CSIRO PUBLISHING

International Journal of Wildland Fire
http://dx.doi.org/10.1071/WF11044

Spatial variability in wildfire probability across the western United States

Marc-André Parisien[A,B,G], *Susan Snetsinger*[C], *Jonathan A. Greenberg*[D], *Cara R. Nelson*[C], *Tania Schoennagel*[E], *Solomon Z. Dobrowski*[F] *and Max A. Moritz*[B]

[A]Northern Forestry Centre, Canadian Forest Service, Natural Resources Canada, 5320 122nd Street, Edmonton, AB, T5H 3S5, Canada.
[B]Department of Environmental Science, Policy and Management, University of California – Berkeley, 137 Mulford Hall 3114, Berkeley, CA 94720, USA.
[C]Department of Ecosystem and Conservation Sciences, College of Forestry and Conservation, University of Montana, 32 Campus Drive, Missoula, MT 59812, USA.
[D]Department of Geography, University of Illinois at Urbana-Champaign, 607 South Mathews Avenue, MC 150, Urbana, IL 61801, USA.
[E]Department of Geography, University of Colorado, Boulder, Boulder, CO 80309, USA.
[F]Department of Forest Management, College of Forestry and Conservation, University of Montana, Missoula, MT 59812,USA.
[G]Corresponding author. Email: marc-andre.parisien@nrcan-rncan.gc.ca

Abstract. Despite growing knowledge of fire–environment linkages in the western USA, obtaining reliable estimates of relative wildfire likelihood remains a work in progress. The purpose of this study is to use updated fire observations during a 25-year period and a wide array of environmental variables in a statistical framework to produce high-resolution estimates of wildfire probability. Using the MaxEnt modelling technique, point-source fire observations that were sampled from area burned during the 1984–2008 time period were related to explanatory variables representing ignitions, flammable vegetation (i.e. fuels), climate and topography. Model results were used to produce spatially explicit predictions of wildfire probability. To assess the effect of humans on the spatial patterns of wildfire likelihood, we built an alternative model that excluded all variables having a strong anthropogenic imprint. Results showed that wildfire probability in the western USA is far from uniform, with different areas responding to different environmental drivers. The effect of anthropogenic factors on wildfire probability varied by region but, on the whole, humans appear to inhibit fire activity in the western USA. Our results not only provide what appear to be robust predictions of wildfire likelihood, but also enhance understanding of long-term controls on wildfire activity. In addition, our wildfire probability maps provide better information for strategic planning of land-management activities, especially where fire regime knowledge is sparse.

Additional keywords: climate, fuels, ignitions, MaxEnt algorithm, spatial modelling, topography.

Received 23 March 2011, accepted 18 August 2011, published online 8 February 2012

Introduction

Recent analyses of high-quality fire and environmental data suggest that variability in the relative likelihood of wildfire occurrence across the western USA is more extreme than previously thought (Littell *et al.* 2009; Parisien and Moritz 2009; Finney *et al.* 2011). Not only does average fire frequency vary among biomes, but it can also fluctuate enormously within biomes, especially where rugged topography or concentrated anthropogenic influence dominate the landscape (Schoennagel *et al.* 2004). For example, in a span of a few kilometres, fire frequency on one side of a mountain range can be orders of magnitude greater than on the other side (Heyerdahl *et al.* 2001; Mermoz *et al.* 2005). Although disturbance dynamics of specific sites or landscapes are well documented, wildfire dynamics and probabilities are unknown for many areas (Collins *et al.* 2010). For this reason, a regional to continent-wide continuous depiction of wildfire likelihood is difficult to obtain and, as a consequence, its spatial variability is often unknown or assumed to be low. An unfortunate outcome of this knowledge gap is that national land-management policies may be well adapted to certain areas but not others (Schoennagel and Nelson 2011).

Over the last decade, several investigators have begun addressing the issue of spatial variability in fire activity across

North America. For instance, fire frequency metrics – computed for large ecological or administrative areas from spatial fire atlases in the United States (Stephens 2005; Bartlein *et al.* 2008) and Canada (Stocks *et al.* 2002; Parisien *et al.* 2006) – have provided a coarse estimate of subcontinental fire variability. There has also been an effort to classify fire regimes into a few synthetic condition classes in order to evaluate the departure from historical fire regimes (Hardy *et al.* 2001). Some investigators have developed wildfire–climate relationships to map monthly or annual predicted ignition probability (Balshi *et al.* 2009; Preisler *et al.* 2009), whereas others have linked long-term climatology to decadal patterns in area burned (Skinner *et al.* 1999; Gedalof *et al.* 2005). By developing fire–climate models for ecological zones in the western USA, Littell *et al.* (2009) have implicitly accounted for the effect of dominant vegetation type on annual area burned. Parisien and Moritz (2009) explicitly incorporated vegetation classes, in combination with mapped climate normals, to map relative wildfire likelihood across the USA. In contrast to these studies, which are statistically based, Finney *et al.* (2011) used a fire simulation model to produce wildfire probability estimates for 134 landscapes that were composited to cover the entire conterminous USA (Finney *et al.* 2011).

Despite recent progress in understanding continent-wide North American wildfire likelihood and wildfire–climate relationships, studies to date often ignore ignition sources and the presence of sufficient biomass for combustion ('fuel'). Although three environmental factors must coincide for fire to occur – available fuels, periods when weather conditions support combustion and ignitions – they rarely act independently of one another (Moritz *et al.* 2005). For instance, whereas climate has a direct effect on fuel moisture and combustion, these same conditions also affect fire indirectly by controlling patterns in vegetation (Krawchuk *et al.* 2009; Bradstock 2010). This dual influence of climate is best illustrated using the Sahara desert: even though this area experiences some of the most extreme fire weather in the world, its climates are completely unsuitable for the presence of flammable vegetation.

Realistic wildfire likelihood estimates require that models incorporate anthropogenic activities that shape our landscapes and disturbance regimes, but these anthropogenic drivers are seldom accounted for (but see Cardille *et al.* 2001; Syphard *et al.* 2008). Isolating the influence of people on fire regimes is not a trivial task, as human activities often strongly covary with changes in vegetation types or climatic factors (Girardin *et al.* 2009; Meyn *et al.* 2010). Nevertheless, every year humans ignite most wildfires in North America (Stocks *et al.* 2002; Stephens 2005). Conversely, people may reduce wildfire activity through fire suppression, but suppression effectiveness appears to be highly variable among areas, ranging from a drastic reduction in area burned in some areas to virtually no detectable long-term effects in others (Stephens and Ruth 2005; Finney *et al.* 2009).

Perhaps more significantly, humans may have modified fire regimes indirectly through land-use change. In fact, Marlon *et al.* (2008) have reported a global decrease in fire activity during the last century, attributing this trend to a general reduction in fuel continuity in fire-prone areas as a result of land-use change. Although much of North America (particularly in the West) is viewed as largely undisturbed wildlands, this is far from true, as all but the most remote areas have some trace of human influence (Cardille and Lambois 2010). The study of 'natural' fire regimes in areas – or time periods – of low human influence is certainly pivotal to our understanding of fire–climate–vegetation interactions, but accurate estimates of current wildfire probability in North America require anthropogenic activities to be taken into account. For example, even though much of the Great Plains was once fire-dominated (Brown *et al.* 2005), widespread agriculture has rendered this area ill-suited for fire ignition and spread.

The purpose of this study is to build on previous investigations of subcontinental fire activity to spatially evaluate the likelihood of wildfire in the 11 westernmost states of the conterminous USA. This was achieved using a statistical framework that linked high-resolution patterns of area burned from 1984 to 2008 to a set of variables chosen to represent the ignition patterns and vegetation, as well as climate normals, extremes and seasonality, that characterise this 25-year period. In addition to producing high-resolution estimates of wildfire probability, we examined the influence of environmental factors on its spatial variability. In order to assess the anthropogenic imprint on wildfire likelihood, we created a second model for which we omitted all variables affected by human activities. Finally, we examined whether the exclusion of small fires, which burn a minute fraction of the total area, is a reasonable simplification in this type of modelling.

Study area

The study area comprises the 11 westernmost states of the conterminous USA ($\sim 3 \times 10^6$ km^2) (Fig. 1). It covers a broad environmental spectrum that encompasses extreme variation in geology, landform, climate, vegetation and land use (Barbour and Billings 2000). The climate of the area is controlled by two broad-scale gradients: a west–east 'continentality' gradient of decreasing moisture and increased temperature seasonality and a north–south gradient of decreasing moisture and increasing temperature. Overall, most of the precipitation falls during the winter months, except for some dry areas of the south-west. The north-west, which experiences warm summers and mild winters, receives the most precipitation, whereas the deserts of the south-west are the most arid. The terrain also plays a large role on climate, as temperature decreases and precipitation increases with elevation. There is a strong rain shadow effect on the lee side of large mountain ranges, notably the Rocky Mountains and the Sierra Nevada, to the east of which lie the Great Basin and the Great Plains respectively, which are areas typified by dry continental climates.

The climatic and geological gradients of the study area are reflected in the vegetation cover, as well as in the historical fire regimes (Hardy *et al.* 2001). The north-west boasts lush temperate rainforests that have infrequent wildfires. Southward, the area ranging from the coast to the Sierra Nevada Mountains has a mix of forest, grassland and shrubland vegetation that varies considerably with respect to fire return interval and fire severity. The chaparral areas of southern California experience some of the highest fire frequencies of North America. To the east, the Rocky Mountains are dominated by coniferous vegetation that support a fire regime of frequent and fairly low-intensity

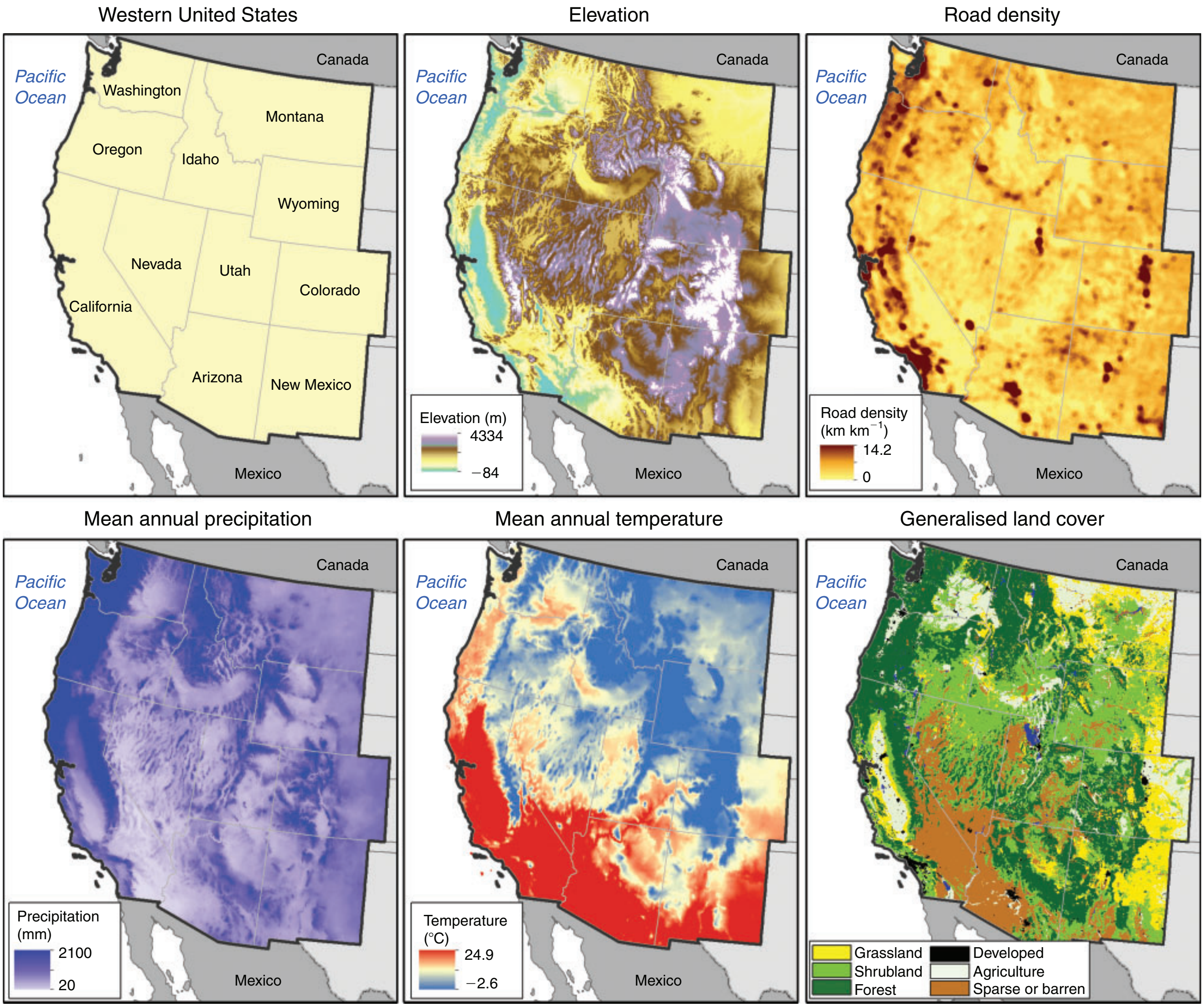

Fig. 1. The study area showing the 11 western USA states, elevation, road density (computed using a 1000-ha circular window), mean annual precipitation, mean annual temperature and land cover that was generalised from the National Gap Analysis.

wildfires at low elevations and periodic stand-renewing wildfires at high elevations. Between the more coastal Sierra Nevada and Cascades Mountains and the more continental Rockies, the vegetation is dominated by drought-adapted vegetation that seldom burns; however, fires do occur within pockets with fairly continuous biomass, including areas dominated by invasive grasses (mainly *Bromus tectorum*). East of the Rocky Mountains, the Great Plains, which were historically fire-prone grasslands, represent a fairly dry area that has undergone massive conversion to agriculture.

Methods

Wildfire probability models were built by relating data points sampled in burned areas from 1984 to 2008 (dependent variable) to a set of explanatory variables that characterised ignition sources, flammable vegetation (i.e. fuels), climate and topography (Table 1). These models were built using the MaxEnt software (Phillips *et al.* 2006), previously shown to be effective in spatial modelling of environmental constraints on fire activity (Parisien and Moritz 2009). The climate variables were averaged for the study's time period and captured spatial patterns in both climatic normals and extremes. The ignitions, vegetation and topography variables were also averaged among years when appropriate (e.g. lightning) but most often used a single representative year if they did not vary significantly from year to year (e.g. road density, topographic roughness). Because it has been shown that the neighbourhood information of the ignitions, vegetation and topography variables may be as important to area burned as the observation at a given location (Parisien *et al.* 2011), four spatial scales of observation (1, 100, 1000 and 10 000 ha) were computed for each variable using a 'moving-window' approach. We deemed it unnecessary to compute moving-window variables for the climate variables, because using the neighbourhood of climate variables does not substantially improve fire predictions (Parisien *et al.* 2011). All data

Table 1. Variables selected for analysis and their description

All climate variables were calculated on a monthly basis and annual averages were based on mean values of every month. Unless otherwise specified, all values were computed for the 1984–2008 time period. All citations for the sources are provided in the text

Category	Input name	Source	Description
Ignitions	Pop_Dens1[A]	Gridded population of the world, version 3	Population density at the 1-ha scale (people km^{-2})
	RdlsVol1[A]	ESRI StreetMap	Roadless volume metric (remoteness) at the 1-ha scale (km^3 $person^{-1}$)
	RdlsVol10000[A]	ESRI StreetMap	Roadless volume metric (remoteness) at the 10 000-ha scale (km^3 $person^{-1}$)
	Lgt_Dens10000	NASA Global Hydrology and Climate Centre	Annual density of lightning strikes (1995–2005) at the 10 000-ha scale (strikes km^{-2} $year^{-1}$)
Climate	MeanTempWettest	PRISM	Temperature of the wettest month (°C)
	MeanTempDriest	PRISM	Temperature of the driest month (°C)
	DiurTempRange	PRISM	Diurnal range in temperature (°C)
	Isotherm	PRISM	Diurnal temperature range ÷ temperature range (×100)
	PcpColdest	PRISM	Precipitation of the coldest month (mm)
	PcpSeas	PRISM	Precipitation seasonality (coefficient of variation)
	WatDef_CV_Ann	PRISM	Coefficient of variation among annual water deficit values (%)
	MaxSPI	PRISM	Maximum monthly standardised precipitation index
	WindDriest99	NOAA	99th percentile wind speed of the driest month ($m\,s^{-1}$)
Topography	SurfArea_Ratio1	USGS/EROS	Ratio of surface to area at the 1-ha scale
Vegetation	GPP100	MODIS (MOD17A3)	Gross primary productivity at the 100-ha scale ($g\,C\,km^{-2}$ $year^{-1}$)
	Fuel_Pct100[A]	USA GAP Analysis Land cover	Percentage land cover of fuels at the 100-ha scale (%)
	Fuel_Pct10000[A]	USA GAP Analysis Land cover	Percentage land cover of fuels at the 10 000-ha scale (%)

[A]Variables that were excluded from the Non-anthropogenic model.

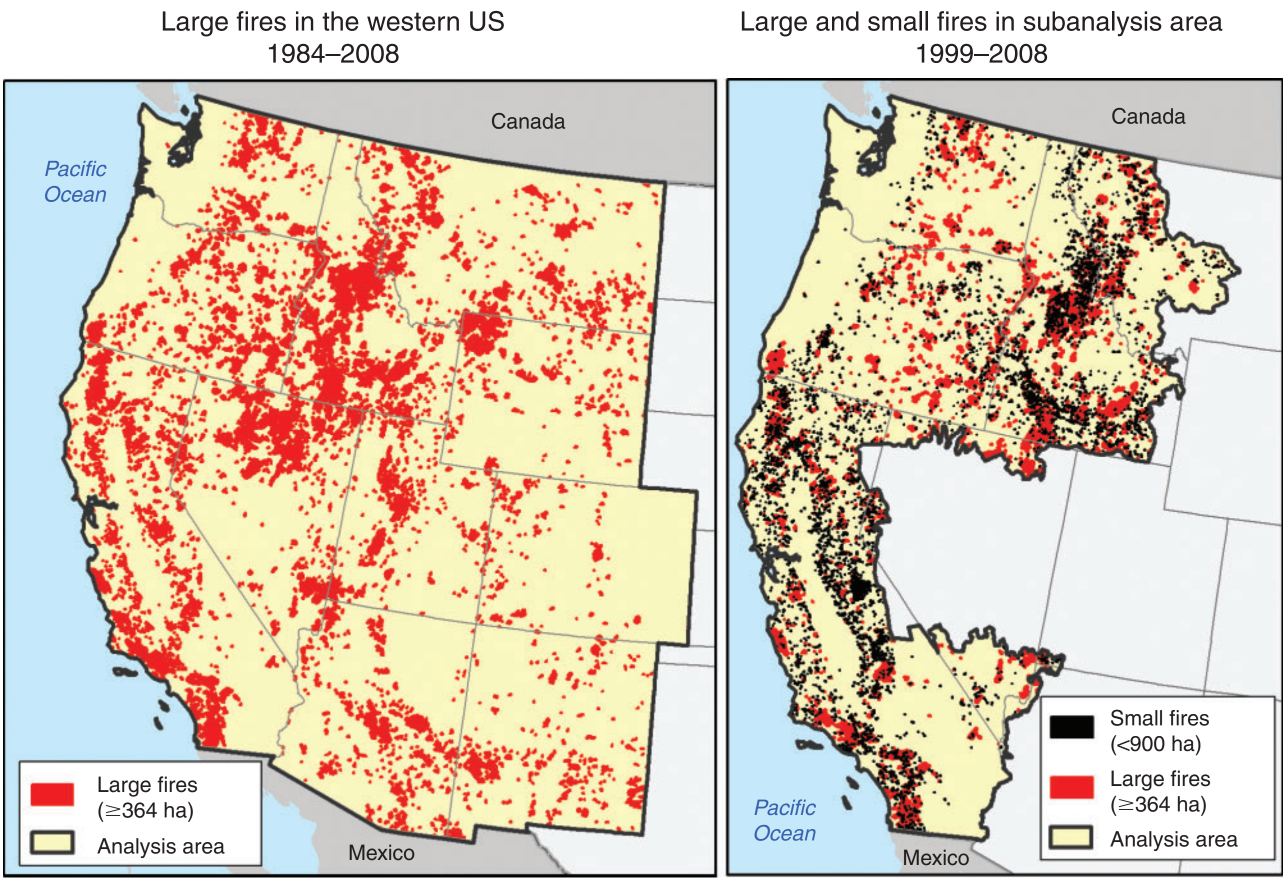

Fig. 2. Location of fires used in the main analysis (left panel; only fires ≥364 ha) and the Small-fires analysis (right panel; both large and small fires) within their respective study areas. Note that the outlines of the fires were exaggerated for visualisation purposes.

used an Albers NAD 1983 equal-area projection and were converted to a 1-km resolution.

Two wildfire probability models were constructed in order to assess the effects of modelling assumptions. The main reference model, termed the 'Full model', was built for the entire western USA and the entire span of fire data (1984–2008), and used all of the selected exploratory variables. Because many of the Full model's ignitions and vegetation variables encompassed a high degree of anthropogenic influence, we sought to examine the effect of these variables on wildfire probability by excluding them in the 'Non-anthropogenic model'. To ensure consistency among areas, only large fires (≥364 ha or 900 acres) were used in the Full and Non-anthropogenic models, because fires below this size threshold were not available for most of the study area.

We tested whether excluding small fires in our model building substantially affected wildfire probability by conducting an additional analysis, which was termed the 'Small-fires analysis.' Small fires (<364 ha), which often suffer from sporadic reporting in time and space, are assumed to have a negligible effect wildfire probability (e.g. Parisien and Moritz 2009), but this claim has never been formally examined. Although it is true that large wildfires are responsible for most of the area burned in the western USA (Stephens 2005), small fires are numerous and may burn in areas that rarely experience large fires (Bartlein *et al.* 2008). Examining small fires also may bridge the gap between wildfire probability as influenced by people and wildfire probability that could occur in the absence of fire suppression. The Small-fires analysis consisted of a comparison of two types of wildfire probability models: one created with only large fires and one created with small and large fire observations. Small fires data were available for a shorter time period than large fires (1999 to 2008) and for a smaller study area (Landfire 1.1.0 Events Geodatabase, US Department of Interior, Geological Survey, see http://www.landfire.gov, accessed 10 June 2010) (Fig. 2), but these were sufficient for our comparative purposes.

Data

The wildfire probability models' dependent variable consisted of 'presence' points that were randomly sampled within recent fire perimeters (see *Spatial modelling* section). Two datasets of mapped fire perimeters were used for this purpose: one for the Full model and Non-anthropogenic model and one for the two models of the Small-fires analysis. The Full and Non-anthropogenic models used the recently compiled data from the Monitoring Trends in Burn Severity (MTBS) project (Eidenshink *et al.* 2007). These data span the 1984–2008 period and cover the entire western USA. They also include prescribed burns, but for the fires ≥364 ha these only consist of ~7% of the observations (and ~2% of the total area burned). The Small-fires analysis combined the large fires from the MTBS for the 1999 to 2008 time period with the small fires (<364 ha) compiled for the Landfire project (http://www.landfire.gov).

Five variables were used to assess the role of ignitions on wildfire probability (Table 1). The only natural ignition source considered is lightning (Ltg_Dens) (Christian *et al.* 2003), whereas four proxies of anthropogenic ignition location were used: population density (Pop_Dens), road density (Rd_Dens), roadless volume (RdlsVol), and distance to the wildland–urban interface (WUI_Dist). Whereas the computation of Pop_Dens (Center for International Earth Science Information Network (CIESIN), Columbia University, and Centro Internacional de Agricultura Tropical (CIAT), Gridded Population of the World Version 3 (GPWv3): Population Density Grids, see http://sedac.ciesin.columbia.edu/gpw, accessed 12 June 2010) and Rd_Dens (ESRI 2008) is straightforward, RdlsVol is a transformation of the distance to road that better characterises the degree of isolation (Watts *et al.* 2007). WUI_Dist was obtained from a database of exurban residential development (Theobald and Romme 2007). Note that these variables capture both the potential for human ignitions and a measure of fire suppression effectiveness, as fires are better detected and more readily accessed in areas of high population and road density respectively (Syphard *et al.* 2007).

In this study, vegetation was used to represent the biomass available for burning. It was therefore desirable to capture the continuous nature of the biomass spectrum, rather than using vegetation classes. Parisien and Moritz (2009) have shown that using categorical vegetation variables can be problematic because of overfitting and because each class is considered equally similar. The first vegetation variable, the percentage land cover fuels (Fuel_Pct), was derived from a simple reclassification of the US GAP Analysis Land Cover (US Geological Survey 2010), where all cover types in which wildland fire spread is unusual were termed 'non-fuel' and all others were 'fuel.' All urban and agricultural areas were non-fuel, as were areas of sparse vegetation cover (e.g. deserts, alpine tundra) and permanent wetlands. This fuel–non-fuel classification was made continuous (i.e. percentage cover) by calculating the moving-windows surfaces from its original 30-m resolution grid. The other vegetation variable used was gross primary productivity (GPP) (Zhao *et al.* 2005), which is the rate at which plants store energy as biomass per unit time or, in other words, the capacity of ecosystems to produce flammable biomass. A major difference between Fuel_Pct and GPP is the degree of anthropogenic influence: the former is strongly affected by humans, whereas the latter is largely (but not entirely) a function of climate.

Climate variables were chosen to represent both the effect of climate on prevailing fuel moisture and its control on vegetation patterns. Although climate exerts both a direct and an indirect influence on fire, it is difficult (if not impossible) to distinguish between these effects at the spatiotemporal frame of this study. The climate variables consisted of metrics describing various permutations of temperature and precipitation (Table 1) (PRISM Group 2004). In addition to mean annual measures (Temp and PcpAnn), the extremes of monthly means of temperature and precipitation were computed. For example, the mean temperature of the wettest month (MeanTempWettest) provides an index of coincidence of a resource (moisture) and energy (heat), whereas the minimum temperature of the coldest month (MinTempColdest) quantifies stress to plants. Two variables were also used to characterise the growing season: the length of the season in days, as defined in McKenney *et al.* (2007), and the cumulative sum of degrees (growing degree days, GrowDegDays) ≥5°C.

Two water-balance metrics, annual evapotranspiration (AET) and water deficit (WatDef), were also used in the models. Considered together, these variables were shown to correlate well with ecosystem types of the western USA (Stephenson 1990) and individual tree species (Lutz *et al.* 2010). Reference evapotranspiration was calculated using the Penman–Monteith model (Allen *et al.* 1998), which uses temperature, radiation, precipitation and wind speed data. Water deficit was calculated as the monthly sum of the difference between reference evapotranspiration and precipitation such that no single month could have a water deficit less than zero. This technique generally follows Stephenson (1990), except that it did not account for soil water or carry-over from month to month. To assess fire spread potential, the 90th, 95th and 99th percentile wind speed were calculated for the mean driest (WindDriest) and warmest (WindWarmest) months, from Kalnay *et al.* (1996). Finally, extreme values of the monthly standardised precipitation index (MinSPI and MaxSPI) measured the departure from mean precipitation normals (McKee *et al.* 1993).

A single explanatory variable, the surface-area ratio (cf. Stambaugh and Guyette 2008), was used to characterise the effect of topographic roughness on wildfire probability. At the spatial extent and resolution of the present study, the effect of topography on fire activity is largely indirect: it exerts its effect on fire mainly by influencing patterns in ignitions, vegetation and weather. However, it can be used as a proxy for several variables that may be missing from the model (Parisien *et al.* 2011). The surface-area ratio variable (SurfArea_Ratio) was calculated from a digital elevation model (US Geological Survey 2000). Because the calculation of this measure is strongly scale-dependent, the moving-window approach was used, as described for ignition and vegetation variables, whereby topographic roughness is measured at four spatial scales: 1, 100, 1000 and 10 000 ha.

Many of the potential explanatory variables initially considered for the modelling (Appendix 1) were similar and thus highly correlated. To avoid incorporating a large number of variables that have overlapping information, we selected a relatively parsimonious subset of variables for model building. This was achieved in a heuristic manner by first cross-correlating the variables and identifying those that were highly correlated (Spearman $R > 0.6$). Within each of these groups, we retained the variable that performed the best in a MaxEnt model that considered only that explanatory variable. As such, a limited set of fairly uncorrelated yet complementary environmental variables were included in the model. In the Non-anthropogenic model, the Pop_Dens, RdlsVol, Fuel_Pct variables were excluded from the Full model to examine the anthropogenic influence on wildfire likelihood, whereas the two models of the Small-fires analysis included the same variables as the Full model.

Spatial modelling

Wildfire probability models were computed in MaxEnt 3.3.3e (Phillips *et al.* 2006). MaxEnt is designed for presence-only data. Presence-only models discern between the environment of burned areas ('fire presence') from that of the entire study area ('background'), as opposed to discerning between burned and unburned areas. Fire presences, which represent the dependent variable, were point-based observations obtained by randomly sampling point locations within fires perimeters. In a presence-only framework, the lack of fire at a given location is not interpreted as an 'absence' of fire by the model, as some of these areas may in fact experience wildfire if they share environmental characteristics with other wildfire-prone locations.

At each fire presence point, MaxEnt estimates wildfire probability by fitting the probability distribution of maximum entropy (the one that is most uniform) to the environmental variables. The algorithm iteratively evaluates the contrasts between the values of the fire presences and those of a background. MaxEnt also has the flexibility to fit non-linear relationships between the response variable and explanatory variables, so that resulting models have the ability to describe complex relationships. However, environmental conditions that exceeded the range of currently observed values were 'clamped' (i.e. held constant) at the maximum value of the range to avoid unfounded extrapolations of wildfire probabilities.

The MaxEnt output represents an estimate of relative, rather than absolute, wildfire probability. Because the models are based on fire patterns and environmental data for a fairly long time period (1984–2008), the mapped fire probabilities are not designed to evaluate specific sets of conditions that lead to large fires in a given year, but instead quantify the relative wildfire likelihood over longer periods (Krawchuk *et al.* 2009; Parisien and Moritz 2009). The probability is relative in that a temporal scale (e.g. 1 year) is not implied. Rather, the probabilities among pixels are relative to one another; that is, a pixel with a wildfire probability of 0.3 is estimated to be three times as likely to experience a fire as a pixel with a value of 0.1.

The entire pool of fire presences consisted of 10 000 points that were randomly located within the burned areas from 1984 to 2008, whereas 50 000 random points were used to characterise the background environment. The same presence points were used for the Full model and Non-anthropogenic model; only the input set of explanatory variables differed between the two models. For the Small-fires analysis, 5000 and 20 000 points were used as fire presences and background respectively.

The effect of spatial autocorrelation in the fire data and the explanatory variables was minimised through the following steps. Only a small random fraction of the total fire presences was used to build individual wildfire probability models. This was replicated for 25 bootstrap subsamples and the ensemble of resulting models was ultimately averaged for analysis. We used Ripley's K function with different-sized subsets to estimate the sampling fraction at which the fire observations were spatially independent. Last, we determined that 500 points sufficiently captured fire–environment relationships while being only faintly clustered across the study area. The fraction of points unused for model building was instead used to calculate the evaluation metrics described below.

The predictive accuracy of the wildfire probability model output was evaluated using several metrics that were computed and averaged for each of the 25 model replicates. The estimated fraction of the area suitable for fire (an approximation of the false positive rate or 1 – specificity) and the omission (false negative rate or 1 – sensitivity) were measured at the wildfire probability threshold that minimises the sum of these error measurements (Liu *et al.* 2005). Interpreted together, these measures give us the expected rate of false negatives for a given predicted suitable area.

A more comprehensive measure of model performance is the area under the curve (AUC) of a plot of sensitivity (true positives) *v.* 1 – specificity (false positives or 1 – true negatives). In a presence–absence framework, the AUC computed with the points unused for model building (i.e. the 'test' AUC) may range from 0.5, where prediction accuracy is no better than if samples were randomly selected, to 1, which indicates perfect classification accuracy. By contrast, in a presence-only framework, as in this study, it is impossible to achieve unity in AUC because absences (hence false positives) are unknown. In fact, the maximum achievable AUC is equal to $1 - a/2$, where a is generally the fraction of the study area that the species (or process) covers (i.e. the prevalence), a measure that is usually unavailable. However, here, we considered a to be the percentage of pixels where fire was observed. This provides a fair, yet underestimated, approximation of prevalence.

Finally, we computed the correlation between the wildfire probability predictions and 1–0 (i.e. fire–no fire) observations, which is known as the point-biserial correlation. Rather than being based solely on rank, such as the AUC, the point-biserial correlation uses the actual predictions to evaluate model performance. To obtain this measure, the 50 000 random points used for the background were assigned a 1 or a 0 whether they were located in a fire pixel or a fire-free pixel respectively. These values were then correlated with the predicted wildfire probability. Although the point-biserial correlation does not provide a stand-alone interpretable value, it does provide a good metric for the comparison of the predictive ability among models.

Data analysis

The relative contribution of the explanatory variables to wildfire probability was assessed in MaxEnt by estimating the change in model gain associated with each variable. The mean and standard deviation (s.d.) of the percentage contribution of each variable were compiled from the 25-model ensemble and were plotted for both the Full model and the Non-anthropogenic model.

To evaluate the relationship between fire activity and environmental factors, the estimated wildfire probability was plotted as a function of a selected explanatory variable. These plots were produced from MaxEnt models where a single explanatory variable was used to predict wildfire probability. This said, these response curves were built to visualise bivariate relationships and were not those used in the wildfire probability models, which are subject to complex interactions with other variables. Twenty-five replicates of the bivariate models were built and the mean and standard deviation of the 25 response curves were plotted.

Wildfire probability maps were produced for the Full and Non-Anthropogenic models, as well as for the models of the Small-fires analysis. The Full and the Non-anthropogenic models were compared on a pixel-wise basis in two ways: first, through simple subtraction of the fire probabilities (absolute change) and, second, by computing the relative change in probability, expressed as percentage change (relative change). The mapped outputs of both Small-fires analysis models (large fires only and large and small fires) were qualitatively compared to visualise the effect of adding small fires to the wildfire probability map. In addition, the variable contributions (mean and s.d.) of these models were plotted and compared, and the rank-order correlation of the variable contributions was assessed (Spearman correlation).

Results

There were a wide range of responses of wildfire probability to explanatory variables, as shown in six of the variables most influential in computing wildfire probability (see below) (Fig. 3). Although some responses were monotonic, notably that of the top variable (Fuel_Pct100), this was not the norm. Many of the variables appeared to be unimodal, whereby the fire response is maximised across intermediate values of the explanatory variable (e.g. MeanTempDriest, PcpColdest, RdlsVol10000, GPP100). In general, weak variables had fire responses that were highly complex and unintuitive (not shown).

Model evaluation metrics show that the Full model and Non-anthropogenic model performed similarly well (Table 2). Suitable area and omission error suggest that when approximately one-third of the study area was considered 'suitable' for fire, approximately one-quarter of the points were predicted to be false positives. The uncorrected AUCs computed from the test portion of observations were 0.792 and 0.742 for the Full and Non-anthropogenic models respectively. However, when adjusted for wildfire prevalence (6.7% of the area), the AUCs were 0.836 and 0.810. AUCs are therefore somewhat larger for the Full model, which suggests that anthropogenic variables are informative with respect to recent wildfire occurrence patterns. The slight superiority of the Full model is also reflected in the point-biserial correlation.

Patterns in modelled wildfire probability are highly heterogeneous throughout the study area (Fig. 4). Wildfire likelihood is at least moderately high ($>$0.3) in most areas, with the exception of the south-west deserts, parts of the coastal northwest and in the expansive agricultural areas, most of which were extensive grasslands in the past.

Explanatory variables from each of the four main environmental factors, ignitions, climate, topography and vegetation, appear to be important in predicting wildfire probability (Fig. 5). The Fuel_Pct100 variable explained nearly 25% of variation in the Full model. However, several other variables had important contributions (e.g. PcpColdest, RdlsVol10000, MeanTempDriest and SurfArea_Ratio1), whereas a few had negligible contributions (e.g. Isotherm, MaxSPI and WindDriest99). Of the variables common to both models, those important for the Full model also contributed substantially to the Non-anthropogenic model, but some variables (SurfArea_Ratio1, GPP100) were disproportionately more important in the latter model.

Wildfire probability patterns between the Full model and the Non-anthropogenic model are broadly similar across the study area (Fig. 4*a*, *b*), but numerous differences are evident at local scales. The absolute change (i.e. subtraction) in wildfire probability between the Full and Non-anthropogenic models shows that over the majority of the landscape, the anthropogenic variables included in the Full model served to reduce wildfire probability, with increases in wildfire probability concentrated

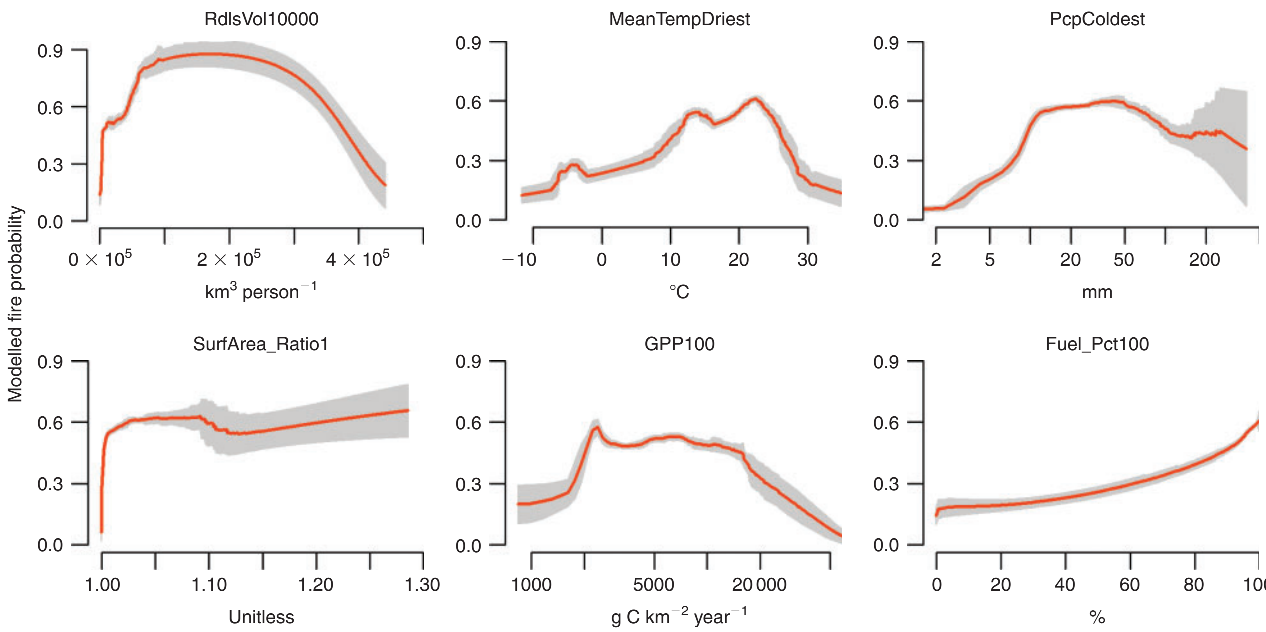

Fig. 3. Predicted wildfire probability for a selected set of explanatory variables. The full variable names and descriptions are found in Table 1. These plots were obtained by building MaxEnt models using only a single independent variable of interest. The line indicates the mean wildfire probability values, whereas the grey shading represents ±s.d., as calculated from 25 replicate runs using random subsets of the data. Note that the *x*-axis of the PcpColdest and GPP100 variables was logarithmically scaled.

Table 2. Performance of MaxEnt models, including maximum test sensitivity plus specificity area

The AUC is the area under the curve of the sensitivity *v.* predicted area (1 – specificity) plot; the 'adjusted AUC' adjusts these values according to geographic prevalence (see Methods). The 'suitable area' represents the fraction of area predicted as suitable and the 'omission error' is the fraction of presence points found in areas predicted to be unsuitable. The probability threshold is minimised according to the sum of these values. The point-biserial correlation evaluates the correspondence between estimated wildfire probability of presence points (ones) and non-presence points (zeros)

	Western USA analysis		Small-fire analysis	
	Full model	Non-anthropogenic model	Large fires	Large + small fires
Suitable area (%)	30.5	32.9	30.0	29.7
Omission error (%)	23.3	25.7	24.2	26.1
AUC	0.792	0.742	0.806	0.795
Adjusted AUC	0.836	0.810	0.839	0.830
Point-biserial correlation	0.346	0.297	0.354	0.348

in large wilderness areas (Fig. 4*c*). The map of relative change in wildfire probability among models shows that most decreases in wildfire probability in the Non-anthropogenic model compared with the Full model occur in developed and agricultural areas, as expected (Fig. 4*d*). In fact, the most drastic decreases (>500%) almost always occur in and around urban areas. In contrast, only minor relative increases in wildfire probability were observed over much of the study area when anthropogenic variables were removed.

The Small-fires analysis, which compared the wildfire probability of models built with only the large fires and with both large and small fires, indicated strong similarities between model outputs (Fig. 6). The spatial predictions were virtually identical between models; differences were highly localised and, even so, relatively minor (e.g. coastal Oregon). Likewise, the variable contributions were extremely similar, as the rank-order correlation between variable ranks was $R = 0.95$ (Spearman correlation).

Discussion

Fire–environment relationships in the western US

This study brings us one step closer to understanding the sub-continental controls on long-term (i.e. multi-decadal) wildfire likelihood in North America. The results provide further support to the idea that large-scale assessments of wildfire likelihood are best described using all of the necessary fire ingredients: an ignition source, hot and dry weather, and sufficient biomass for sustained combustion. Fire universally requires the spatio-temporal coincidence of these same basic elements; however,

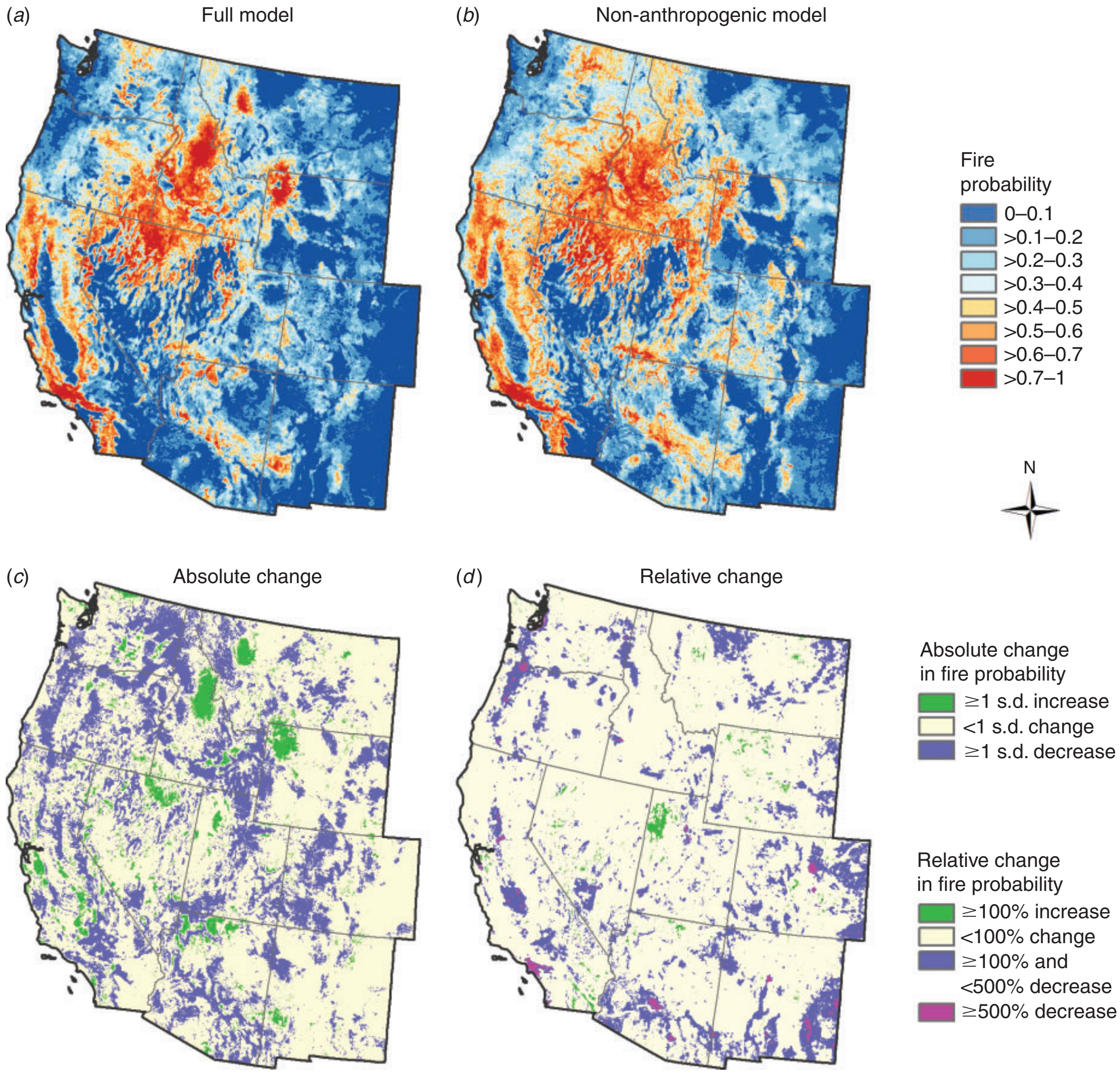

Fig. 4. Mean predicted wildfire probability (based on 25 model replicates) for the Full model (*a*); the Non-anthropogenic model (*b*); the absolute change (*c*); and the relative change (*d*) from the Full model to the Non-anthropogenic model, whereby green indicates an increase and blue represents a decrease in wildfire probability as a result of human-influenced variables. The wildfire probability maps produced in Fig. 4 and their projection information are available as supplementary material to this paper (see http://www.publish.csiro.au/?act=view_file&file_id=WF11044_AC.zip).

understanding the specific effect of each of the variables chosen is not straightforward. The degree to which elements limit fire varies substantially across subcontinental extents, as shown in Australia (Russell-Smith *et al.* 2007) and subequatorial Africa (Archibald *et al.* 2009). Although evaluating region-specific limits on fire activity was not the focus of the present study, our results are strongly coherent with those of previous studies that show that fire–environment relationships are far from constant throughout the western US (McKenzie *et al.* 2004; Littell *et al.* 2009).

It is not surprising that the strongest predictor of wildfire likelihood in the western USA is the percentage cover of fuels. However, this predictor only explains ~25% of model variation. Our results suggest that gradients in topography, climate, biomass (i.e. GPP) and ignitions also play important roles in characterising wildfire likelihood across the study area. In fact, when the percentage fuels variable was omitted from the model (i.e. in the Non-anthropogenic model), only a slight decrease in model performance was observed, suggesting that the percentage fuels information was reflected in a combination of other variables. These results are thus coherent with those of Russell-Smith *et al.* (2007) in Australia and Parisien and Moritz (2009) in the conterminous USA, who found that, because they are dependent on climate and human land use, measures of biomass or flammable vegetation can be omitted from subcontinental models of fire activity.

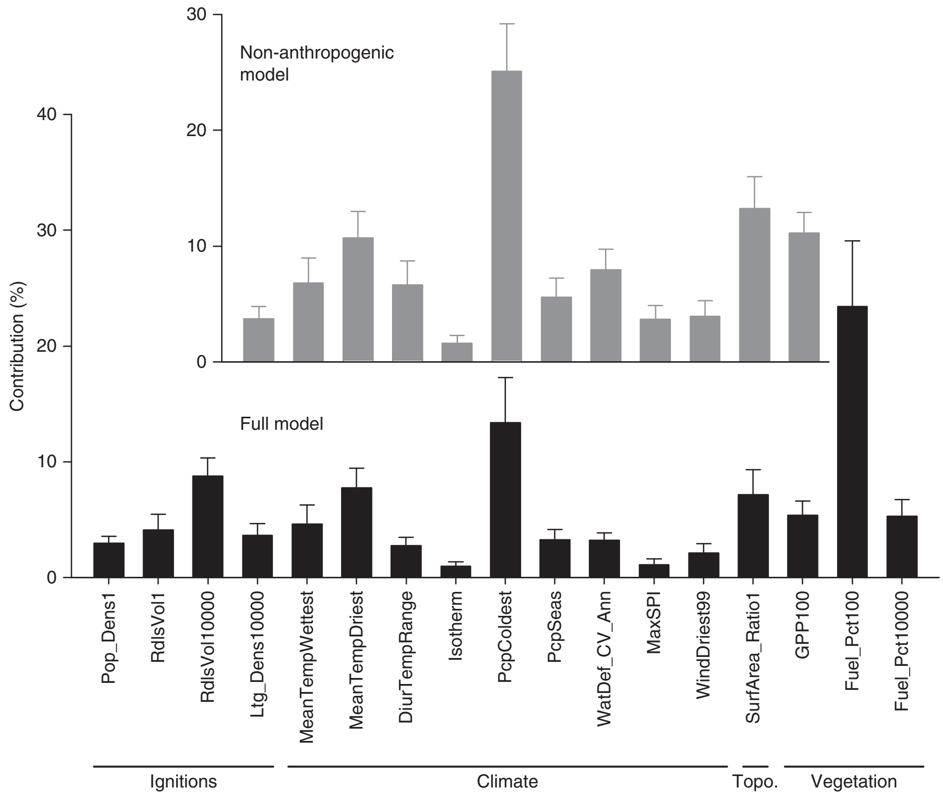

Fig. 5. Variable contributions, expressed as a percentage, of the Full model (bottom plot), and the Non-anthropogenic model (top plot).

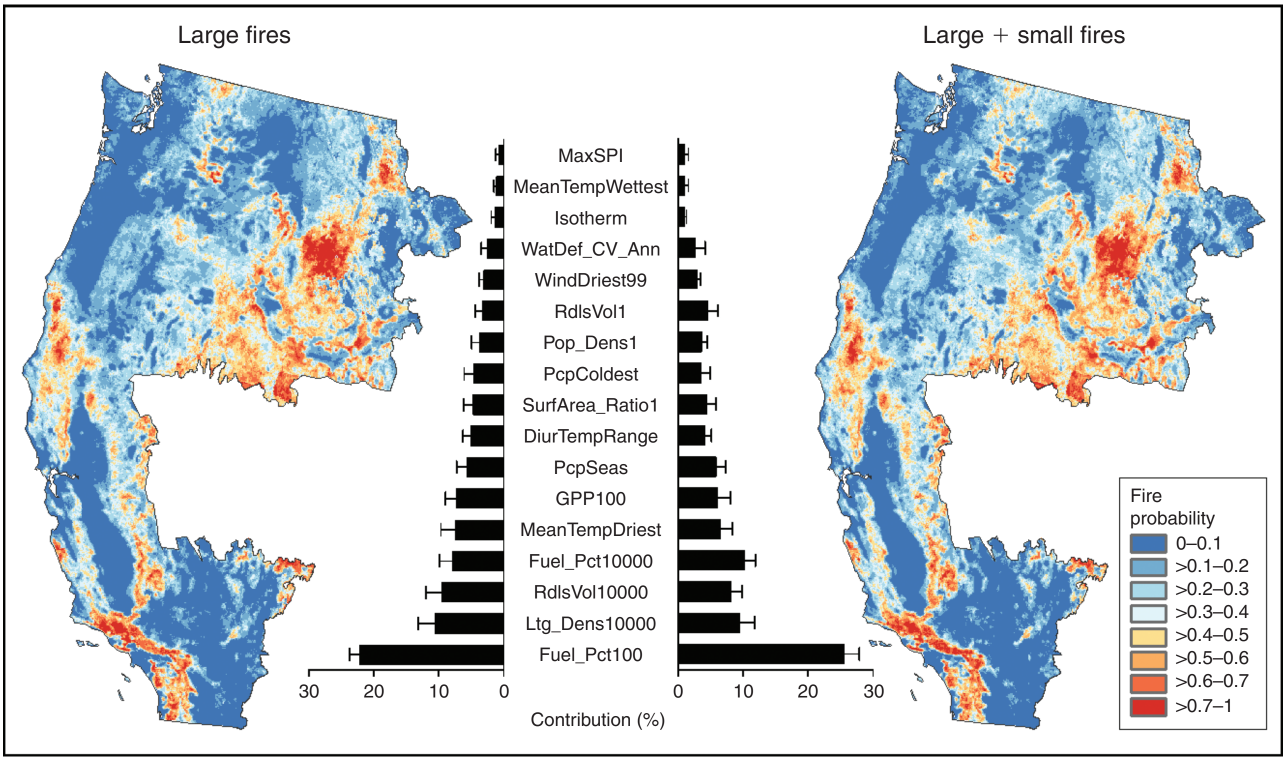

Fig. 6. Wildfire probability maps of the Small-fires analysis study area using only large fire (≥364 ha) data (left) and data from both small and large fires (right). The bar chart represents the comparison of variable contributions for both models. The wildfire probability maps represent the average of 25 model replicates.

Our results further underscore the complex relationship of fire with climate. Despite including a vast array of climate variables in our initial variable selection and retaining nine of them for model building, no single climate variable provided a ‘universal’ or even dominant link to fire in the western USA. This suggests that same set of variables does not operate in the same manner across the entire domain. For instance, the climate drivers in the Pacific Northwest are very different from those in the Southwest. Synthetic variables such as water balance metrics (AET and water deficit) that are strong predictors of the geographical distribution of vegetation types (Stephenson 1998) and annual area burned in some parts of the US Interior West (Littell and Gwozdz 2011), did not, for the most part, graduate to our models of long-term wildfire likelihood. These variables were accounted for by other variables depicting the coincidence of moisture and energy (e.g. precipitation of the coldest month). We surmise that this lack of importance of water balance metrics is blurred by their somewhat bipolar relationship with fire. For example, although recent warming has led to increases in fire activity in some areas, notably through lengthening of drought periods and the fire season (Westerling *et al.* 2006), they may also impose moisture stress that will ultimately limit the continuity of fuels (Crimmin *et al.* 2011).

In an area as topographically diverse as the western USA, it is expected that relief would be a strong predictor of wildfire likelihood. Our results are consistent with those of Stambaugh and Guyette (2008) in the Missouri Ozarks and Dickson *et al.* (2006) in Arizona, who both reported a significant positive association between topographic roughness and wildfire likelihood. In contrast, a negative fire–roughness association was reported in boreal Canada (Parisien *et al.* 2011) and, though weaker, in subequatorial Africa (Archibald *et al.* 2009). These contradictory results suggest that, at the spatiotemporal scale of the present study, topography does not exert a direct influence on wildfire likelihood (i.e. through wind–slope interactions) but rather acts as a proxy for other controls. Everything else being constant, one would expect rugged topography to hinder the spread of large fires. However, in the western USA, rough topography is often associated with less anthropogenic modification of fuel (i.e. fuel fragmentation as a result of land use) and decreased access for fire suppression. Almost all agriculture lands, which represent the most expansive anthropogenic influence in the western USA (Leu *et al.* 2008), occur on flat terrain. In biomes that are largely unsuitable for agriculture, such as boreal forest, flat areas are generally more prone to fire spread than rugged ones, because they typically have a more continuous cover of flammable vegetation.

Compared with climate and vegetation, ignitions generally appeared to play a relatively minor role in predicting wildfire probability in the western USA. Lightning contributed only moderately to the models, but this was largely expected, as there does not appear to be any evident associations in the spatial patterns of area burned and lightning density across the western USA (Christian *et al.* 2003). In fact, because it is usually accompanied by rainfall, lightning density is largely negatively correlated with fire occurrence. Although ‘dry’ lightning would provide a more meaningful metric of lightning ignition potential (Rorig and Ferguson 1999), these data were not available for the spatiotemporal scale of the current study.

In contrast to lightning, human ignition proxies contributed significantly to wildfire probability. In general, our results are consistent with those of Syphard *et al.* (2007), who showed that the likelihood of fire has a non-linear relationship with human influence in California, whereby peak fire activity is associated with intermediate levels of human influence. Although the measures of Syphard *et al.* (2007) are different (i.e. they used population density and fire counts), the best-performing measure of human ignitions used here, the roadless volume (10 000-ha moving window), exhibits a similar inverse-U-shaped relationship to fire. Unlike what was reported in the Upper Midwest (Cardille *et al.* 2001; Sturtevant and Cleland 2007), increased remoteness was not always associated with increased area burned in the western USA. This phenomenon is the result of the low vegetation cover (deserts or high-elevation areas) of some remote areas and perhaps to a limitation of natural ignitions (lightning) in others.

The patterns of decreased wildfire probability as a result of anthropogenic variables exhibit fairly strong spatial concordance with the areas of highest human footprint intensity (Leu *et al.* 2008). Decreases in wildfire probability as a result of anthropogenic variables were, as expected, associated with areas of urban development or expansive agriculture, this phenomenon being particularly pronounced in and around major cities. Although it is impossible to accurately measure the overall effect of humans on area burned in an area as large and complex as the western USA, our results seem to suggest that, given the widespread land-use change, current human activities may have an overall inhibiting effect on wildfire likelihood, as reported in California (Stephens *et al.* 2007). Guyette *et al.* (2002) have proposed a temporal framework to describe the anthropogenic effects on fire regimes in the Missouri Ozarks that may indeed apply to much of the western USA. In brief, they suggest that many areas have experienced a human-dominated surge in fire activity once they became inhabited, but that this increase was eventually met with fuels limitation as human density (and burning) increased, and, in the later stages, decreased abruptly as a result of fuel fragmentation and fire suppression.

Methodological aspects of wildfire probability modelling

Although the wildfire probability presented in this study was deliberately presented as a relative measure of wildfire likelihood, it is possible to transform pixel-wise probabilities into expected values of fire frequency or annual area burned (i.e. something closer to an absolute value). For instance, a simple method of obtaining an annual wildfire probability is through a linear transformation in which the value of each pixel is multiplied by a scalar obtained from the ratio of the mean annual fire frequency calculated from historical data (say 1984–2008) to that of the mean relative probability. This technique adjusts the relative wildfire probability according to the historical area burned by directly manipulating the final predictions. Another way to obtain an annual wildfire probability from the MaxEnt predictions is to apply an adjustment factor to the logit of the prediction that results in a mean output probability equal to that of the prevalence (i.e. annual area burned) (Elith *et al.* 2011). This more complicated correction can accommodate non-linear changes, but its behaviour relative to a simple scalar remains unexplored. Regardless, both techniques are prone to the same uncertainties and assumptions regarding prevalence – the 'true' distribution of fire – which is unknown. For example, they both assume that the current or recent burning rate calculated from fire data is reliable (typically not the case) and stationary through time. As such, we prefer to use the more conservative relative wildfire probability estimates.

Incorporating better and more comprehensive information into fire–environment models is likely to improve estimates of wildfire likelihood. However, the accuracy of wildfire likelihood predictions is, as always, contingent on model assumptions and data quality. The spatial predictions presented here should be interpreted for the specific spatiotemporal frame of study and should not necessarily be assumed to remain stable far into the future. Nonetheless, the wildfire probability predictions do appear to be fairly robust. Modelling explorations in which different sampling point subsets and different variables were included yielded strongly consistent patterns of fire probabilities (results not shown), as long as key information (i.e. climate) was not omitted. In fact, a comparison of the Full model's spatial predictions with those of the Small-fires analysis shows broadly similar patterns, even if the latter only uses a fraction of the fire observations (10 *v.* 25 years of data).

Importantly, the Small-fires analysis indicates that the current standard of excluding small fires ($<$364 ha) when assessing wildfire probability at subcontinental scales is a reasonable modelling shortcut. Currently, data on small fires is limited in availability and quality; thus, the time and expense of including small fires in wildfire probability models is often cost-prohibitive. Our Small-fires analysis results are coherent with the general belief that small fires contribute a small proportion of the total area burned in the western USA ($\sim$4% of the area burned of the Small-fires analysis study area), although this fraction is larger in the coastal Northwest (Strauss *et al.* 1989), and that models with and without the small fires yield virtually identical fire probabilities and have similar environmental drivers.

Conclusion

The use of comprehensive and spatially precise wildfire data, a wide array of variables characterising the wildfire environment, and a distribution modelling technique that can capture complex relationships (MaxEnt) should yield a gain in accuracy compared with previous wildfire probability distributions of the study area. The wildfire probability maps of this study indeed represent some of the most spatially refined estimates of wildfire likelihood for the study area to date. Reassuringly, the predicted patterns of wildfire likelihood are coherent with our current understanding of western USA fire regimes and the explanatory variables generally support what other studies of wildfire–environment relationships have found. Our results emphasise the complexity in these relationships and support the idea that area-specific fire–environment linkages in the western USA appear to be as variable as the fire regimes themselves (Hardy *et al.* 2001; Littell *et al.* 2009).

The wildfire probability estimates of this study are not those of a 'pristine' system but that of the modern and human-dominated area (Cardille and Lambois 2010). Although we cannot claim to have elucidated the effect of anthropogenic influence on western USA wildfire likelihood, adding information pertaining to the human imprint significantly affected wildfire probability patterns. In spite of the inherent difficulty in assessing the fire-proneness of areas where fire has been largely excluded, modelling can be used to at least partly fill in the gap either across a landscape (Collins *et al.* 2010) or, as shown in this study, at a subcontinental extent. Furthermore, the wildfire probability computed in this study provides a likelihood estimate that can be used in the strategic planning of fire or land-management activities. For example, a reliable spatial assessment of wildfire likelihood can help prioritise fuel treatment placement. From a land-management perspective, mapped wildfire probability is invaluable in the planning of urban areas, roads, forestry operations and even ecological restoration. In addition, with respect to the last item, wildfire likelihood estimates can help predict whether fire is likely to have been excluded from a naturally fire-prone area and should, therefore, be part of a restoration plan.

Acknowledgements

We are grateful to all our colleagues and agencies that provided data for the analysis. In particular, we thank David Theobald for sharing his wildland–urban interface map, as well as Pia Papadopol and Kevin Lawrence for producing the growing season variables. We are also indebted to Robert Hijmans, who adapted the R 'raster' package to fit our needs. Jennifer Lecker, Brenda Lundberg and Brad Quayle provided excellent and on-going support for the fire data processing. This study was partly funded by the Joint Fire Science Program (Project 06–4-1–04) and the Wilburforce Foundation.

References

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration: guidelines for computing crop water requirements. Food and Agriculture Organization of the United Nations, FAO Irrigation and Drainage Paper 56. (Rome, Italy)

Archibald S, Roy DP, van Wilgen BM, 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

Balshi MS, Mcguire AD, Duffy P, Flannigan M, Walsh J, Melillo J (2009) Assessing the response of area burned to changing climate in western boreal North America using a Multivariate Adaptive Regression Splines (MARS) approach. *Global Change Biology* **15**, 578–600. doi:10.1111/J.1365-2486.2008.01679.X

Barbour MG, Billings WD (Eds) (2000) 'North American Terrestrial Vegetation, 2nd Edn.' (Cambridge University Press: Cambridge, UK)

Bartlein PJ, Hostetler SW, Shafer SL, Holman JO, Solomon AM (2008) Temporal and spatial structure in a daily wildfire-start data set from the western United States (1986–96). *International Journal of Wildland Fire* **17**, 8–17. doi:10.1071/WF07022

Bradstock RA (2010) A biogeographic model of fire regimes in Australia: current and future implications. *Global Ecology and Biogeography* **19**, 145–158. doi:10.1111/J.1466-8238.2009.00512.X

Brown KJ, Clark JS, Grimm EC, Donovan JJ, Mueller PG, Hansen BCS, Stefanova I (2005) Fire cycles in North American interior grasslands and their relation to prairie drought. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 8865–8870. doi:10.1073/PNAS.0503621102

Cardille JA, Lambois M (2010) From the redwood forest to the Gulf Stream waters: human signature nearly ubiquitous in representative US landscapes. *Frontiers in Ecology and the Environment* **8**, 130–134. doi:10.1890/080132

Cardille JA, Ventura SJ, Turner MG (2001) Environmental and social factors influencing wildfires in the Upper Midwest, United States. *Ecological Applications* **11**, 111–127. doi:10.1890/1051-0761(2001)011[0111:EASFIW]2.0.CO;2

Christian HJ, Blakeslee RJ, Boccippio DJ, Boeck WL, Buechler DE, Driscoll KT, Goodman SJ, Hall JM, Koshak WJ, Mach DM, Stewart MF (2003) Global frequency and distribution of lightning as observed from space by the Optical Transient Detector. *Journal of Geophysical Research* **108**, 4005. doi:10.1029/2002JD002347

Collins BM, Stephens SL, Moghaddas JJ, Battles J (2010) Challenges and approaches in planning fuel treatments across fire-excluded forested landscapes. *Journal of Forestry* **108**, 24–31.

Crimmins SM, Dobrowski SZ, Greenberg JA, Abatzoglou JT, Mynsberge AR (2011) Changes in climatic water balance drive downhill shifts in plant species' optimum elevations. *Science* **331**, 324–327. doi:10.1126/SCIENCE.1199040

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

Eidenshink J, Schwind B, Brewer K, Zhu Z-L, Quayle B, Howard S (2007) A project for monitoring trends in burn severity. *Fire Ecology* **3**, 3–21. doi:10.4996/FIREECOLOGY.0301003

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Diversity & Distributions* **17**, 43–57. doi:10.1111/J.1472-4642.2010.00725.X

ESRI (2008) 'Tele Atlas North America: StreetMap 2008 North America.' (ESRI: Redlands, CA)

Finney M, Grenfell IC, McHugh CW (2009) Modeling containment of large wildfires using generalized linear mixed-model analysis. *Forest Science* **55**, 249–255.

Finney MA, McHugh CW, Grenfell IC, Riley KL, Short KC (2011) A simulation of probabilistic wildfire risk components for the continental United States. *Stochastic Environmental Research and Risk Assessment* **25**, 973–1000. doi:10.1007/S00477-011-0462-Z

Gedalof Z, Peterson DL, Mantua NJ (2005) Atmospheric, climatic, and ecological controls on extreme wildfire years in the north-western United States. *Ecological Applications* **15**, 154–174. doi:10.1890/03-5116

Girardin MP, Ali AA, Carcaillet C, Mudelsee M, Drobyshev I, Hely C, Bergeron Y (2009) Heterogeneous response of circumboreal wildfire risk to climate change since the early 1900s. *Global Change Biology* **15**, 2751–2769. doi:10.1111/J.1365-2486.2009.01869.X

Guyette RP, Muzika RM, Dey DC (2002) Dynamics of an anthropogenic fire regime. *Ecosystems* **5**, 472–486.

Hardy CC, Schmidt KM, Menakis JP, Sampson RN (2001) Spatial data for national fire planning and fuel management. *International Journal of Wildland Fire* **10**, 353–372. doi:10.1071/WF01034

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

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Leetmaa A, Reynolds R, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Jenne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**, 437–471. doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Krawchuk MA, Moritz MA, Parisien M-A, Van Dorn J, Hayhoe K (2009) Global pyrogeography: the current and future distribution of wildfire. *PLoS ONE* **4**, e5102. doi:10.1371/JOURNAL.PONE.0005102

Leu M, Hanser SE, Knick ST (2008) The human footprint in the West: a large-scale analysis of anthropogenic impacts. *Ecological Applications* **18**, 1119–1139. doi:10.1890/07-0480.1

Littell JS, Gwozdz R (2011) Climatic water balance and regional fire years in the Pacific Northwest, USA: linking regional climate and fire at landscape scales. In 'The Landscape Ecology of Fire, Ecological Studies 213'. (Eds D McKenzie, C Miller, DA Falk) pp. 117–139. (Springer Science+Business Media)

Littell JS, McKenzie D, Peterson DL, Westerling AL (2009) Climate and wildfire area burned in western US ecoprovinces, 1916–2003. *Ecological Applications* **19**, 1003–1021. doi:10.1890/07-1183.1

Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **28**, 385–393. doi:10.1111/J.0906-7590.2005.03957.X

Lutz JA, van Wagtendonk JW, Franklin JF (2010) Climatic water deficit, tree species ranges, and climate change in Yosemite National Park. *Journal of Biogeography* **37**, 936–950. doi:10.1111/J.1365-2699.2009.02268.X

Marlon JR, Bartlein PJ, Carcaillet C, Gavin DG, Harrison SP, Higuera PE, Joos F, Power MJ, Prentice IC (2008) Climate and human influences on global biomass burning over the past two millennia. *Nature Geoscience* **1**, 697–702. doi:10.1038/NGEO313

McKee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration of time scales. In 'Proceedings of the Eighth Conference on Applied Climatology', 17–23 January 1993, Anaheim, CA. pp. 179–186. (American Meteorological Society: Boston, MA)

McKenney DW, Papadopol P, Lawrence K, Campbell K, Hutchinson MF (2007) Customized spatial climate models for Canada. Canadian Forest Service, Natural Resources Canada, Great Lakes Forestry Centre, Technical Note 108. (Sault Ste Marie, ON)

McKenzie D, Gedalof Z, Peterson DL, Mote P (2004) Climatic change, wildfire, and conservation. *Conservation Biology* **18**, 890–902. doi:10.1111/J.1523-1739.2004.00492.X

Mermoz M, Kitzberger T, Veblen TT (2005) Landscape influences on occurrence and spread of wildfires in Patagonian forests and shrublands. *Ecology* **86**, 2705–2715. doi:10.1890/04-1850

Meyn A, Schmidtlein A, Taylor SW, Girardin MP, Thonicke K, Cramer W (2010) Spatial variation of trends in wildfire and summer drought in British Columbia, Canada, 1920–2000. *International Journal of Wildland Fire* **19**, 272–283. doi:10.1071/WF09055

Moritz MA, Morais ME, Summerell LA, Carlson JM, 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

Parisien M-A, 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

Parisien M-A, Peters VS, Wang Y, Little JM, Bosch EM, Stocks BJ (2006) Spatial patterns of forest fires in Canada, 1980–1999. *International Journal of Wildland Fire* **15**, 361–374. doi:10.1071/WF06009

Parisien M-A, Parks SA, Krawchuk MA, Flannigan MD, Bowman LM, Moritz MA (2011) Scale-dependent factors controlling area burned in boreal Canada. *Ecological Applications* **21**, 789–805.

Phillips SJ, Anderson RP, Shapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**, 231–259. doi:10.1016/J.ECOLMODEL.2005.03.026

Preisler HK, Burgan RE, Eidenshink JC, Klaver JM, Klaver RW (2009) Forecasting distributions of large Federal-lands fires utilizing satellite and gridded weather information. *International Journal of Wildland Fire* **18**, 508–516. doi:10.1071/WF08032

PRISM Group (2004) Climate normals, 1971–2000. (Oregon State University: Corvallis, OR). Available at http://www.prismclimate.org [Verified 1 June 2010]

Rorig ML, Ferguson SA (1999) Characteristics of lightning and wildland fire ignition in the Pacific Northwest. *Journal of Applied Meteorology* **38**, 1565–1575. doi:10.1175/1520-0450(1999)038<1565:COLAWF>2.0.CO;2

Russell-Smith J, Yates CP, Whitehead PJ, Smith R, Craig R, Allan GE, Thackway R, Frakes I, Cridland S, Meyer MC, Gill AM (2007) Bushfire 'Down Under': patterns and implications of contemporary Australian burning. *International Journal of Wildland Fire* **16**, 361–377. doi:10.1071/WF07018

Schoennagel T, Nelson CR (2011) Restoration relevance of National Fire Plan treatments across the western US. *Frontiers in Ecology and the Environment* **9**, 271–277. doi:10.1890/090199

Schoennagel T, Veblen TT, Romme WH (2004) The interaction of fire, fuels and climate across Rocky Mountain forests. *Bioscience* **54**, 661–676. doi:10.1641/0006-3568(2004)054[0661:TIOFFA]2.0.CO;2

Skinner WR, Stocks BJ, Martell DL, Bonsal B, Shabbar A (1999) The association between circulation anomalies in the mid-troposphere and area burned by wildland fire in Canada. *Theoretical and Applied Climatology* **63**, 89–105. doi:10.1007/S007040050095

Stambaugh MC, Guyette RP (2008) Predicting spatio-temporal variability in fire return intervals using a topographic roughness index. *Forest Ecology and Management* **254**, 463–473. doi:10.1016/J.FORECO.2007.08.029

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

Stephens SL, Ruth LW (2005) Federal forest fire policy in the United States. *Ecological Applications* **15**, 532–542. doi:10.1890/04-0545

Stephens SL, Martin RE, Clinton NE (2007) Prehistoric fire area and emissions from California's forests, woodlands, shrublands, and grasslands. *Forest Ecology and Management* **251**, 205–216. doi:10.1016/J.FORECO.2007.06.005

Stephenson NL (1990) Climatic control of vegetation distribution: the role of the water balance. *American Naturalist* **135**, 649–670. doi:10.1086/285067

Stephenson NL (1998) Actual evapotranspiration and deficit: biologically meaningful correlates of vegetation distribution across spatial scales. *Journal of Biogeography* **25**, 855–870. doi:10.1046/J.1365-2699.1998.00233.X

Stocks BJ, Mason JA, Todd JB, Bosch EM, Wotton BM, Amiro BD, Flannigan MD, Hirsch KG, Logan KA, Martell DL, Skinner WR (2002) Large forest fires in Canada, 1959–1997. *Journal of Geophysical Research—Atmospheres* **108**(D1), FFR5-1–FFR5-12.

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

Sturtevant BR, Cleland DT (2007) Human and biophysical factors influencing modern fire disturbance in northern Wisconsin. *International Journal of Wildland Fire* **16**, 398–413. doi:10.1071/WF06023

Syphard AD, Radeloff VC, Keeley JE, Hawbaker TJ, Clayton MK, Stewart SI, Hammer RB (2007) Human influence on California fire regimes. *Ecological Applications* **17**, 1388–1402. doi:10.1890/06-1128.1

Syphard AD, Radeloff VC, Keuler NS, Taylor RS, Hawbaker TJ, Stewart SI, Clayton MK (2008) Predicting spatial patterns of fire on a southern California landscape. *International Journal of Wildland Fire* **17**, 602–613. doi:10.1071/WF07087

Theobald DM, Romme WH (2007) Expansion of the US wildland–urban interface. *Landscape and Urban Planning* **83**, 340–354. doi:10.1016/J.LANDURBPLAN.2007.06.002

US Geological Survey (2000) Global 30 Arc-Second Elevation Data Set. (EROS Data Center Distributed Active Archive Center: Sioux Falls, SD). Available at http://edcdaac.usgs.gov/gtopo30/gtopo30.html [Verified 20 June 2010]

US Geological Survey (2010) National Land Cover, Version 1. National Biological Information Infrastructure, Gap Analysis Program (GAP). Available at http://gapanalysis.usgs.gov/data/land-cover-data [Verified 20 June 2010]

Watts RD, Compton RW, McCammon JH, Rich CL, Wright SM, Owens T, Ouren DS (2007) Roadless space of the conterminous United States. *Science* **316**, 736–738. doi:10.1126/SCIENCE.1138141

Westerling AL, Hidalgo HG, Cayan DR, Swetnam TW (2006) Warming and earlier spring increases western US forest wildfire activity. *Science* **313**, 940–943. doi:10.1126/SCIENCE.1128834

Zhao MS, Heinsch FA, Nemani RR, Running SW (2005) Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sensing of Environment* **95**, 164–176. doi:10.1016/J.RSE.2004.12.011

Appendix 1. List of all explanatory variables initially considered for modelling

A fraction of these variables graduated to the final models to avoid using variables that were highly correlated. All climate variables were calculated on a monthly basis and annual averages were based on mean values of every month. Unless otherwise specified, all values were computed for the 1984 to 2008 time period. All citations for the sources are provided in the text

Category	Input name	Source	Description
Ignitions[A]	Lgt_Dens[S]	NASA Global Hydrology and Climate Centre	Annual density of lightning strikes (1995–2005) at the 1, 100, 1000 and 10 000-ha scales (strikes km^{-2} $year^{-1}$)
	Pop_Dens[S]	Gridded population of the world, version 3	Population density at the 1, 100, 1000 and 10 000-ha scales (people km^{-2})
	Rd_Dens[S]	ESRI StreetMap	Road density at the 1, 100, 1000 and 10 000-ha scales (people km^{-2})
	RdlsVol[S]	ESRI StreetMap	Roadless volume metric (remoteness) at the 1, 100, 1000 and 10 000-ha scales (km^3 $person^{-1}$)
	WUI_Dist[S]	University of Colorado	Distance to the nearest wildland–urban interface area at the 1, 100, 1000 and 10 000-ha scales (m)
Climate	Temp	PRISM	Annual temperature (°C)
	DiurTempRange	PRISM	Diurnal range in temperature (°C)
	TempRange	PRISM	Annual range in mean monthly temperature (°C)
	Isotherm	PRISM	DiurTempRange/TempRange (×100)
	TempSeas	PRISM	Temperature seasonality (coefficient of variation)
	MaxTempWarmest	PRISM	Maximum temperature of the warmest month (°C)
	MinTempColdest	PRISM	Minimum temperature of the warmest month (°C)
	MeanTempWettest	PRISM	Temperature of the wettest month (°C)
	MeanTempDriest	PRISM	Temperature of the driest month (°C)
	MeanTempWarmest	PRISM	Temperature of the warmest month (°C)
	MeanTempColdest	PRISM	Temperature of the coldest month (°C)
	GrowDegDays	Natural Resources Canada	Cumulative sum of degrees (or growing degree days) ≥5°C during the growing season (°C)
	GrowDays	Natural Resources Canada	Growing season length (number of days)
	PcpAnn	PRISM	Annual precipitation (mm)
	PcpWettest	PRISM	Precipitation of the wettest month (mm)
	PcpDriest	PRISM	Precipitation of the driest month (mm)
	PcpSeas	PRISM	Precipitation seasonality (coefficient of variation)
	PcpWarmest	PRISM	Precipitation of the warmest month (mm)
	PcpColdest	PRISM	Precipitation of the coldest month (mm)
	WatDef	PRISM	Annual water deficit (mm)
	WatDef_CV_Mon	PRISM	Coefficient of variation among monthly means of water deficit values (%)
	WatDef_CV_Ann	PRISM	Coefficient of variation among annual water deficit values (%)
	AET	PRISM	Mean annual actual evapotranspiration (mm)
	AET_CV_Mon	PRISM	Coefficient of variation among monthly means of actual evapotranspiration values (%)
	AET_CV_Ann	PRISM	Coefficient of variation among annual actual evapotranspiration values (%)
	MinSPI	PRISM	Minimum monthly standardised precipitation index
	MaxSPI	PRISM	Maximum monthly standardised precipitation index
	WindWarmest90	NOAA	90th percentile wind speed of the warmest month (m s^{-1})
	WindWarmest95	NOAA	95th percentile wind speed of the warmest month (m s^{-1})
	WindWarmest99	NOAA	99th percentile wind speed of the warmest month (m s^{-1})
	WindDriest90	NOAA	90th percentile wind speed of the driest month (m s^{-1})
	WindDriest95	NOAA	95th percentile wind speed of the driest month (m s^{-1})
	WindDriest99	NOAA	99th percentile wind speed of the driest month (m s^{-1})
Topography[A]	SurfArea_Ratio[S]	USGS and EROS	Ratio of surface to area at the 1, 100, 1000 and 10 000-ha scales
Vegetation[A]	GPP[S]	MODIS (MOD17A3)	Gross primary productivity at the 1, 100, 1000, and 10 000-ha scales (g C km^{-2} $year^{-1}$)
	Fuel_Pct[S]	USA GAP Analysis Land cover	Percentage land cover of fuels at the 1, 100, 1000 and 10 000-ha scales (%)

[A]Denotes the 'scale-dependent' variable. Values of for these variables were calculated using a moving-window approach at four spatial scales ([S]): 1, 100, 1000 and 10 000 ha. In the text, the scale-dependent variables are written as Fuel_Pct1, Fuel_Pct100, Fuel_Pct1000, and Fuel_Pct10000.