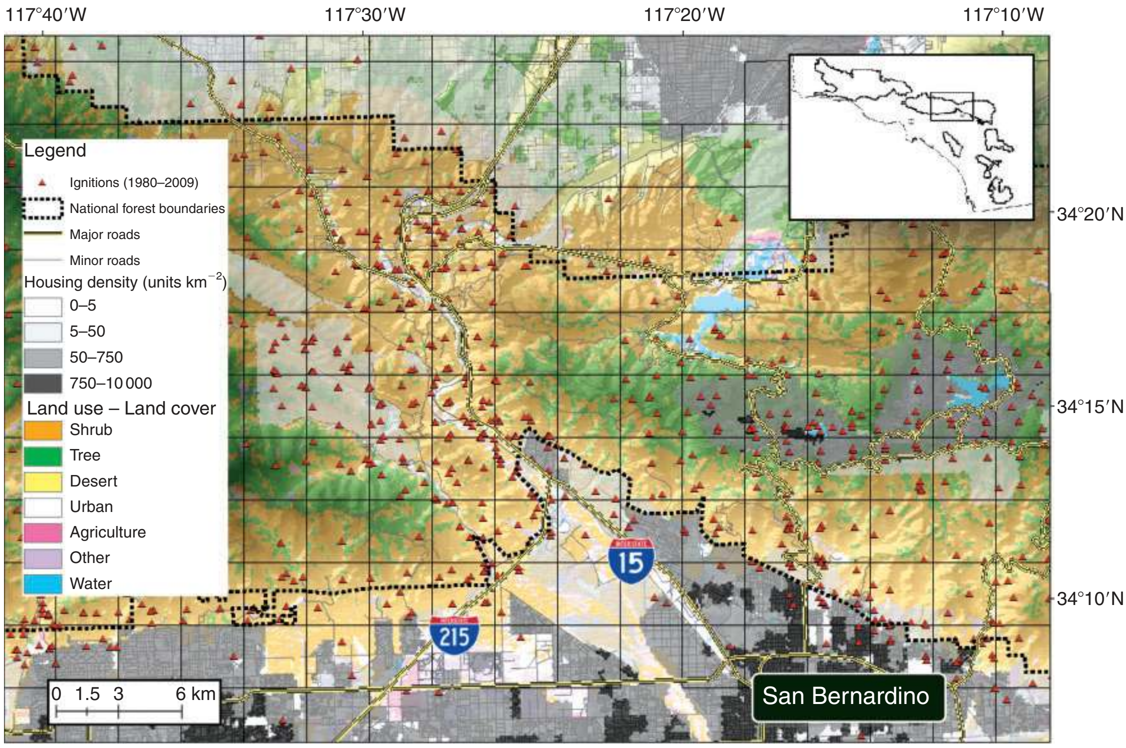

Fig. 3. The distribution of ignitions, roads and housing in the San Bernardino National Forest. This subset of the study area shows the 3×3-km grid framework used to summarise ignitions data and associated predictors. Observed ignitions (red triangles) between 1980 and 2009 were overlaid on GIS layers for roads, housing density and land cover types.

included agricultural land, urban, desert, wetland, water and barren (Fig. 2*e*). We calculated the fraction of each class within each 3×3-km grid cell. The monthly averages of precipitation and daily maximum and minimum temperature were taken from the gridded Parameter-Elevation Regressions on Independent Slopes Model (PRISM) dataset at 800-m resolution (Daly *et al.* 2002). In our analysis we summarised winter precipitation (September–March; Fig. 2*f*), and annual mean maximum and minimum temperatures for each 3×3-km cell over the 1981–2009 period.

Model building: logistic and Poisson regressions

Our input dataset consisted of the 15 explanatory variables described above, which were spatially averaged within each 3×3-km grid cell (Fig. 3). We sought to determine the influence of those predictors on two dependent variables: (i) the occurrence of ignition (presence or absence within a cell) and (ii) the frequency of ignition (number of ignitions within a cell).

We used a logistic regression approach to model the presence or absence of ignitions (Kleinbaum *et al.* 2002; Hosmer and Lemeshow 2005). Logistic regression has been used to successfully model the probability of fire occurrence at a range of geographic scales (Chou *et al.* 1993; Chuvieco *et al.* 1999; Vasconcelos *et al.* 2001). Logistic regression is expressed as:

$$\mathrm{logit}(p_i) = \ln\left(\frac{p_i}{1-p_i}\right) = a_0 + b_1 \cdot x_{1i} + \ldots + b_j \cdot x_{ji} \qquad (1)$$

where p_i is the probability of an ignition in the cell i and x_{ji} is the value of the jth predictor in the cell i. The underlying distribution is binomial and the logit function is defined as the natural logarithm (ln) of the probability of ignition occurrence.

We found exponential relationships between the explanatory variables and the number of ignitions (Fig. 4) and thus used a Poisson regression model (Agresti 2002) for ignition frequency expressed as:

$$\log(y) = a_0 + b_1 \cdot x_{1i} + \ldots + b_j \cdot x_{ji} \qquad (2)$$

Poisson models provide several advantages including the ability to represent a skewed distribution and the restriction of predicted values to non-negative numbers (Gardner *et al.* 1995). The Poisson probability distribution of observing any specific count y for an outcome Y and where λ describes the average rate of ignitions is given by:

$$\Pr(Y = y) = \frac{\lambda^y e^{-\lambda}}{y!} \qquad (3)$$

Model building: variable selection and model validation

We sought to identify which variables were most important for controlling the spatial distribution of ignitions, independent of their interaction with other explanatory variables. We first examined the correlation matrix among explanatory variables

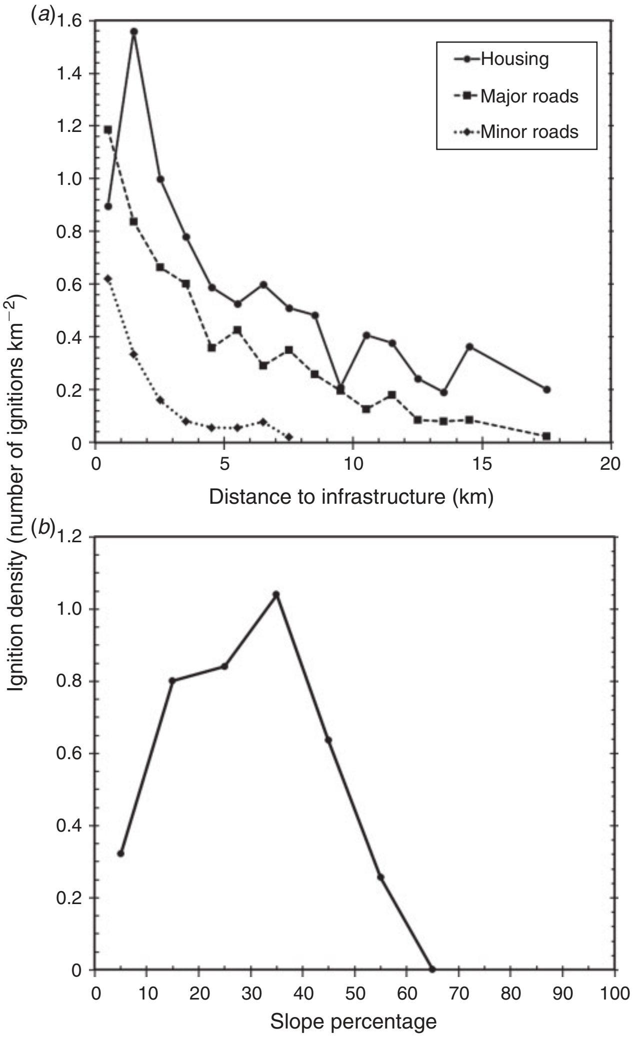

Fig. 4. Relationships between ignition density and (*a*) distance to human infrastructure (major roads, minor roads and low housing density), and (*b*) slope percentage over the major study domain (i.e. National Forests). Fig. 4*a* was computed by averaging the distance between each ignition location (FIRESTAT dataset) and the nearest human infrastructure within each 3 × 3-km grid.

for high pairwise correlations (Table 1) and then performed univariate logistic and Poisson regressions for all predictors using the R statistical package (R Development Core Team 2012; Table 2). We then developed logistic and Poisson multiple regression models that included all significant terms ($P < 0.001$) based on the univariate regression analyses. The original sample of 2625 grid cells (1483 with fire ignitions) was randomly split into fitting and validation subsamples (70 : 30 ratio). We sequentially selected the variables using a backward stepwise selection procedure (R package MASS; Crawley 2005) based on the Akaike Information Criterion (AIC; Akaike 1974). The backward selection process started with the full model and sequentially excluded explanatory variables based initially on the correlation with the response variable and other explanatory variables. We quantified the relative importance of retained variables by estimating their percentage contribution to the model goodness of fit (i.e. maximised log-likelihood).

Our goal was to develop simplified logistic and Poisson models, with a reduced set of explanatory variables as a compromise between model fit and model complexity. Logistic nested models were compared and examined using inferential and descriptive statistics. We used the likelihood ratio test and the Wald statistic to assess overall model fit and the respective contribution of individual predictors to fitted models. Receiver operating characteristic (ROC) analysis was performed to quantify the area under the curve (AUC) – a measure of the predictive capability of the logistic model to identify cells that had an ignition event. We also used the ROC curve to select the optimal threshold probability or cut-off value for the probability that an ignition would occur in a given cell. Using the validation samples we cross-validated the best multiple regression model and built contingency tables of observed and expected responses to evaluate model accuracy, precision, sensitivity and specificity (Hilbe 2009). As with the multiple logistic regression, we tested the significance of nested Poisson models and the significance of individual parameters using the likelihood ratio test and the Wald statistic. We also calculated Pearson's correlation coefficients for validation samples between observed and predicted values to assess model goodness of fit.

Model building: performance improvement

Poisson regression is a form of generalised linear models that assumes the conditional variance to equal the conditional mean. Therefore, a Poisson model is usually too restrictive when predicting count data, which manifests as data over-dispersion (i.e. the variance exceeds the mean) or as estimates of considerably fewer zero counts than are actually observed in the sample (Long 1997). As an alternative approach to Poisson regression, we tested a negative binomial (NB) regression (Agresti 2002; Gelman and Hill 2007; Hardin and Hilbe 2007), which uses a dispersion parameter φ to handle the variance independent of the mean.

Based on their respective contributions to model fit, we restricted the number of predictors to a subset of four explanatory variables while ensuring model performance was not significantly altered. We compared the performance of Poisson and NB models to ordinary linear regression, using Pearson's correlation coefficients, the root mean square error (RMSE), and the percentage of bias between simulated and observed data. We estimated how the four main predictors influenced model performance by comparing the nested models to the 4-parameters final model.

Predictive ignition frequency mapping

We applied the resulting NB model to estimate the ignition frequency for all 3 × 3-km grids within the National Forest boundaries, and analysed the spatial distribution of the residual ignition frequency. We used the Pearson correlation coefficient to estimate how much of the spatial variability in fire frequency derived from FRAP fire perimeter data was explained by the predicted variability of ignitions within National Forest boundaries. We also separately considered Santa Ana and non-Santa Ana fires following Jin *et al.* (2014) to compare the controls of ignition patterns on fire frequency for both fire types.

Table 1. Correlations among the 15 explanatory variables used in the analysis

The table indicates the Pearson's correlations between all independent variables. Please refer to the main text for more information on individual variables. Interactions with significant correlations ($P < 0.001$) superior to 0.5 (inferior to −0.5 for negative correlations) are indicated in bold

d.MajR														
0.54	**d.minR**													
0.38	0.36	**d.Housing**												
−0.42	−0.43	−0.36	**Droad**											
−0.22	−0.18	−0.22	**0.8**	**Dpopulation**										
0.18	0.26	0.14	−0.33	−0.28	**elevation**									
0.33	0.43	0.1	−0.49	−0.28	0.41	**slope**								
0.03	0.07	0.15	−0.09	−0.15	0.16	0.01	**south-westness**							
0.11	0.11	0.1	−0.19	−0.16	**0.66**	0.27	0.12	**tree**						
0.25	0.19	0.13	−0.42	−0.31	−0.16	0.31	−0.19	**−0.50**	**shrub**					
−0.18	−0.18	0.13	0.06	−0.03	−0.24	−0.36	−0.02	−0.12	−0.15	**grass**				
−0.32	−0.25	−0.31	**0.66**	**0.52**	−0.32	−0.46	0.13	−0.29	**−0.60**	−0.02	**other**			
−0.16	−0.16	−0.25	0.25	0.25	**−0.81**	−0.11	−0.16	**−0.63**	0.26	0.05	0.25	**Tmin**		
−0.18	−0.25	−0.14	0.31	0.21	**−0.85**	−0.44	−0.11	**−0.67**	0.16	0.17	0.38	**0.69**	**Tmax**	
0.26	0.34	0.1	−0.23	−0.09	0.34	**0.58**	−0.12	0.28	0.22	−0.18	−0.45	−0.16	**−0.51**	**Prec**

Table 2. Univariate logistic and Poisson regression results for all variables influencing the occurrence and frequency of fire ignitions in Southern California National Forests

Values and direction (i.e. positive or negative) of the coefficients indicate the influence of covariates (the driver variables) towards the response variables (ignition occurrence or frequency)

Explanatory variable	Fire occurrence (Logistic model) Binary response variable		Fire frequency (Poisson model) Continuous response variable	
	Coefficient	*P*-value	Coefficient	*P*-value
Human accessibility				
Distance to major roads (km)	−0.11 ± 0.010	<0.0001	−0.15 ± 0.005	<0.0001
Distance to minor roads (km)	−0.32 ± 0.040	<0.0001	−0.41 ± 0.020	<0.0001
Distance to housing (km)	−0.05 ± 0.005	<0.0001	−0.07 ± 0.003	<0.0001
Urban development				
Population density (1000 persons km^{-2})	−0.007 ± 0.004	0.09	0.004 ± 0.001	<0.01
Road density (km roads km^{-2})	0.006 ± 0.002	<0.05	0.017 ± 0.001	<0.0001
Topography				
Elevation (m)	0.44 ± 0.090	<0.0001	0.01 ± 0.003	<0.001
Slope (%)	0.007 ± 0.004	0.08	0.004 ± 0.001	<0.001
South-westness (0–1)	−0.14 ± 0.120	0.225	−0.07 ± 0.04	<0.05
Land cover types				
Tree (%)	0.67 ± 0.19	<0.0001	0.3 ± 0.06	0.629
Shrub (%)	0.11 ± 0.14	0.436	0.12 ± 0.04	<0.05
Grass (%)	−1.97 ± 0.42	<0.0001	−1.46 ± 0.18	<0.001
Others (%)	−0.42 ± 0.17	<0.05	0.19 ± 0.06	<0.001
Climate				
Temperature maximum (°C)	−0.03 ± 0.01	<0.05	0.02 ± 0.006	<0.001
Temperature minimum (°C)	−0.04 ± 0.02	<0.05	0.04 ± 0.006	<0.001
Winter precipitation (mm year^{-1})	0.57 ± 0.41	<0.01	0.39 ± 0.09	<0.001

Results

Spatial distribution of fire ignitions

Most of the wildland fire ignitions in the National Forests during 1980–2009 occurred near major roads and close to urban housing (Figs 1–3). Ignition points were clustered around populated areas, major infrastructure and highways, implying a strong influence by human factors (Fig. 3). The WUI, defined as areas where housing meets or intermingles with undeveloped wildlands (Stewart *et al.* 2007), had particularly high ignition densities (i.e. near the border of National Forests in Los Angeles, San Bernardino and Orange counties; Figs 1–3). Ignitions were much less frequent in sparsely populated areas such as Santa Barbara and Ventura counties (Fig. 1). The WUI covered only 5% of National Forest area but accounted for 40% of ignitions. Ignition density was considerably higher within the WUI (0.6 ignitions km^{-2}) than in more remote areas (0.03 ignitions km^{-2}).

A quantitative analysis of the relationship between ignition density and human variables confirmed that ignitions were most common near roadways and housing (Fig. 4). Approximately 60% of all ignitions occurred within 1 km of a major road, and ignition density declined with distance more rapidly from minor roads than from major roads (Fig. 4). Approximately 75% of ignitions occurred within 5 km of areas with a density of housing units greater than 6.2 km^{-2}. Ignition density peaked ~2 km and decreased exponentially in areas further away from housing (Fig. 4). Ignition density was highest in areas with intermediate levels of topographic complexity, with slopes between 20 and 40% (Fig. 4).

Influence of human and biophysical variables on ignition occurrence

Univariate logistic regressions showed that all human-related variables except population density were significant in explaining ignition occurrence ($P < 0.05$) (Table 2). Ignition occurrence was positively correlated with road density and negatively correlated with distance from major roads, minor roads and low-density housing. The presence or absence of ignitions also was related to vegetation type, with a significantly higher likelihood of ignition in forest and lower probability in grassland and non-vegetated areas (Table 2). We estimated that 53% of 1980–2009 ignitions occurred in shrublands and 22% in forests (Fig. 2*e*). Elevation was the only topographic variable significantly correlated with ignition occurrence: the likelihood of ignition increased with elevation. Precipitation was significantly correlated with ignition presence (Table 2).

The final model based on stepwise backward selection had 10 significant variables ($P < 0.0001$):

$$\text{logit}(p_i) = -4.65 - 0.12 \times d_{MajR} + 0.001 \times elev - 0.03 \times d_{Housing} + 2.04 \times shrub - 0.22 d_{minR} + 1.88 \times tree + 0.12 \times T_{max} + 0.02 \times slope + 0.02 \times D_{roads} - 0.002 \times D_{population} \quad (4)$$

where d_{MajR} is the distance to major roads (km), *elev* is elevation (m), $d_{Housing}$ is distance to nearest housing area with density greater than 6.2 units km^{-2} (km), *shrub* is the percentage cover of shrubland, d_{minR} is the distance to minor roads (km), *tree* is the percentage cover of forest, T_{max} is the annual average daily maximum temperature (°C) from 1980 to 2009, *slope* is the percentage slope, D_{roads} is road density (kilometres of roads per square kilometre) and $D_{population}$ is population density (number of persons per square kilometre). The analysis of modelled variance indicated that not all variables contributed equally to the model fit: d_{MajR}, *elev*, d_{Hou}, *shrub* and d_{minR} together explained over 87% of the model variance (Fig. 5*a*). The variables *tree*, *slope*, T_{max} and T_{min} were highly correlated with *elev* (Table 1), and their contribution to the model may have been masked by the apparent strong influence of elevation (Fig. 5*a*). Likewise, the contributions of D_{roads} and $D_{population}$ may have been masked by multicollinearities with d_{MajR}, $d_{Housing}$, *shrub* and d_{minR} (Table 1; Fig. 5*a*).

The ROC analysis using only the five strongest parameters resulted in an AUC of 0.72, which indicated that the reduced model was reasonably able to distinguish where ignitions were most likely to occur. Our cross-validation demonstrated that the model correctly predicted 67% of the observed distribution of ignition occurrence.

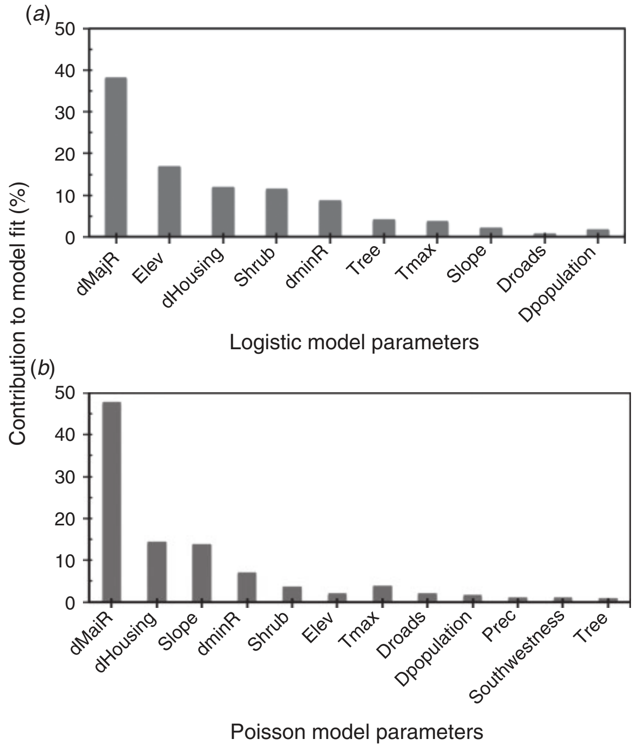

Fig. 5. Relative importance of explanatory variables as quantified by their relative contribution to model ignition occurrence (*a*) and ignition frequency (*b*). Variables are ranked by significance order estimated using automated stepwise selection. Significant variables included distance to major roads (*dMajR*), distance to housing (*dHousing*), distance to minor roads (*dminR*), population density (*Dpopulation*), road density (*Droads*), slope percentage, elevation, tree and shrub cover, annual average of daily maximum temperature (*Tmax*), cumulative winter precipitation (*prec*) and south-westness index.

Influence of human and biophysical variables on ignition frequency

We found that all variables related to human presence significantly explained variability of ignition frequency (Table 2). The variables d_{MajR}, $d_{Housing}$ and d_{minR} were the most influential human factors for ignition frequency (Figs 5*b*, 6). In contrast to the logistic regression results, all climate variables were significant and positively correlated with ignition frequency. Ignitions were more frequent in areas with warmer temperatures and higher precipitation (Table 2). The slope and shrub cover also had significant and positive influences on ignition frequency and contributed substantially to explaining the variability of ignition patterns (Table 2; Fig. 5*b*). The final Poisson regression following backward selection retained all explanatory variables ($P < 0.001$) except annual average daily minimum temperature:

$$\log(Ignitions_{FREQ}) = -4.38 - 0.14 \times d_{MajR} - 0.05 \times d_{Housing} + 0.02 \times slope - 0.23 \times d_{minR} + 1.13 \times shrub + 0.001 \times elev + 0.15 \times T_{max} + 0.02 \times D_{roads} - 0.001 \times D_{population} + 0.84 \times prec + 0.23 \times swindex + 0.67 \times tree \quad (5)$$

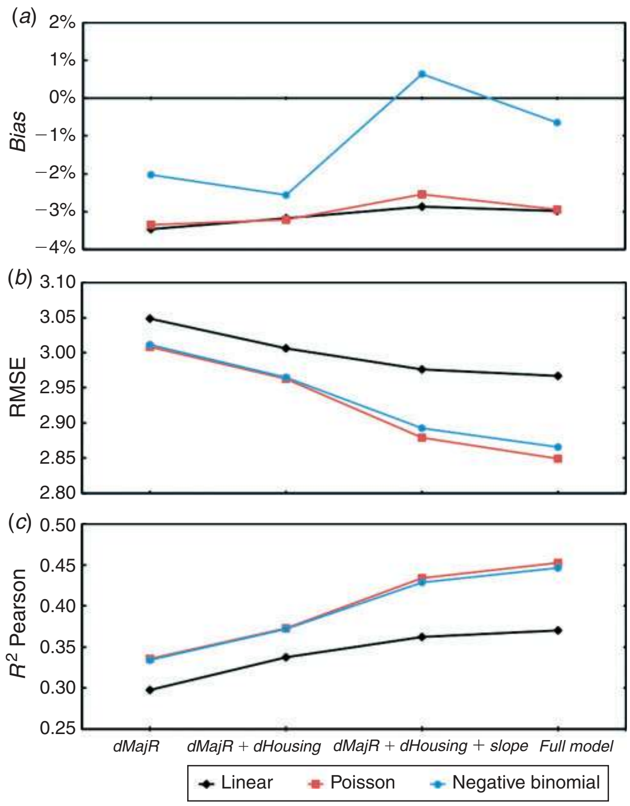

Fig. 6. Performance of reduced and full models with significant variables successively added into linear, Poisson and negative binomial regressions. Statistics shown were (*a*) percentage bias relative to the observation, (*b*) root mean square error and (*c*) Pearson R^2. Variables retained in the final models included distance to major roads (*dMajR*), distance to housing (*dHousing*), slope percentage and distance to minor roads (*dminR*).

where d_{MajR}, $d_{Housing}$, *slope*, d_{minR}, *shrub*, *elev*, T_{max}, D_{road}, $D_{population}$ and *tree* were defined as above, *prec* is the annual average cumulative winter precipitation (September–March) (mm year^{-1}) and *swindex* is the south-westness index. The most influential predictors of ignitions frequency were d_{MajR}, $d_{Housing}$, *slope* and d_{minR}: these variables combined explained ~85% of model variance (Fig. 5*b*). The comparison of Poisson and NB models to a linear model showed a clear improvement of model performance, with a reduced AIC value (Table 3). The linear model was inferior to the other models: the estimated coefficients of linear regression showed significantly higher standard errors (Table 3). Our results indicated that over-dispersion was better captured by the NB, as the dispersion estimate was closer to 1. The NB model improved the fit compared to the Poisson model with a significantly reduced bias while showing similar Pearson correlation and RMSE values ($R^2 = 0.45$; $RMSE = 2.79$) (Fig. 6). The form of the model equation for NB regression was the same as that for Poisson regression:

$$\log(Ignitions_{FREQ}) = 0.97 - 0.11 \times d_{MajR} - 0.04 \times d_{Housing} + 0.03 \times slope - 0.25 \times d_{minR} \quad (6)$$

Table 3. Summary of fitted regression models for ignition frequency data

The top part of the table gives coefficient estimates (with standard errors) for each explanatory variable. The second portion of the table compares model performance and reports the number of estimated parameters, maximised log-likelihood, AIC criterion and estimates of dispersion after model fitting

Model predictors	Linear model	Generalised linear models	
		Poisson	Negative binomial
(Intercept)	2.10 ± 0.06	1.00 ± 0.03	0.97 ± 0.07
distance to major roads	−0.83 ± 0.08	−0.14 ± 0.007	−0.11 ± 0.01
distance to housing	−0.49 ± 0.07	−0.04 ± 0.003	−0.04 ± 0.004
slope percentage	0.58 ± 0.07	0.03 ± 0.001	0.03 ± 0.003
distance to minor roads	−0.28 ± 0.09	−0.28 ± 0.03	−0.25 ± 0.04
Degrees of freedom	5	5	6
log-likelihood	−5322.3	−4120.7	−3223.4
AIC criterion	10 656.7	8251.4	6446.7
Dispersion φ	–	3.43	0.78

The distance to major road alone explained 33% of the observed spatial variance in ignition frequency; distance to housing and slope explained another 10%; and distance to minor roads explained the remainder. We caution that the primary influence of the variables retained in the final Poisson model may be confounded by multicollinearity (Table 1). For example, the importance of *slope* may have been overestimated due to implicit contributions from *elev*, *prec*, T_{max} or land cover.

Predictive mapping of ignition frequency

The spatial distribution of ignition frequency predicted using the NB model showed good agreement with the observed ignition pattern ($R^2 = 0.45$; Figs 6, 7*a*,*b*). The proximity to human infrastructure strongly determined ignition frequency (Fig. 7*b*; Table 3). The model accurately captured the relative lack of ignitions in remote, interior areas (e.g. Los Padres National Forest). Likewise, the model accurately predicted high ignition frequency in many areas near major roads and housing (e.g. Los Angeles County). The model underestimated ignition frequency along high-traffic transportation corridors (e.g. Interstate 5) and in close proximity to some populated urban areas (e.g. San Bernardino) (Fig. 7*c*). Thus, the variables used to predict ignition frequency (i.e. distance to major roads and distance to housing) were insufficient for discriminating areas where human pressure exceeded a certain threshold.

Relationship between ignition frequency and fire frequency

We found a significant positive relationship between ignition frequency and fire frequency. The gridded ignition frequency observations within National Forests explained 3.0% of the spatial variance of observed fire frequency ($P < 0.001$, $n = 2625$). For this same domain, the NB model explained 5.3% of the observed fire frequency variance ($P < 0.001$). We found that many areas with a high potential risk of ignition did not burn between 1980 and 2009 (Fig. 8), indicating that a substantial component of burned area variability was not explained by the drivers of ignition.

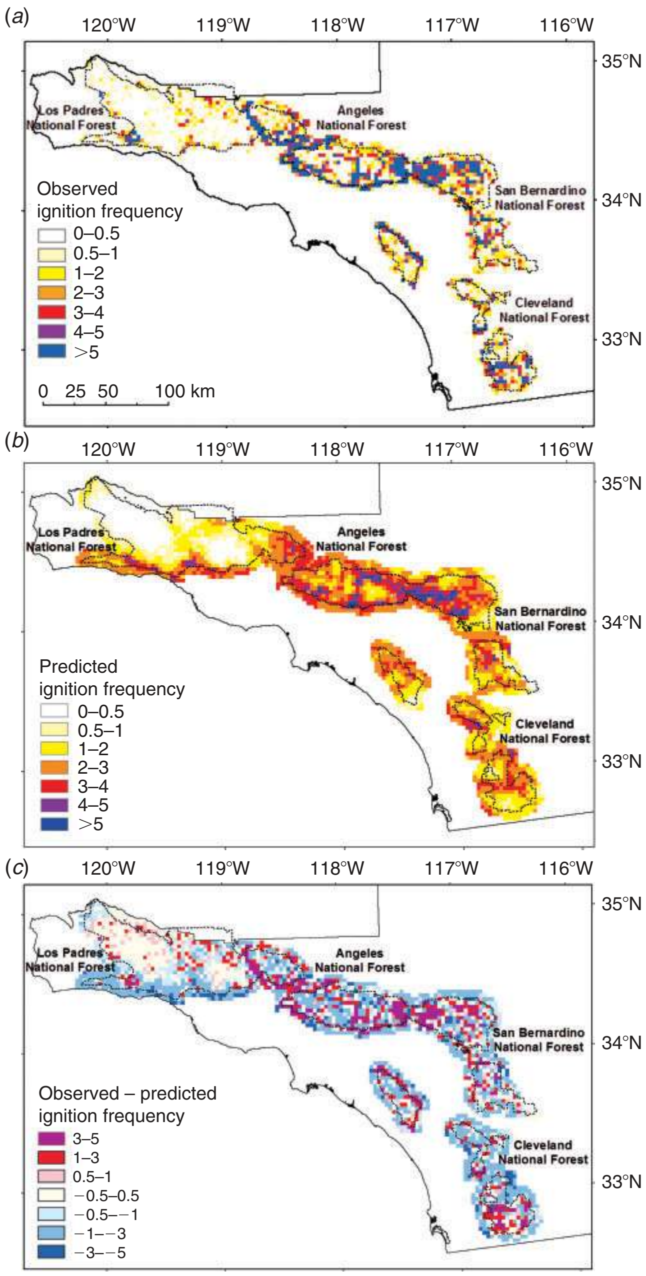

Fig. 7. Spatial maps of ignition frequency in unit of numbers of ignitions per grid cell across the National Forests at a 3-km resolution during 1980–2009 as (*a*) recorded in the FIRESTAT database and (*b*) predicted by the negative binomial model. The residual ignition frequency (observation – prediction) is shown in (*c*).

We repeated the previously described correlation analysis separately for Santa Ana *v.* non-Santa Ana fires (Jin *et al.* 2014). Santa Ana fires accounted for 45.0% of total burned area across Southern California and 18.0% of the number of fires over the 30-year study period. For non-Santa Ana fires, observed ignitions explained 6.5% ($P < 0.001$) of observed fire frequency and the estimated ignition patterns from the NB model explained 12.2% of fire frequency ($P < 0.001$) within the National Forests.

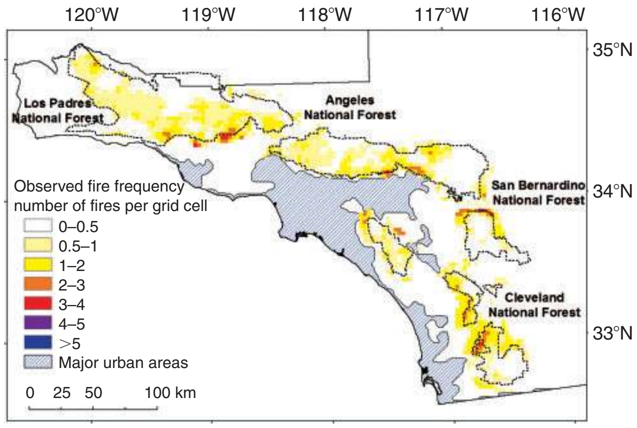

Fig. 8. Historical area-weighted fire frequency in USFS lands based on 1980–2010 fire perimeters data (FRAP 2010).

Ignitions were less important, although significant, in controlling the spatial pattern of Santa Ana fires, with observed and predicted ignitions explaining 3.1 and 4.9% ($P < 0.001$) of observed fire frequency.

Discussion

Our modelling approach allowed us to identify the combination of factors influencing the spatial distribution of ignitions. We found that proximity to roads and housing were the dominant controls for ignition frequency. All variables describing human accessibility and urban development were significantly correlated with ignition frequency. These results provided evidence that human activities are the primary source of ignition in Southern California and are consistent with studies that also found increased ignition frequency near transportation corridors (Stephens 2005) and WUIs (Syphard *et al.* 2007). Environmental variables resulted in a higher density of ignitions for mid-level slopes and forest land cover types. Past studies that considered generalised linear models to predict the spatial distribution of ignition frequency arrived at similar conclusions (Yang *et al.* 2007; Syphard *et al.* 2008). Although topographic features usually influence fire intensity (i.e. spread rate) and the distribution of burns across the landscape (Beaty and Taylor 2001; Alexander *et al.* 2006), variations in elevation cause variations in fuel type, moisture and phenology, which in turn control the conditions for fire ignition (Swetnam *et al.* 2011). Our logistic approach confirmed that ignition occurrence is most strongly determined by distance to major roads and housing, elevation and the proportion of shrub cover. Elevation may be capturing the secondary influence of temperature, slope and tree cover because these variables were collinear. The logistic model captured ~70% of ignition likelihood at 3-km resolution, which we considered satisfactory considering the heterogeneity and the large area investigated. Ignition occurrence was strongly conditioned by fuel type, with 75% of ignitions occurring in forest and shrubland.

Our NB approach improved model performance over commonly used linear and Poisson models. The NB model better captured the clustering patterns of ignitions around urban development and transportation corridors with a reduced set of predictors. Nevertheless, locations with particularly high ignition frequency, such as areas adjacent to major highways and parts of the WUI, were underestimated. It may be possible to improve model accuracy in these areas by incorporating traffic volume data as a proxy of human activity (www.traffic-counts.dot.ca.gov, accessed 13 August 2013).

Besides the errors input to the underestimation of ignitions frequency, two contrasting patterns of ignition merit discussion: a peri-urban ignition pattern in counties with dense development and a wildland ignition pattern in counties with sparse housing. Contrasting human influences on ignition patterns depending on human settlement density were reported by Badia-Perpinyá and Pallares-Barbera (2006). This partly explains why the scattered patterns of ignition were more difficult to capture in more rural areas. Using a fragmentation index to describe the interspersion of human infrastructures within wildland areas may help to refine model predictions.

Predicted ignition frequency explained a small but significant amount (i.e. 12%) of the observed spatial patterns in non-Santa Ana fire frequency within the National Forests. Wildland areas that are likely to experience greater numbers of ignitions coincide with areas characterised by recurrent burning. Nevertheless, in more remote areas than the WUI, such as Los Padres National Forest, where fuel fragmentation is not a limiting factor, fires tend to spread away from ignition sources and burn more frequently (Syphard *et al.* 2008). For Santa Ana fires, the lower correlation of ignition frequency to fire frequency suggested that the ignition controls on burned area patterns were considerably weaker relative to other factors as compared to non-Santa Ana fires. As burned area *per se* is not a function of ignition probability only (Archibald *et al.* 2009), additional variables related to fuel moisture, fuel continuity, fuel load and wind speed (Moritz *et al.* 2010) need to be considered for modelling burned area. Similarly, interactions between biophysical factors such as wind speed and precipitation or vegetation type may need consideration given the environmental heterogeneity in the region. Here, we addressed single-term effects of explanatory variables as we sought to build a simplified model of ignition frequency patterns. Although we also investigated quadratic and interactive terms between biophysical variables and between human variables, the results were not conclusive. An important future step is to combine our estimates of ignition frequency with other data sources to model the spatial distribution of burned area using a similar framework.

The predictive maps of ignition frequency generated in this study are synthetic measures of the spatial influence of human and environmental drivers on the current landscape. An important related question is how ignition patterns will evolve during future decades. Increasing human influence through densification and expansion of the WUI is expected to directly affect the wildland ignition regime (Hammer *et al.* 2007, Radeloff *et al.* 2010). California's population is projected to increase to ~49 million in 2025, a 44% increase from 2000. Although most population growth will occur in urban centres, housing density within 10 km of wildlands is projected to increase by ~80% by 2030 in California (Miller *et al.* 2011). We found distance to nearest housing area with density greater than 6.2 units km^{-2} was the second most influential control of ignition risk. As a consequence of future housing growth at the periphery of rural and wilderness areas, it is likely that ignition risk will increase. Traffic trends are likely to follow the WUI expansion, which would imply higher traffic volumes along fast-growing corridors such as Ventura, Orange and San Diego Counties (Crane *et al.* 2002). Our analysis demonstrated that major roads carry higher ignition risk than secondary roads. As a result, rising traffic among highways, such as Interstate 5, which crosses the Los Padres National Forest, and Interstate 15, which contours the San Bernardino National Forest (Fig. 3), will likely increase the ignition rate in these areas. Our approach may prove useful for both fire mitigation and urban planning. For example, it may be used to project ignition risk based on projections of future climatic and human activity across Southern California.

Acknowledgements

This study was supported by NASA Interdisciplinary Science grant NNX10AL14G to the University of California, Irvine. We thank the USFS and the California Department of Forestry and Fire Protection for providing the compiled ignition and fire perimeter data. We thank researchers from UCI's Earth System Science department for their comments on the earlier version of the manuscript, and two anonymous reviewers for their valuable comments.

References

Agee JK (1993) 'Fire Ecology of Pacific Northwest Forests.' (Island Press: Washington, DC)

Agresti A (2002) 'Categorical Data Analysis.' (Wiley: Hoboken, NJ)

Akaike H (1974) A new look at the statistical model identification. *Transactions on Automatic Control* **19**, 716–723. doi:10.1109/TAC.1974.1100705

Alexander JD, Seavy NE, Ralph CJ, Hogoboom B (2006) Vegetation and topographical correlates of fire severity from two fires in the Klamath–Siskiyou region of Oregon and California. *International Journal of Wildland Fire* **15**, 237–245. doi:10.1071/WF05053

Archibald S, Roy DP, Van Wilgen BW, Scholes RJ (2009) What limits fire? An examination of drivers of burnt area in southern Africa. *Global Change Biology* **15**, 613–630. doi:10.1111/J.1365-2486.2008.01754.X

Arroyo MTK, Zedler PH, Fox MD (1995) 'Ecology and Biogeography of Mediterranean Ecosystems in Chile, California, and Australia.' (Springer: London)

Badia-Perpinyá A, Pallares-Barbera M (2006) Spatial distribution of ignitions in Mediterranean periurban and rural areas: the case of Catalonia. *International Journal of Wildland Fire* **15**, 187–196. doi:10.1071/WF04008

Bailey HP (1966) 'The Climate of Southern California.' (University of California Press: Berkeley, CA)

Bar Massada A, Radeloff VC, Stewart SI, Hawbaker TJ (2009) Wildfire risk in the wildland–urban interface: a simulation study in northwestern Wisconsin. *Forest Ecology and Management* **258**, 1990–1999. doi:10.1016/J.FORECO.2009.07.051

Beaty RM, Taylor AH (2001) Spatial and temporal variation of fire regimes in a mixed conifer forest landscape, Southern Cascades, California, USA. *Journal of Biogeography* **28**, 955–966. doi:10.1046/J.1365-2699.2001.00591.X

Beers TW, Dress PE, Wensel LC (1966) Aspect transformation in site productivity research. *Journal of Forestry* **64**, 691–692.

Bradstock RA, Gill AM, Kenny BJ, Scott J (1998) Bushfire risk at the urban interface estimated from historical weather records: consequences for the use of prescribed fire in the Sydney region of south-eastern Australia. *Journal of Environmental Management* **52**, 259–271. doi:10.1006/JEMA.1997.0177

Callaway R, Davis F (1993) Vegetation dynamics, fire, and the physical environment in coastal central California. *Ecology* **74**, 1567–1578. doi:10.2307/1940084

Catry FX, Rego FC, Bação FL, Moreira F (2009) Modeling and mapping wildfire ignition risk in Portugal. *International Journal of Wildland Fire* **18**, 921–931. doi:10.1071/WF07123

Cayan DR, Maurer EP, Dettinger MD, Tyree M, Hayhoe K (2008) Climate change scenarios for the California region. *Climatic Change* **87**, 21–42. doi:10.1007/S10584-007-9377-6

Cayan DR, Das T, Pierce DW, Barnett TP, Tyree M, Gershunov A (2010) Future dryness in the southwest US and the hydrology of the early 21st century drought. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 21 271–21 276. doi:10.1073/PNAS.0912391107

Chou Y, Minnich R, Chase R (1993) Mapping probability of fire occurrence in San Jacinto Mountains, California, USA. *Environmental Management* **17**, 129–140. doi:10.1007/BF02393801

Chuvieco E, Salas FJ, Carvacho L, Rodriguez-Silva F (1999) Integrated fire risk mapping. In 'Remote Sensing of Large Wildfires'. (Ed. E Chuvieco) pp. 61–100. (Springer: Berlin)

Countryman CM (1972) The fire environment concept. USDA Forest Service, Pacific Southwest Forest and Ranger Experiment Station, Technical Paper. (Berkeley, CA) Available at http://www.firemodels.org/downloads/behaveplus/publications/Countryman/Countryman_1972_TheFireEnvironmentConcept_ocr.pdf [Verified 13 August 2013].

Crane R, Valenzuela A Jr, Chatman D, Schweitzer L, Wong P (2002) California Travel Trends and Demographics Study. University of California, Los Angeles, Institute of Transportation Studies School of Public Policy and Social Research, Final Report 74A0034. (Los Angeles, CA)

Crawley MJ (2005) Multiple regression. In 'Statistics: An Introduction using R'. pp. 195–208. (Wiley: Hoboken, NJ)

Daly C, Gibson WP, Taylor GH, Johnson GL, Pasteris P (2002) A knowledge-based approach to the statistical mapping of climate. *Climate Research* **22**, 99–113. doi:10.3354/CR022099

Davis GW, Richardson DM (1995) 'Mediterranean-type Ecosystems: the Function of Biodiversity.' (Springer: London)

Dellasala DA, Williams JE, Williams CD, Franklin JF (2004) Beyond smoke and mirrors: a synthesis of fire policy and science. *Conservation Biology* **18**, 976–986. doi:10.1111/J.1523-1739.2004.00529.X

Di Castri F, Goodall DW, Specht RL (1981) 'Mediterranean-type Shrublands.' (Elsevier: Amsterdam)

Dickson BG, Prather JW, Xu Y, Hampton HM, Aumack EN, Sisk TD (2006) Mapping the probability of large fire occurrence in northern Arizona, USA. *Landscape Ecology* **21**, 747–761. doi:10.1007/S10980-005-5475-X

Franklin J (1998) Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *Journal of Vegetation Science* **9**, 733–748. doi:10.2307/3237291

FRAP (2002) California Department of Forestry – Fire and Resource Assessment Program Multi-source Land Cover data. Available at http://frap.cdf.ca.gov/data/frapgisdata/download.asp?rec=fveg02_2 [Verified 9 May 2013]

FRAP (2010) California Department of Forestry – Fire and Resource Assessment Program GIS database of fire perimeter polygons. Available at http://frap.cdf.ca.gov/data/frapgisdata/download.asp?rec=fire [Verified 9 May 2013]

Gardner W, Mulvey EP, Shaw EC (1995) Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological Bulletin* **118**, 392–404. doi:10.1037/0033-2909.118.3.392

Gelman A, Hill J (2007) 'Data Analysis using Regression and Multilevel/Hierarchical Models.' (Cambridge University Press: Cambridge, UK)

Gesch D, Oimoen M, Greenlee S, Nelson C, Steuck M, Tyler D (2002) The national elevation dataset. *Photogrammetric Engineering and Remote Sensing* **68**, 5–11.

Haidinger TL, Keeley JE (1993) Role of high fire frequency in destruction of mixed chaparral. *Madrono* **40**, 141–147.

Haight RG, Cleland DT, Hammer RB, Radeloff VC, Rupp TS (2004) Assessing fire risk in the wildland-urban interface. *Journal of Forestry* **102**, 41–48.

Hammer RB, Stewart SI, Winkler RL, Radeloff VC, Voss PR (2004) Characterizing dynamic spatial and temporal residential density patterns from 1940–1990 across the North Central United States. *Landscape and Urban Planning* **69**, 183–199. doi:10.1016/J.LANDURBPLAN.2003.08.011

Hammer RB, Radeloff VC, Fried JS, Stewart SI (2007) Wildland–urban interface housing growth during the 1990s in California, Oregon, and Washington. *International Journal of Wildland Fire* **16**, 255–265. doi:10.1071/WF05077

Hardin JW, Hilbe J (2007) 'Generalized Linear Models and Extensions.' (StataCorp Press: College Station, TX)

Heyerdahl EK, Brubaker LB, Agee JK (2001) Spatial controls of historical fire regimes: a multiscale example from the interior west, USA. *Ecology* **82**, 660–678. doi:10.1890/0012-9658(2001)082[0660:SCOHFR]2.0.CO;2

Hilbe JM (2009) 'Logistic Regression Models.' (Chapman & Hall/CRC: Boca Raton, FL)

Hosmer DW, Lemeshow S (2005) Introduction to the Logistic Regression Model. In 'Applied Logistic Regression', 2nd edn. pp. 1–30. (Wiley: Hoboken, NJ)

Hughes M, Hall A (2010) Local and synoptic mechanisms causing Southern California's Santa Ana winds. *Climate Dynamics* **34**, 847–857. doi:10.1007/S00382-009-0650-4

Jin Y, Randerson JT, Faivre N, Capps S, Hall A, Goulden ML (2014) Contrasting controls on wildland fires in Southern California during periods with and without Santa Ana winds. *Journal of Geophysical Research – Biogeosciences* **119**, 432–450. doi:10.1002/2013JG002541

Keeley JE (1982) Distribution of lightning- and man-caused wildfires in California. In 'Proceedings of the International Symposium on the Dynamics and Management of Mediterranean Type Ecosystems', 22–26 June 1981, San Diego, CA. (Eds CE Conrad, WC Oechel) USDA Forest Service, Pacific Southwest Forest and Range Experiment Station, General Technical Report PSW-GTR-058, pp. 431–437. (Berkeley, CA)

Keeley J, Fotheringham C (2001) Historic fire regime in southern California shrublands. *Conservation Biology* **15**, 1536–1548. doi:10.1046/J.1523-1739.2001.00097.X

Keeley JE, Zedler PH (2009) Large, high intensity fire events in southern California shrublands: debunking the fine-grained age-patch model. *Ecological Applications* **19**, 69–94. doi:10.1890/08-0281.1

Keeley J, Fotheringham C, Morais M (1999) Reexamining fire suppression impacts on brushland fire regimes. *Science* **284**, 1829–1832. doi:10.1126/SCIENCE.284.5421.1829

Keeley JE, Safford H, Fotheringham CJ, Franklin J, Moritz M (2009) The 2007 Southern California wildfires: lessons in complexity. *Journal of Forestry* **107**, 287–296.

Kleinbaum DG, Klein M, Pryor ER (2002) 'Logistic Regression: a Self-Learning Text.' (Springer: New York)

Long JS (1997) 'Regression models for categorical and limited dependent variables.' (SAGE Publications Inc.: Thousand Oaks, CA)

Martínez J, Vega-Garcia C, Chuvieco E (2009) Human-caused wildfire risk rating for prevention planning in Spain. *Journal of Environmental Management* **90**, 1241–1252. doi:10.1016/J.JENVMAN.2008.07.005

Miller C, Abatzoglou J, Brown T, Syphard A (2011) Wilderness Fire Management in a Changing Environment. In 'The Landscape Ecology of Fire'. (Eds D McKenzie, C Miller, DA Falk) pp. 269–294. (Springer: London)

Moritz MA (2003) Spatiotemporal analysis of controls on shrubland fire regimes: age dependency and fire hazard. *Ecology* **84**, 351–361. doi:10.1890/0012-9658(2003)084[0351:SAOCOS]2.0.CO;2

Moritz MA, Morais ME, Summerell LA, Carlson J, Doyle J (2005) Wildfires, complexity, and highly optimized tolerance. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 17 912–17 917. doi:10.1073/PNAS.0508985102

Moritz MA, Moody TJ, Krawchuk MA, Hughes M, Hall A (2010) Spatial variation in extreme winds predicts large wildfire locations in chaparral ecosystems. *Geophysical Research Letters* **37**, L04801. doi:10.1029/2009GL041735

Narayanaraj G, Wimberly MC (2012) Influences of forest roads on the spatial pattern of human- and lightning-caused wildfire ignitions. *Applied Geography* **32**, 878–888. doi:10.1016/J.APGEOG.2011.09.004

Pan LL, Chen SH, Cayan D, Lin MY, Hart Q, Zhang MH, Liu YB, Wang JZ (2011) Influences of climate change on California and Nevada regions revealed by a high-resolution dynamical downscaling study. *Climate Dynamics* **37**, 2005–2020. doi:10.1007/S00382-010-0961-5

Parisien MA, Moritz MA (2009) Environmental controls on the distribution of wildfire at multiple spatial scales. *Ecological Monographs* **79**, 127–154. doi:10.1890/07-1289.1

Perry GLW (1998) Current approaches to modelling the spread of wildland fire: a review. *Progress in Physical Geography* **22**, 222–245. doi:10.1177/030913339802200204

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

Pyne SJ (2001) 'Fire: a Brief History.' (University of Washington Press: Washington, DC)

Pyne SJ, Andrews PJ, Laven RD (1996) 'Introduction to Wildland Fire', 2nd edn. (Wiley: New York)

R Development Core Team (2012) 'R: a Language and Environment for Statistical Computing' (R Foundation for Statistical Computing: Vienna, Austria)

Radeloff VC, Hammer RB, Stewart SI, Fried JS, Holcomb SS, McKeefry JF (2005) The wildland-urban interface in the United States. *Ecological Applications* **15**, 799–805. doi:10.1890/04-1413

Radeloff VC, Stewart SI, Hawbaker TJ, Gimmi U, Pidgeon AM, Flather CH, Hammer RB, Helmers DP (2010) Housing growth in and near United States protected areas limits their conservation value. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 940–945. doi:10.1073/PNAS.0911131107

Spracklen DV, Mickley LJ, Logan JA, Hudman RC, Yevich R, Flannigan MD, Westerling AL (2009) Impacts of climate change from 2000 to 2050 on wildfire activity and carbonaceous aerosol concentrations in the western United States. *Journal of Geophysical Research* **114**, D20301. doi:10.1029/2008JD010966

Stephens SL (2005) Forest fire causes and extent on United States Forest Service lands. *International Journal of Wildland Fire* **14**, 213–222. doi:10.1071/WF04006

Stewart SI, Radeloff VC, Hammer RB, Hawbaker TJ (2007) Defining the wildland–urban interface. *Journal of Forestry* **105**, 201–207.

Strauss D, Bednar L, Mees R (1989) Do one percent of forest fires cause ninety-nine percent of the damage? *Forest Science* **35**, 319–328.

Sugihara NG, VanWagtendonk JW, Fites-Kaufman J, Shaffer KE, Thode AE (2006) The Future of Fire in California's Ecosystems. In 'Fire in California's Ecosystems'. (Eds N Sugihara, J Van Wagtendonk, KE Shaffer, J Fites-Kaufman, AE Thode) pp. 538–543. (The University of California Press: Berkeley, CA)

Swetnam T, Falk D, Hessl A, Farris C (2011) Reconstructing Landscape Pattern of Historical Fires and Fire Regimes. In 'The Landscape Ecology of Fire'. (Eds D McKenzie, C Miller, DA Falk) pp. 165–192. (Springer: London)

Syphard A, Radeloff V, Keeley J, Hawbaker T, Clayton M, Stewart S, Hammer R (2007) Human influence on California fire regimes. *Ecological Applications* **17**, 1388–1402. doi:10.1890/06-1128.1

Syphard A, Radeloff V, Keuler N, Taylor R, Hawbaker T, Stewart S, Clayton M (2008) Predicting spatial patterns of fire on a southern California landscape. *International Journal of Wildland Fire* **17**, 602–613. doi:10.1071/WF07087

Syphard AD, Keeley JE, Massada AB, Brennan TJ, Radeloff VC (2012) Housing arrangement and location determine the likelihood of housing loss due to wildfire. *PLoS ONE* **7**, e33954. doi:10.1371/JOURNAL.PONE.0033954

Syphard AD, Bar Massada A, Butsic V, Keeley JE (2013) Land use planning and wildfire: development policies influence future probability of housing loss. *PLoS ONE* **8**(8), e71708. doi:10.1371/JOURNAL.PONE.0071708

US Census Bureau (2000) 'Census 2000 TIGER/Line Files' (US Census Bureau: Washington, DC)

US Census Bureau (2001) 'Census 2001 Population and Housing Block Groups/Shape Files' (US Census Bureau: Washington, DC)

US Census Bureau (2012) 2010 Census of Population and Housing, Population and Housing Unit Counts, CPH-2-6, California (US Government Printing Office: Washington, DC) Available at http://www.census.gov/prod/cen2010/cph-2-6.pdf [Verified 3 June 2014]

USDA Forest Service (2010) FIRESTAT GIS database of fire origins from the US Department of Agriculture and Forest Service. Available at http://www.fs.fed.us/r5/rsl/projects/frdb/layers/fire.html [Verified 15 April 2013]

Vasconcelos MJP, Silva S, Tome M, Alvim M, Pereira JC (2001) Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks. *Photogrammetric Engineering and Remote Sensing* **67**, 73–81.

Verdú F, Salas J, Vega-García C (2012) A multivariate analysis of biophysical factors and forest fires in Spain, 1991–2005. *International Journal of Wildland Fire* **21**, 498–509. doi:10.1071/WF11100

Westerling AL (2006) Warming and earlier spring increase western U.S. forest wildfire activity. *Science* **313**, 940–943. doi:10.1126/SCIENCE.1128834

Westerling A, Bryant B (2008) Climate change and wildfire in California. *Climatic Change* **87**, 231–249. doi:10.1007/S10584-007-9363-Z

Westerling A, Bryant B, Preisler H, Hidalgo H, Das T, Shrestha S (2009) Climate change, growth, and California wildfire. California Energy Commission, Draft Paper. (Sacramento, CA)

Whelan RJ (1995) 'The Ecology of Fire.' (Cambridge University Press: Cambridge, UK)

Yang J, He HS, Shifley SR, Gustafson EJ (2007) Spatial patterns of modern period human-caused fire occurrence in the Missouri Ozark Highlands. *Forest Science* **53**, 1–15.