CSIRO PUBLISHING

International Journal of Wildland Fire **2021**, *30*, 307–321
https://doi.org/10.1071/WF20186

The importance of small fires for wildfire hazard in urbanised landscapes of the northeastern US

Amanda R. Carlson [A,D], *Megan E. Sebasky* [B], *Matthew P. Peters* [C] and *Volker C. Radeloff* [A]

[A]SILVIS Laboratory, Department of Forest and Wildlife Ecology, University of Wisconsin–Madison, 1630 Linden Drive, Madison, WI 53706, USA.
[B]Wisconsin Department of Natural Resources, 101 S Webster Street, Madison, WI 53707, USA.
[C]USDA Forest Service, Northern Research Station, 359 Main Road, Delaware, OH 43015, USA.
[D]Corresponding author. Email: carlson28@wisc.edu

Abstract. Frequent, small wildfires can pose dangers to homes in the wildland–urban interface, but are not often included in wildfire hazard models. We assessed patterns of small wildfire occurrence probability in the Northeast region of the United States, focusing on (1) spatial and seasonal variations; (2) differences between small and large fires (size threshold of 4 ha); and (3) how predicted probabilities are influenced by inconsistent wildfire definitions in urbanised landscapes. We analysed fire incident report data from 2005 to 2017 to parameterise maximum entropy (MaxEnt) models based on land cover, topography, climatic water deficit, soil moisture and road density. Overall, wildfire occurrence was highest in areas with lower agricultural cover and with more low-density urban development (explaining 53.5 and 28.6% of variance, respectively, in our region-wide model), while larger fires were concentrated in areas with intermediate levels of development, higher climatic water deficit and more rugged topography. These patterns were largely consistent when we assessed models for individual states, but differences in wildfire reporting patterns led to differences in the effect of urban development on fire probability. Our results provide novel understanding of small wildfire patterns in the Northeast and demonstrate the need to more reliably quantify these hazards.

Keywords: fire frequency, fire management, fire occurrence, fire regimes, maximum entropy, Mid-Atlantic, Midwest, modelling, New England, Northeast, small wildfire, wildland–urban interface.

Received 15 December 2020, accepted 2 March 2021, published online 23 March 2021

Introduction

Wildfire is an important natural disturbance process in many vegetated ecosystems, but both natural and human-caused fires can pose major hazards to human life and property (Moritz *et al.* 2014; McLauchlan *et al.* 2020). There is a growing need to make communities more fire-adapted as wildfires around the globe become more destructive owing to climate change (Abatzoglou and Williams 2016), growth of housing in the wildland–urban interface (WUI; Radeloff *et al.* 2005; Kramer *et al.* 2018; Radeloff *et al.* 2018), and increased fuel loads due to past fire suppression (Stephens *et al.* 2016). In the United States, concerns over wildfire danger focus primarily on the Western region, where the frequency and destructiveness of large fires has increased in the 21st century (Bowman *et al.* 2017). However, frequent, small fires that occur near homes and structures can also pose threats and strain suppression resources (Ager *et al.* 2019; Mietkiewicz *et al.* 2020). Hence, it is important to better understand wildfire hazard not only where there is potential for large burned areas, but where ignitions have the potential to threaten valued resources (Kolden 2020).

Large wildfires are relatively rare in regions with cool and humid climates, without extensive areas of fire-adapted vegetation, and with few lightning-caused ignitions (Malamud *et al.* 2005; Cary *et al.* 2009). In such regions, wildfire occurrence is likely to be strongly associated with human-caused ignitions near homes, roads and other types of development (Cardille and Ventura 2001; Yang *et al.* 2008; Bar Massada *et al.* 2013). Indeed, fire frequency is highest in areas with low- to intermediate-density development where sources of human ignitions intermingle with flammable vegetation (Syphard *et al.* 2007; Hawbaker *et al.* 2013). Human-caused wildfire danger in these landscapes depends more on the frequency of ignitions in close proximity to homes than on large fires spreading to homes from remote wildland areas (Balch *et al.* 2017; Mietkiewicz *et al.* 2020).

Predicting where wildfires are likely to occur is a critical component of hazard and risk assessments. The US Department of Agriculture Forest Service defines wildfire 'risk' as the potential loss of resources and assets to wildfire, and 'hazard' as potential damage caused by wildfire as a function of both the probability of occurrence and the likely intensity of burning (Scott *et al.* 2013). Quantifying hazard is required to calculate risk, in combination with susceptibility of resources and assets. Patterns of wildfire hazard are determined by patterns in flammable fuels, frequency of weather conditions favourable to fire and ignition sources (Thompson *et al.* 2011). National-scale risk assessments have estimated wildfire hazard across the USA using FSim, a simulation model that predicts fire spread potential based on fuel cover and ignition maps (Finney *et al.* 2011; Dillon *et al.* 2015). Fuel cover input maps are provided by the LANDFIRE program (Rollins 2009), which broadly categorises developed land cover types (e.g. agriculture and urban) as 'non-burnable'. FSim is therefore well suited for predicting likelihoods of large fires spreading through contiguous patches of vegetation, but has limited ability to assess wildfire hazard in urban areas. Alternative approaches are necessary to assess wildfire occurrence in regions with large numbers of human-caused ignitions in fragmented vegetation.

Explicitly modelling small fires is important because their patterns and drivers may differ from those of larger fires. For example, higher road densities may be associated with higher densities of fire ignitions, but may also reduce the potential for large fires by allowing easier firefighter access and facilitating suppression (Spyratos *et al.* 2007). Across North America, human population density is positively correlated with fire ignition frequency since 1984 but negatively correlated with occurrences of very large fires for the same time period (>400 ha; Parisien *et al.* 2016). Low- to medium-density housing development can reduce large wildfire likelihood by fragmenting wildland vegetation and limiting the potential for fire spread (Syphard *et al.* 2007). However, small wildfire occurrence can increase probabilities of large fires occurring during periods of favourable weather conditions (Nagy *et al.* 2018). Furthermore, frequent occurrences of small fires may be indicative of heavy investment in fire suppression to prevent larger fires.

Capturing patterns of small wildfire occurrence requires appropriate model types and spatial scales. FSim models for the USA predict burn probability at a quite coarse spatial resolution (270 m) owing to the computational demands of the model (Dillon *et al.* 2015). This scale of modelling matches the spatial resolution of fire records derived from satellite image analysis, such as the Monitoring Trends in Burn Severity data product (mtbs.gov), which only includes fires larger than 400 ha. Satellite sensors cannot reliably capture fires smaller than the grain size of the sensor (e.g. smaller than the 1-km resolution of MODIS (Moderate Resolution Imaging Spectroradiometer); Hawbaker *et al.* 2008), and finer-resolution sensors may have longer repeat cycles that miss fires with short burning durations (e.g. shorter than 16-day repeat cycle of Landsat; Hawbaker *et al.* 2020). Alternatively, fire occurrences can be modelled as point processes using species distribution models. Point locations of fires can be determined using fire incident reports, which provide a more complete record of small fire occurrences than satellite-based records. Species distribution models can be used to estimate long-term spatial drivers of ignitions (e.g. Syphard *et al.* 2008; Bar Massada *et al.* 2013; Parisien *et al.* 2016) or influences of spatial drivers and weather on spatio-temporal patterns of fire occurrence (e.g. Miranda *et al.* 2012; Peters *et al.* 2013) and extrapolate these predictions into the future. This type of model may be most appropriate where fires are typically small and single point locations can be representative of an entire burned area.

A key limitation for modelling small wildfire hazards is the difficulty of collecting complete and consistent occurrence records at regional or national scales. Compiling records across a multitude of federal, state and municipal fire response agencies presents logistical challenges, while additional challenges may arise from varying definitions of the types of wildland fires of interest for hazard assessments. In the USA, the most complete record of fires of all sizes is the fire occurrence database maintained by the Forest Service (Short 2014, 2018). However, the database sometimes misses incident records for fires fought by municipal fire departments, which are often smaller fires in close proximity to developed areas, such as in municipal parks, drainage ditches, alleyways, construction sites, vacant lots, or roadsides (Karen Short, pers. comm.). These types of wildfires are not typically considered in national-scale wildfire risk assessments, because larger fires can be more reliably mapped with satellite imagery and pose major threats to communities (Finney *et al.* 2011; Dillon *et al.* 2015). However, fires originating in highly human-modified landscapes also represent an ecologically important interactions between humans and the environment, although they may be ambiguously defined as 'wildland fires' (Pyne 2019).

Our goal was to assess patterns and drivers of small wildfire occurrence across the Northeast region of the USA. Small wildfires are an important contributor to overall wildfire hazard in the Northeast owing to the limited occurrence of large fires, the prevalence of small, human-caused ignitions, and high exposure of homes in the WUI (Malamud *et al.* 2005; Martinuzzi *et al.* 2015; Balch *et al.* 2017). The Northeast accounts for only ~1.5% of area burned in the USA but ~14% of wildfire incidents (NIFC 2020), and the annual number of homes threatened or destroyed by wildfire is similar to the West despite burned areas being substantially smaller (St Denis *et al.* 2020). Specifically, we asked:

1. How do overall and seasonal wildfire occurrence probabilities vary across the Northeast in relation to human and biophysical predictors?
2. How do patterns and drivers of small fire occurrences differ from those of larger fires?
3. How might inconsistencies in wildfire definitions and reporting standards among the 20 states in the region influence region-wide assessments of fire occurrence probabilities, and subsequent hazard and risk assessments?

Methods

Study area

Our study area included the 20 states in the Northeast region defined by the National Cohesive Wildland Fire Management Strategy (Fig. 1*a*). We grouped the region into four sub-regions in order to better categorise state-by-state patterns across diverse fuel types, climates and urbanisation patterns. Sub-regions included: New England (including states of Connecticut, Massachusetts, Maine, New Hampshire, Rhode Island and Vermont), the Mid-Atlantic (Delaware, Maryland, New Jersey, New York, Pennsylvania and West Virginia), the Lower Midwest (Illinois, Indiana, Iowa, Missouri and Ohio) and the Upper Midwest (Michigan, Minnesota and Wisconsin).

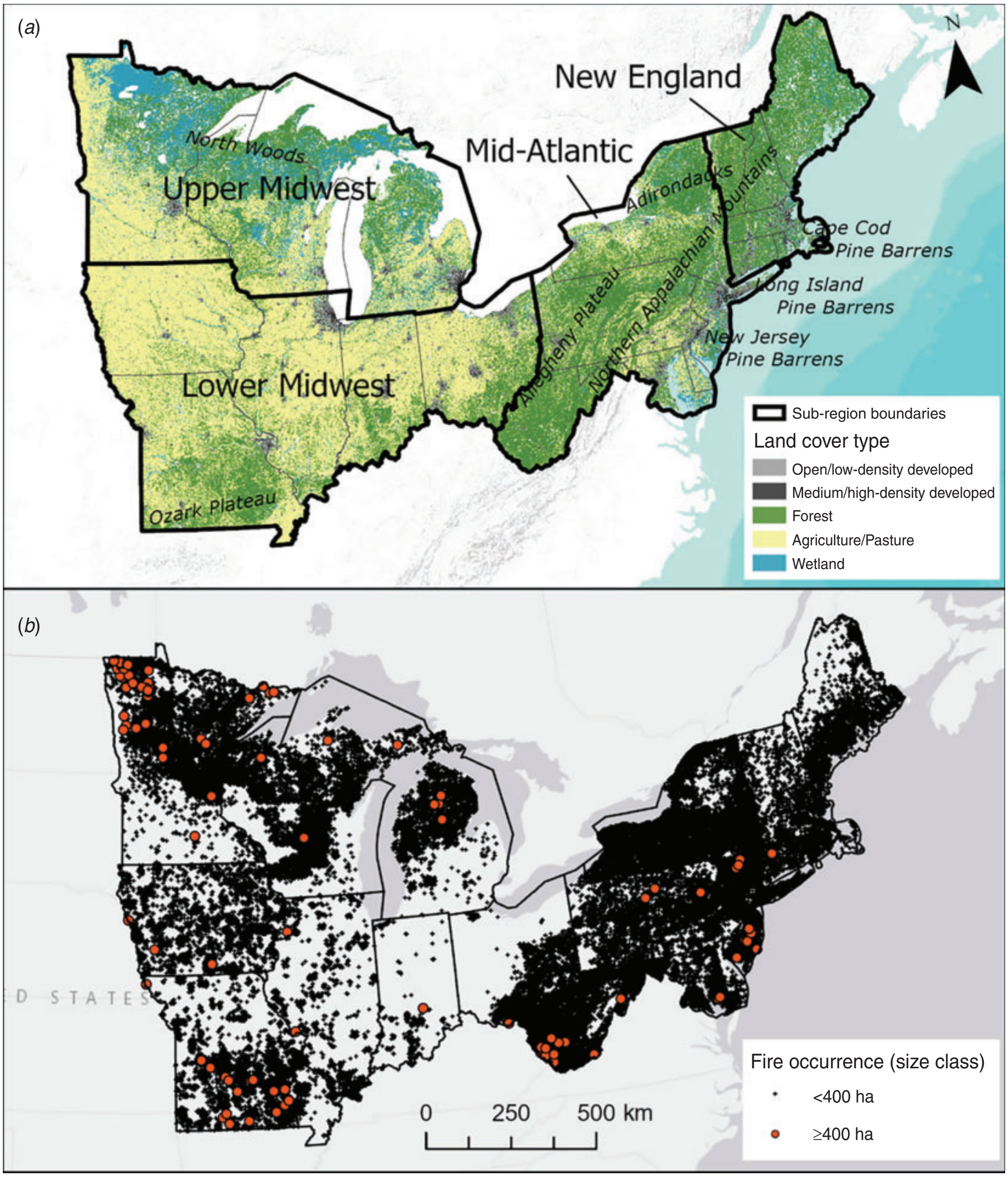

Fig. 1. Maps of the Northeast region showing (*a*) boundaries of sub-regions used to group results according to broad similarities in climate and vegetation, along with land cover types from the 2016 National Land Cover Dataset (mrlc.gov) and selected geographic features; and (*b*) locations of all fire occurrences from 2005 to 2017 for fires >0.1 ha from the USFS fire occurrence database. Fires ≥400 ha are highlighted to compare patterns of fires large enough to be included in the Landsat-based Monitoring Trends in Burn Severity dataset (mtbs.gov).

The Northeast is broadly characterised by a humid temperate climate with a wide range of ecosystem types, including sub-boreal conifer forest, tallgrass prairie, temperate deciduous hardwood and mixed forest, Ozark and Appalachian mountain forests, oak savanna prairie and Atlantic coastal plains (Omernik 1987). Historically, fires burned probably almost annually in the prairie ecosystems around the southern Great Lakes (Upper and Lower Midwest sub-regions) and in the marshlands along the Mid-Atlantic coast (Grimm 1984; Leitner *et al.* 1991; Frost 1995). Similarly, dry upland pine and oak forests of the Great Lakes and Atlantic coastal regions burned probably at least once per decade in the pre-European era. However, fire return intervals were much longer in the sub-boreal hardwood forests of the Upper Midwest, the Adirondacks and northern New England, burning only every few decades (Little 1974; Guyette *et al.* 2005). Native Americans and early European settlers used fire extensively to manage forests and prairies, but a policy of fire suppression enacted by the US Forest Service (USFS) in the early 20th century led to the widespread decline of fire-adapted oak and pine forest and savannas (Whitney 1987; Nowacki and

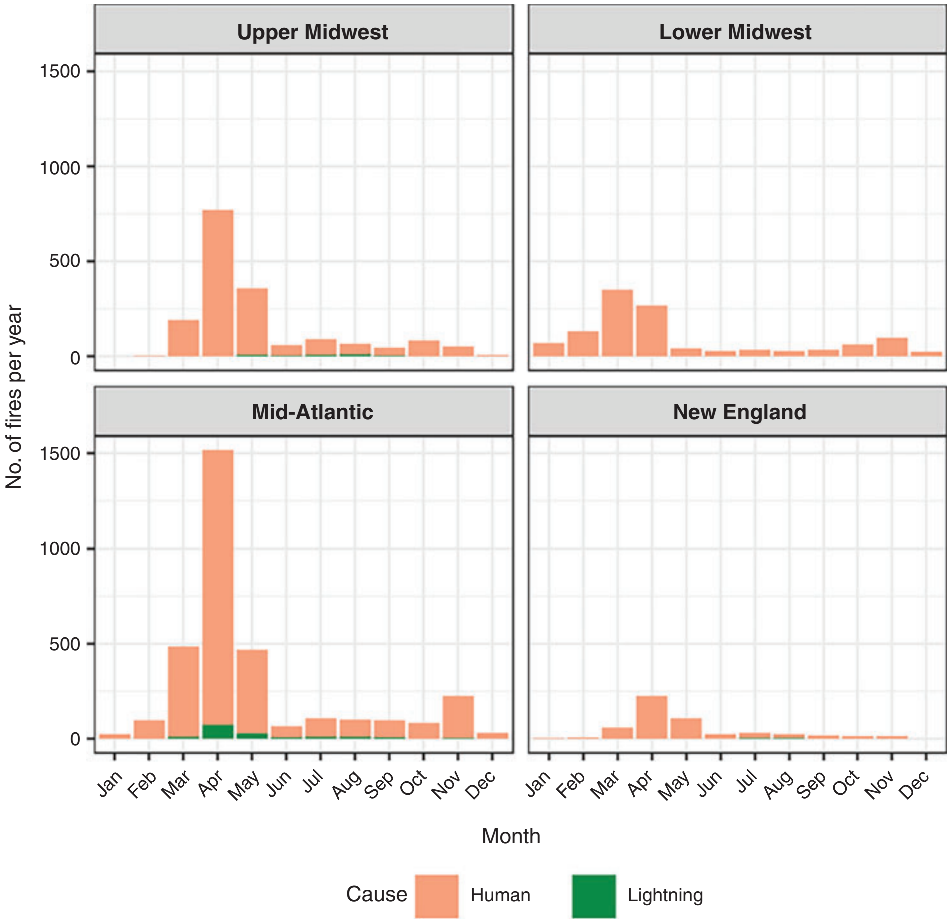

Fig. 2. Monthly patterns of fires reported in the US Forest Service fire occurrence database, summarised for four sub-regions of the Northeast region. Averages are calculated for the period 2005–2017. Records include all fires with a final size >0.1 ha.

Abrams 2008). Furthermore, wildland vegetation has been highly fragmented throughout the region owing to widespread conversion to agriculture or urban land cover (Motzkin *et al.* 1999; Rhemtulla *et al.* 2007). Loss of native vegetation to agriculture is most extensive in the Lower Midwest. The most extensive WUI areas are in the Mid-Atlantic and in New England, where extensive urbanisation intermingles with forest cover (Fig. 1*a*).

Data

We assessed fire occurrences using point locations from the USFS fire occurrence database (Short 2018). Records within the database have been subjected to quality assurance, which includes removal of duplicate fires and correction or removal of records with obvious location errors (Short 2014). However, there are clear differences in reporting rates among states (Fig. 1*b*), and we focused our analysis on individual states to account for this. The database provides point locations (latitude and longitude), firefighting agency, fire cause, final fire size, and dates of discovery and control. We examined annual numbers of wildfires reported by each state to determine year ranges for which reporting appeared to be consistent (i.e. where the number of fires for a range of consecutive years did not have any abrupt increases that were likely due to a shift in reporting methods). We determined that reporting rates were consistent for all 20 states over the period 2005–2017, and we used records from these years for all analyses. We excluded fires <0.1 ha in order to ensure that occurrence records did not contain, for example, illegal campfires or unpermitted debris burns that did not escape control. For this range of fire years and sizes, the majority of occurrences in all four sub-regions were in spring (Mar–May; 72% of all fires), were human-caused (95% of all fires with known causes; Fig. 2), and were <4 ha (Table 1).

We obtained a WUI classification from the University of Wisconsin–Madison SILVIS laboratory (http://silvis.forest.wisc.edu/data/wui-change/; Martinuzzi *et al.* 2015). The WUI dataset is based on block-level housing density from the 2010 census for the conterminous USA. WUI areas are separated into two distinct classes: (1) the intermix, where housing intermingles with wildland vegetation; and (2) the interface, where housing is adjacent to wildland vegetation (Radeloff *et al.* 2005; Stewart *et al.* 2007).

We obtained 30-m-resolution land cover layers from the 2016 National Land Cover Dataset (NLCD; Yang *et al.* 2018). Although LANDFIRE distinguishes fuel types in greater detail, the NLCD distinguishes between open, low-, medium- and

Table 1. Fire occurrence records by state summarised by the total number of records from 2005 to 2017 and percentages of fires in each size class

Some rows may not sum to 100% owing to rounding. States are grouped by sub-regions. State abbreviations are as follows: CT, Connecticut; DE, Delaware; IA, Iowa; IL, Illinois; IN, Indiana; MA, Massachusetts; MD, Maryland; ME, Maine; MI, Michigan; MN, Minnesota; MO, Missouri; NH, New Hampshire; NJ, New Jersey; NY, New York; OH, Ohio; PA, Pennsylvania; RI, Rhode Island; VT, Vermont; WI, Wisconsin; WV, West Virginia

Sub-region	State	No. occurrences	<1 ha	1–4 ha	4–40 ha	>40 ha
Upper Midwest	MI	3946	29.2%	43.9%	23.9%	3.0%
	MN	13 169	34.8%	37.3%	20.3%	7.6%
	WI	5534	47.4%	33.9%	16.8%	1.9%
Lower Midwest	IA	4021	12.8%	40.7%	36.4%	10.1%
	IL	1533	19.2%	45.7%	28.9%	6.2%
	IN	489	21.3%	45.0%	26.0%	7.8%
	MO	6527	6.9%	48.6%	29.9%	14.6%
	OH	2844	18.6%	54.2%	24.4%	2.8%
Mid-Atlantic	DE	60	21.7%	48.3%	28.3%	1.7%
	MD	1864	41.1%	35.5%	17.8%	5.6%
	NJ	3523	49.1%	35.1%	12.8%	3.0%
	NY	25 095	36.7%	56.1%	6.9%	0.3%
	PA	5392	26.7%	47.7%	22.9%	2.7%
	WV	6946	23.7%	33.8%	28.4%	14.0%
New England	CT	1694	42.7%	45.9%	10.4%	0.9%
	MA	1305	35.9%	50.8%	12.0%	1.2%
	ME	2434	45.4%	41.2%	12.6%	0.8%
	NH	726	39.3%	45.7%	13.5%	1.5%
	RI	313	31.6%	57.2%	11.2%	0.0%
	VT	344	24.1%	53.2%	20.9%	1.7%

high-density development classes and offers a more high-level classification that we favoured for easier interpretation of our model results. We combined similar NLCD classes to create the following simplified classes: high/medium intensity developed, low/open intensity developed, agriculture, grassland/herbaceous, shrub/scrub, evergreen forest, mixed forest, deciduous forest and wetlands. Reported fire location coordinates did not always match the exact location of the fire and were often biased towards adjacent developed areas (e.g. highways, empty lots, which may have been locations of firefighter staging areas rather than the actual burned site; Karen Short, pers. comm.). To account for spatial errors, we calculated the percentage cover of each class within 120 m of the point location to associate fire locations with vegetation within a neighbouring area and reduce bias towards developed land cover types.

We included additional explanatory variables to account for soil moisture, topography, road density and climate (Table S1, Supplementary material). We obtained 30-m-resolution soil available water supply at 150-cm depth from the gridded Soil Survey Geographic dataset (gSSURGO; Soil Survey Staff 2020). We obtained elevation from the National Elevation Dataset (Gesch *et al.* 2018) and derived a terrain roughness index (Riley *et al.* 1999). As a proxy for human accessibility and potential for ignitions, we calculated road density using the National Transportation Dataset (US Geological Survey 2017). We assessed climatic drivers of fire occurrence based on monthly climate water deficit data (actual minus potential evapotranspiration) from the TerraClim climate dataset (Abatzoglou *et al.* 2018), which we averaged over all months within the study period 2005–2017. We also averaged seasonal water deficit for spring (Mar–May), summer (Jun–Aug) and fall (autumn; Sep–Nov).

WUI area comparisons

We used WUI areas to summarise broad patterns of wildfire occurrence in relation to human development and wildland vegetation. For each sub-region, we tallied the percentage of total area in each WUI class (intermix and interface) and the percentage of fires occurring in each WUI class. We additionally calculated these percentages for a subset of 'large' fires, which we defined as fires with a final size >4 ha (10 acres). Large fires made up 11% of total fire occurrences in the region (see Table 1 for state-by-state percentages).

Fire occurrence models

We modelled fire occurrence probabilities with maximum entropy models (MaxEnt). We chose MaxEnt because it is optimised to handle occurrence records as 'presence-only' data (Phillips *et al.* 2006), which is important because the absence of a fire record in a given location is unlikely to indicate that it would be impossible for a fire to occur there. Because our fire occurrence dataset spans only 13 years for which there are reliable data, and because locations in the database may not exactly match the exact ignition location, we assumed that locations in the study area with no recorded fire occurrences do not represent 'true absences' of fire.

All models were computed in *R* (R Core Team 2019) with the 'dismo' package (Hijmans 2017). We parameterised models to randomly sample 10 000 background points (removing duplicates) and used all available covariate features, excluding 'discrete'. The regularisation multiplier, which imposes a penalty on coefficients to prevent over-fitting, was set to the default value of 1.0. We fitted models using random training subsets (75%) of relevant occurrence records, and validated them with the remaining 25% of records. We fitted models for each state in the Northeast region using spatial subsets of fire occurrence records and explanatory data layers. We checked for collinearity among explanatory variables by calculating Pearson's correlations (ρ) using values extracted from the full set of fire occurrence locations. We included all variables in the models after determining that $|\rho| < 0.7$ for all variable pairs. For each state, we fitted one model for all fires and a second model using large fires only (>4 ha). We did not run large-fire models for four states where there were ≤15 validation occurrences, or ≤60 total occurrences (Delaware, New Hampshire, Rhode Island and Vermont; Table 1).

In addition to models for individual states, we fitted models for the entire Northeast Region using all occurrence points within the 20-state boundary. Comparing region-wide fire occurrence probabilities with probabilities mapped across individual states allowed us to identify states where fire occurrence patterns differed from the overall regional patterns. As with our state models, we fitted region-wide models for all fires (n = 87 759) and for large fires only (n = 20 075). We additionally fitted seasonal region-wide models based on

Table 2. Percentages of area in each sub-region classified as wildland–urban intermix or interface and percentage of fires from 2005 to 2017 that occurred in each class

Region	% Intermix area	% Intermix fires – all	% Intermix fires >4 ha	% Interface area	% Interface fires – all	% Interface fires >4 ha
Upper Midwest	6.7	30.1	10.1	1.6	5.0	1.1
Lower Midwest	5.9	16.1	7.8	1.7	4.9	2.0
Mid-Atlantic	23.1	33.2	29.7	6.8	14.3	3.9
New England	27.2	47.6	30.8	3.9	15.3	11.4

temporal subsets of fire occurrences in spring, summer and fall months with the corresponding seasonal averages of climatic water deficit.

For each state and region-wide model, we generated predictive maps of relative probabilities of fire occurrence. Relative probability values were based on probability distributions for occurrences and background points fitted by the MaxEnt algorithm, which are scaled from 0 to 1 (Elith *et al.* 2011). Because relative probabilities are scaled to the individual area defined by the model, outputs from separate models cannot be directly compared. Probabilities for individual states therefore cannot be compared with values in other state-wide models or region-wide models, and we only used our model results to compare relationships of fire occurrences with explanatory variables between states.

Model validation

We evaluated model performance using the area under the curve (AUC) of the receiver-operating characteristic, as well as omission rates based on minimum and 10th percentile threshold values. Both metrics were calculated using the withheld validation subsets. AUC is a threshold-independent validation metric based on plots of accurate occurrence predictions against false positive predictions, with values between 0 and 1. A score of 0.5 indicates that model performance is equivalent to that of a random prediction, while a score of 1 would indicate an exact match between predictions and validation data. We additionally calculated model omission rates by defining a suitability threshold for binary predictions using both the minimum and the 10th percentile value of suitability scores at training locations. Omission rates can highlight model inaccuracies resulting from overfitting that the AUC misses (Radosavljevic and Anderson 2014). The minimum value of training suitability scores is an intuitive threshold for classifying binary predictions, while the 10th percentile value is a more conservative threshold designed to exclude spurious observations.

We further evaluated region-wide model performance by dividing predicted fire occurrence probability values into 'low', 'medium' and 'high' categories. Category thresholds were separately defined for each model using equal-interval division of the range of predicted probabilities (resulting in category thresholds of approximately 0.33 and 0.67). We quantified model predictive performance by determining percentages of all occurrence locations (training and validation sets combined) in each probability category. We then tabulated percentages of each category across sub-regions in order to characterise wildfire hazards geographically for the entire study region.

Results

WUI area comparisons

In all sub-regions, a disproportionately high number of fires occurred in the WUI relative to proportions of total WUI area. The discrepancies were larger when considering fires of all sizes *v.* large fires only. In the Upper Midwest and the Mid-Atlantic, large fires occurred less frequently in the interface than expected given the proportion of total interface area (Table 2). For fires of all sizes, the proportion of occurrences in the intermix WUI was more than four times higher than expected given the proportion of total intermix WUI area in the Upper Midwest, nearly three times higher in the Lower Midwest and nearly twice as high in New England (Table 2). With the exception of New England, fire occurrences in the interface WUI in other sub-regions did not exceed the proportion of total interface WUI area as much as in the intermix.

MaxEnt model performance

Model performances for individual states were fair to good, based on AUC values and omission rates (Table 3). AUCs for state models predicting occurrence of all fires (>0.1 ha) ranged from 0.64 to 0.88, and from 0.62 to 0.89 for models of large fires only (>4 ha). On a state-to-state basis, there was little difference in AUC between models predicting all fires *v.* models predicting large fires only (Table 3). However, omission rates were higher for large-fire models than for all-fire models. High omission rates for two state models (the Delaware all-fires model and Massachusetts large-fires model) were likely due to a limited number of validation records (n = 15 and 16, respectively), either due to the small areas of these states or due to low rates of fire reporting. However, omission rates were also high for large-fire models in some states where the number of observations was not limiting (e.g. Maryland and Ohio). States in the Upper and Lower Midwest generally had higher AUCs than states in the Mid-Atlantic and New England (Table 3).

The region-wide models performed well for all fires, large fires only and seasonal fires based on AUCs and omission rates (Table 4). The all-fires model reliably predicted fire occurrences, with 83.8% of all reported fire locations occurring in pixels classified as having high probability and only 2.0% of fire locations occurring in areas classified as low probability. However, the large-fires model was not as effective at discriminating fire occurrence probabilities, with 51.9% of reported large-fire locations falling in high-probability areas and 45.3% in medium-probability areas. Spring seasonal model performance was similar to the all-fires model, with 88.7% of fire locations occurring in high-probability areas and only 2.1% occurring in

Table 3. Test AUC scores and omission rates for state models predicting occurrences of all fires and for models predicting large fires only (>4 ha), by state
Missing values indicate states where there were fewer than 50 occurrence records for large fires. Omission rates are based on suitability thresholds derived from both the minimum and 10th percentile of training occurrences. State abbreviations are described in the heading for Table 1

Sub-region	State	All fires AUC	Omission rate (min.)	Omission rate (10th percentile)	Fires >4 ha AUC	Omission rate (min.)	Omission rate (10th percentile)
New England	CT	0.761	0.00%	10.61%	0.746	0.00%	16.67%
	MA	0.725	0.31%	13.80%	0.569	5.88%	58.82%
	ME	0.861	0.17%	11.38%	0.797	7.69%	15.39%
	NH	0.751	0.55%	10.44%	-	-	-
	RI	0.701	0.00%	12.66%	-	-	-
	VT	0.774	0.00%	15.12%	-	-	-
Mid-Atlantic	DE	0.793	13.33%	13.33%	-	-	-
	MD	0.675	0.22%	10.11%	0.764	5.17%	29.31%
	NJ	0.695	0.23%	10.90%	0.729	3.13%	17.19%
	NY	0.865	0.00%	14.55%	0.756	1.56%	15.63%
	PA	0.736	0.00%	11.57%	0.721	2.47%	13.58%
	WV	0.755	0.00%	9.44%	0.802	0.00%	10.38%
Lower Midwest	IA	0.867	0.00%	10.77%	0.845	0.00%	10.67%
	IL	0.82	0.00%	12.57%	0.771	1.35%	14.87%
	IN	0.886	0.00%	12.20%	0.911	4.00%	4.00%
	MO	0.811	0.00%	10.17%	0.792	0.00%	11.49%
	OH	0.876	0.00%	9.56%	0.848	2.27%	20.46%
Upper Midwest	MI	0.813	0.10%	11.86%	0.758	0.86%	15.52%
	MN	0.853	0.00%	8.41%	0.800	0.00%	8.79%
	WI	0.76	0.00%	11.49%	0.709	0.86%	16.38%

Table 4. Test AUC scores and omission rates for region-wide models predicting occurrences of all fires and for models predicting large fires only (>4 ha), for all fire records and for fires occurring in spring, summer and fall months
Omission rates are based on suitability thresholds derived from both the minimum and 10th percentile of training occurrences

Model	AUC	Omission rate (min.)	Omission rate (10th percentile)
All fires	0.801	0.00%	9.99%
Large (>4 ha) fires	0.781	0.00%	10.32%
Spring fires	0.813	0.00%	10.06%
Summer fires	0.810	0.00%	10.35%
Fall fires	0.786	0.00%	10.95%

low-probability areas. The majority of summer and fall fire locations fell in high or medium-probability areas, but were not predicted as accurately as spring fires (48.6% medium and 42.3% high for summer fires; 49.9% medium and 43.3% high for fall fires).

State models

Environmental drivers of wildfire occurrence varied considerably among states (Fig. 3*a*). Nevertheless, the percentage of low/open developed land cover within 120 m of the fire location explained the largest proportion of the observed patterns for almost all states (mean contribution 33%). Climatic water deficit was also an important variable for most states (mean contribution 14%), while agricultural land cover was especially important in the Lower Midwest states (mean contribution for sub-region 27%). Differences in explanatory variables between states were also reflected in the mapped probabilities of occurrence (Fig. 4*a*).

Patterns of large wildfire occurrence differed greatly from those of all fires and relative importance of explanatory variables was highly variable among all state-level models (Fig. 3*b*). Agricultural land cover was an important explanatory variable in the lower Midwest states, while low/open developed land cover was a less important variable overall than in all-fires models (Fig. 3). Large fire occurrence tended to be more negatively correlated with both open/low- and medium/high-density developed land cover than all fires (see Supplementary material). There was a particularly stark difference between occurrence probabilities for small and large fires in the state of New York, where small fires were highly clustered around cities and large fires were more restricted to eastern Long Island and the Hudson Valley (Fig. 4). Differences between small and large fire occurrences were also pronounced in the New Jersey Pine Barrens, coastal Maryland, and the Appalachians in Ohio and West Virginia.

Region-wide models

Similarly to the individual state models, region-wide models resulted in substantial differences in occurrence patterns between fires of all sizes and larger fires (Fig. 5). Highest occurrence probabilities of large fires were in forested areas in the Upper Midwest (i.e. the North Woods region), the Ozark Plateau, the Appalachian Mountain regions of eastern Ohio, West Virginia and Pennsylvania, and along the southern Mid-Atlantic coast. For fires of all sizes, occurrence probabilities

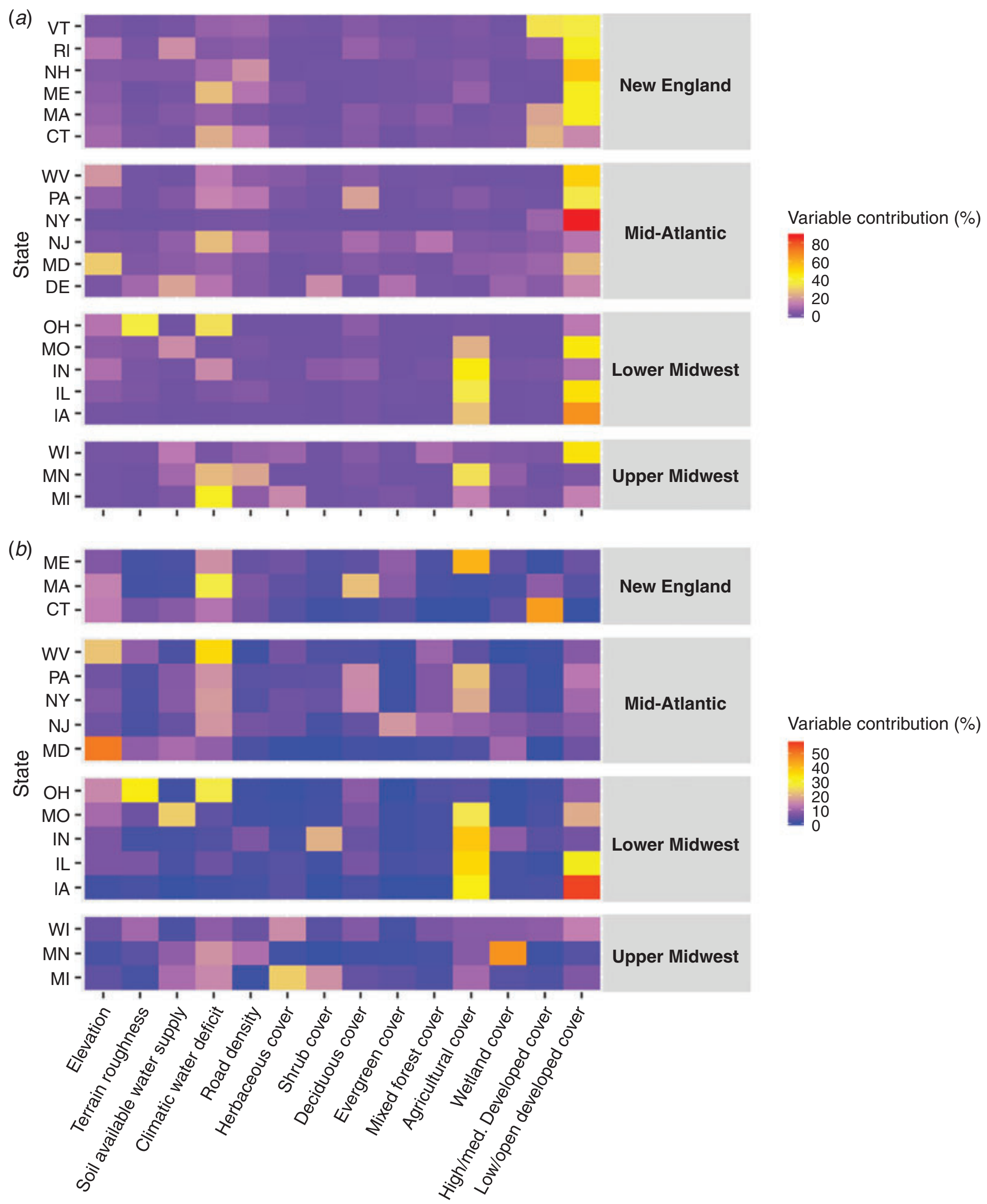

Fig. 3. Variable contributions to MaxEnt models fit to individual states for (*a*) all fires; and (*b*) fires >4 ha.

were highest near cities. Occurrence probabilities for all fires and for large fires were generally low in more sparsely populated areas (i.e. the northernmost Appalachian Mountains and Adirondacks) and in the agricultural portions of the Midwest. Across the region, the highest occurrence probabilities for fires of all sizes were predicted in the Mid-Atlantic and New England sub-regions, while the highest probabilities of larger fire occurrence were predicted mainly in the Mid-Atlantic and Upper Midwest (Fig. 6*a*).

Geographical patterns of relative fire occurrence probability varied substantially among seasons. High occurrence probability was most prevalent in the spring for all sub-regions, while the Mid-Atlantic and New England had higher summer and fall probabilities than the Upper and Lower Midwest (Fig. 6*b*). Spring fire occurrence probabilities were higher throughout the region than those for summer and fall, which were more spatially clustered (Fig. 7). Summer and fall fire occurrence probabilities were highest along the Atlantic coast, around cities in the Midwest, and in some forested portions of the Upper Midwest, the Ozarks and Appalachians, and low throughout most of the Lower Midwest and New England.

The most important variables predicting occurrence of all fires were agricultural land cover (contributing 53.5% of the model's explanatory power), low/open developed land cover (28.6%), climatic deficit (6.7%), road density (5.1%), soil available water supply (3.7%) and elevation (1.2%). Response

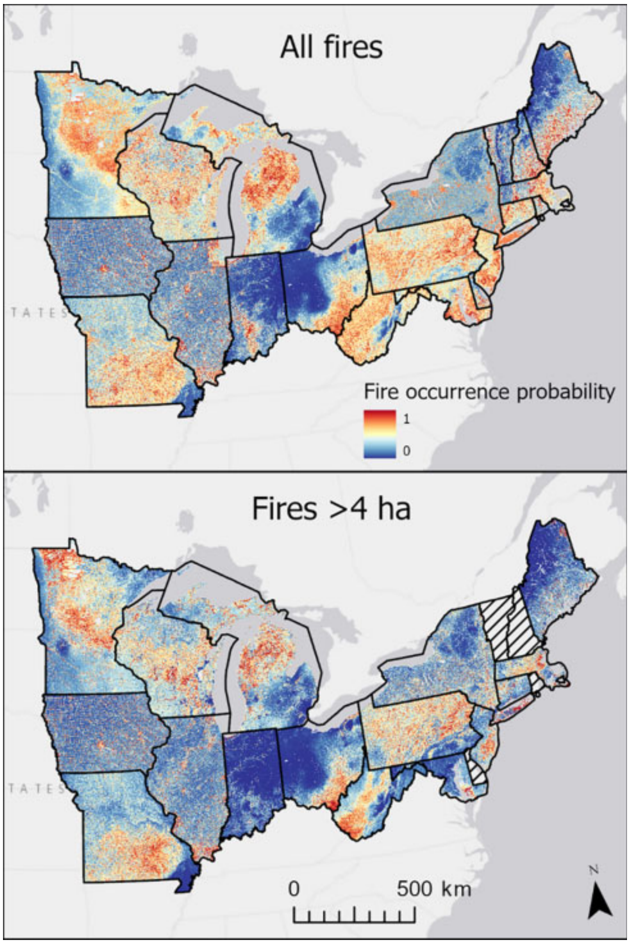

Fig. 4. Map of MaxEnt model predictions of relative wildfire occurrence probability for individual state models. Probabilities are scaled from 0 to 1 for each state (i.e. relative probabilities for different states may correspond with different absolute probabilities). Hashed symbols indicate states where models could not be fitted owing to a low number of fire locations.

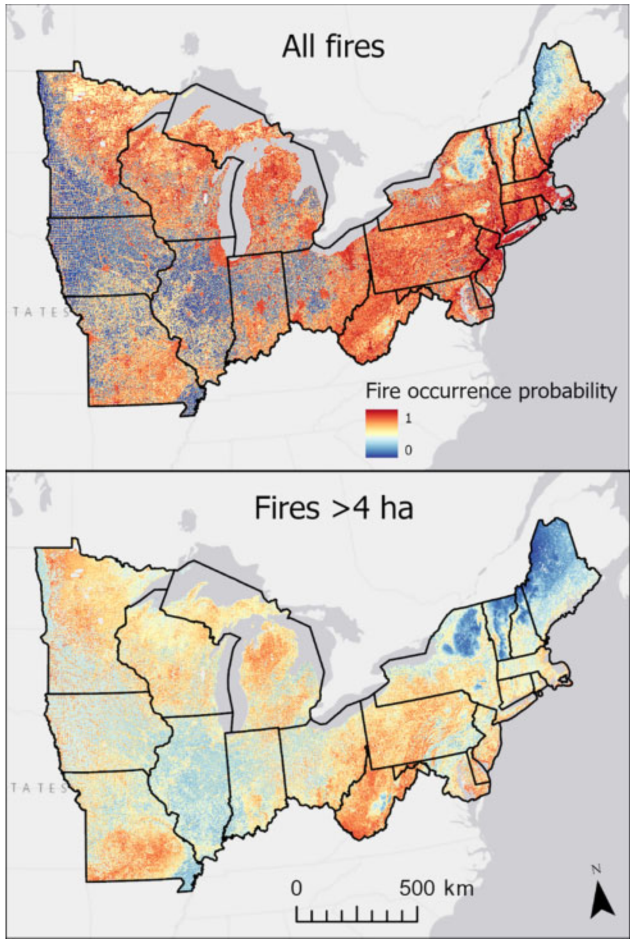

Fig. 5. Map of MaxEnt model predictions of relative wildfire occurrence probability for models fitted to the Northeast region.

curves showed that higher occurrence probabilities were associated with higher proportions of low/open developed land cover and lower proportions of agricultural land cover, as well as higher climatic deficits, higher road densities and soils with low available water supply (Fig. 8*a*). The most important variables predicting large fire occurrences were agricultural land cover (contributing 35.6% of the model's explanatory power), climatic water deficit (14.9%), elevation (11.5%), terrain roughness (10.8%), mixed forest cover (8.7%) and low/open density development (8.1%). Occurrence probability decreased with increasing agricultural land cover and mixed forest cover, increased with increasing climatic deficit and terrain roughness, and was highest at intermediate levels of low/open developed land cover across the region (Fig. 8*b*).

Discussion

Our models identified clear patterns in small wildfire occurrence throughout the Northeast. Wildland fire activity was strongly influenced by human development, with wildfires occurring more frequently in WUI than in non-WUI areas and near roads and cities in both region-wide and individual state MaxEnt models. Agricultural land cover and low-density urban development were the most important predictors of fire occurrences in almost all state models and in region-wide models, with most fires occurring in non-agricultural areas with higher proportions of low-density urban land cover. However, while the occurrence probabilities for fires of all sizes increased with increasing proportions of low/open developed land cover, probability of larger fires peaked at intermediate levels of developed cover (Fig. 8). Furthermore, models predicting occurrences of large fires only resulted in substantially different patterns of wildfire hazard and had different explanatory variable contributions than models of fires of all sizes. These differences highlight the importance of considering small wildfires in regional and national-scale hazard and risk assessments.

Our results indicate that wildfire occurrence in the Northeast is most strongly driven by human ignitions, similarly to what others have found (Cardille and Ventura 2001; Yang *et al.* 2008; Miranda *et al.* 2012). In addition to the WUI being where structures face the greatest risk from wildfires (Kramer *et al.* 2018), the presence of human development in areas with flammable vegetation also increases the frequency of wildfire ignitions (Mietkiewicz *et al.* 2020). In the Northeast, higher

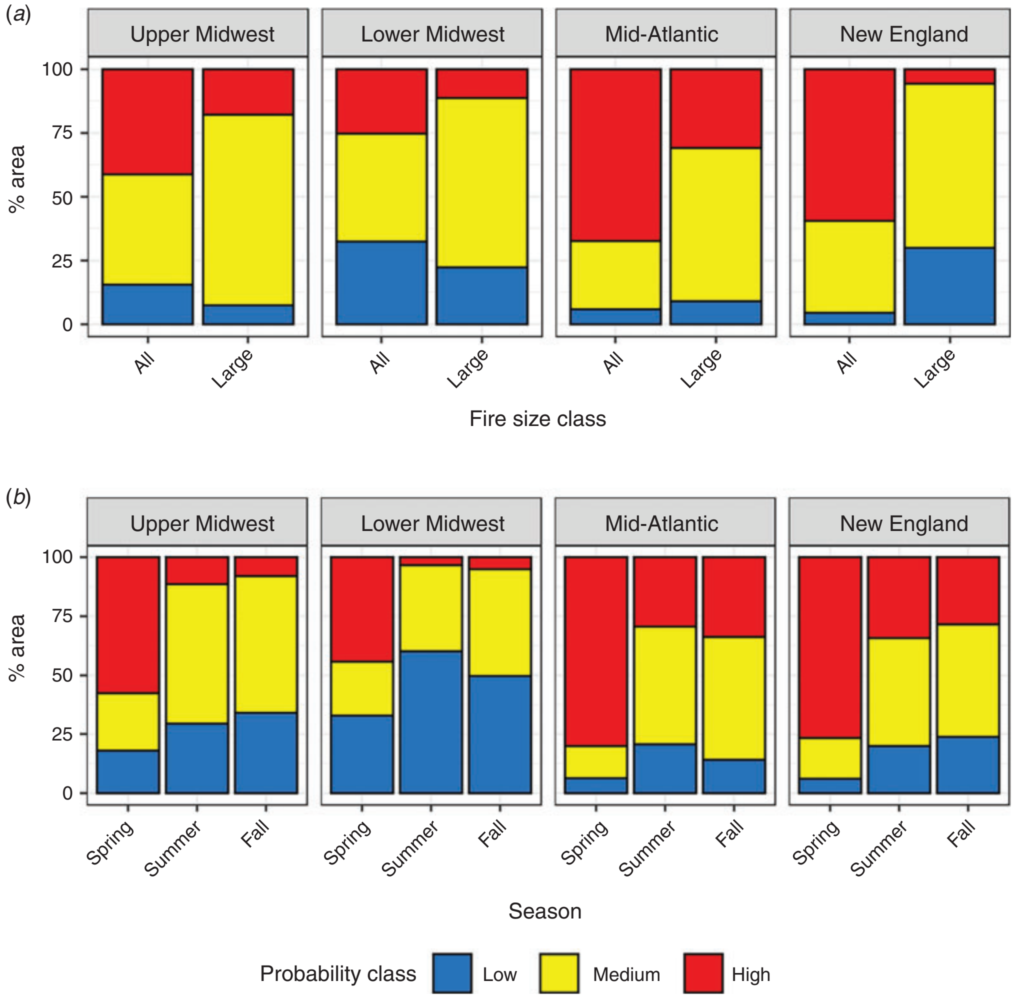

Fig. 6. Percentages of area classified as low, medium, and high occurrence probability for region-wide models predicting occurrences of (*a*) all fires and large fires only for all seasons; and (*b*) fires of all sizes by season, grouped by sub-region.

proportions of fire in the intermix *v.* interface WUI may indicate that increased ignitions result from the intermingling of human development with wildland vegetation, but that the proximity to larger patches of vegetation in the interface does not increase fire occurrence to as great an extent. This result is consistent with prior findings that wildfire occurrence is typically highest in areas with low to intermediate development densities, because there is a combination of frequent human ignitions and flammable vegetation (Syphard *et al.* 2009, 2013). Our models showed that fire was less strongly associated with high/medium-density development, and we assume that this is due to lower availability of flammable vegetation in these areas. However, the result may also be influenced by underreporting of occurrences based on inconsistency in defining what constitutes a 'wildland fire' in highly developed areas.

The influence of climatic water deficit on large fire occurrence in our models indicates that climate change will likely have an effect on future wildfire activity in the Northeast. Warming temperatures and decreased precipitation have a clear effect on wildfire activity in the West (Abatzoglou and Williams 2016) and similar climate changes would likely produce similar effects in the Northeast. While it is unclear whether precipitation will increase or decrease in the eastern USA over the 21st century, warming temperatures may increase fire activity by increasing atmospheric evaporative demand and reducing fuel moisture (Liu *et al.* 2010) and by causing more extreme short-term weather conditions (Jain *et al.* 2017). Our models demonstrated that fire occurrence increases with increasing climatic water deficits at the regional scale, although responses were varied in individual state models (see Supplementary material). This effect was particularly strong for summer and fall fires, during months when climatic water deficits are highest. Fire hazard in the Northeast is greatest in the spring months when snowmelt exposes dry fuels before summer green-up (Haines *et al.* 1983), but our models indicate that summer and fall drought, which may become more frequent and severe in the future owing to climate change, have the potential to shift fire hazard later in the year.

Our large-fire models resulted in substantially different patterns of occurrence than the all-fire models, indicating that

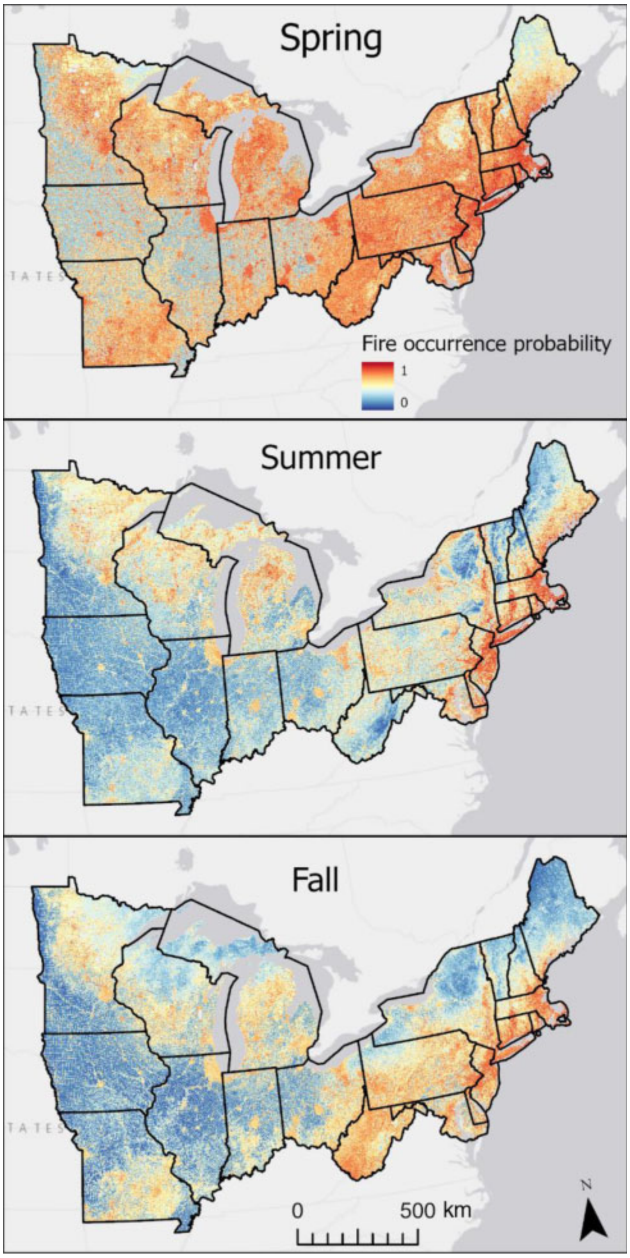

Fig. 7. Maps of MaxEnt model predictions of seasonal wildfire occurrence probabilities for models fitted to the entire Northeast region for all wildfire sizes.

there are different factors determining wildfire ignitions *v.* the probability of fires growing to a large size. Small fires may be more likely to occur near urban areas because fuels are more fragmented in these areas, but also because road accessibility allows firefighters to respond to events more quickly (Plucinski 2019). Although large wildfires pose greater threats to structures, frequent small fire ignitions contribute to suppression costs and can strain firefighting resources. The occurrence of large fires is also strongly correlated with the frequency of human-caused ignitions in the eastern USA (Nagy *et al.* 2018). Models of small fire ignitions can therefore be a valuable complement to simulation-based models that predict additional components of large-fire hazard, such as the expected rate of spread or intensity of burning (Scott *et al.* 2013).

Our fire occurrence probability maps highlighted the considerable differences in variable importance among models for different states (Figs 3 and 4). In a few states (Illinois, Iowa, Massachusetts, New York and Rhode Island), wildfire occurrence was tightly clustered around urban population centres, while some other states (e.g. Michigan, New Jersey, Ohio, Pennsylvania and Wisconsin) showed the opposite pattern. Some of these differences corresponded to differences in patterns of reported wildfire locations. For example, density of wildfire records in New York was noticeably higher than in adjacent states, and occurrence probabilities for New York were noticeably higher around cities (Fig. S1, Supplementary material). Model results for New Jersey did not show the same clustered pattern in occurrence probability, and there was not a high density of fire occurrence records in the highly urbanised Northeast portion of the state. Differences between the region-wide model and models for individual states were greatest for the Mid-Atlantic and southern New England states, where state models showed stark differences in whether or not fire occurrences were explained by proportions of low/open or high/medium-density developed land cover. Validation scores indicated that our models were reasonably accurate according to the available data, but differences in observed fire occurrence patterns among states raise the question of how closely the model predictions match true occurrence probabilities in some states.

Differences in wildfire reporting methods between states potentially reflect ambiguities in defining wildland fires in urban areas in addition to the logistical challenges of compiling fire records. The 13-year record of fire occurrences used in our models was obtained by compiling incident reports from thousands of disparate federal, state and local firefighting agencies, which resulted in inconsistent reporting patterns both within and among states (Fig. 1*b*). In some instances, there are clear differences in reporting methods – for example, the state of New Jersey only creates incident records for fires that required response from the state wildland firefighting agency rather than local municipal fire departments alone (Marie Cook, pers. comm.), while the state of New York collects all local fire department events reported through the National Fire Incident Reporting System (NFIRS; https://www.nfirs.fema.gov/; Christine Purpura, pers. comm.). Standardising wildfire reporting among states would have clear benefits for assessing wildfire hazards. However, this requires clear definitions of what types of events constitute a 'wildland fire', which may be ambiguous in urban areas where uncontrolled burning may occur in areas not typically considered as 'wildland' vegetation, such as city parks or vacant lots (Pyne 2019).

Implications for wildfire planning and management

Quantitative wildfire hazard and risk assessments are increasingly important tools for managing a future with the potential for growing wildfire danger. Our wildfire occurrence probability models demonstrated the need to account for small fires and small burnable areas (0.1–4 ha in size) in these assessments in the Northeast, where small, human-caused fires represent a large

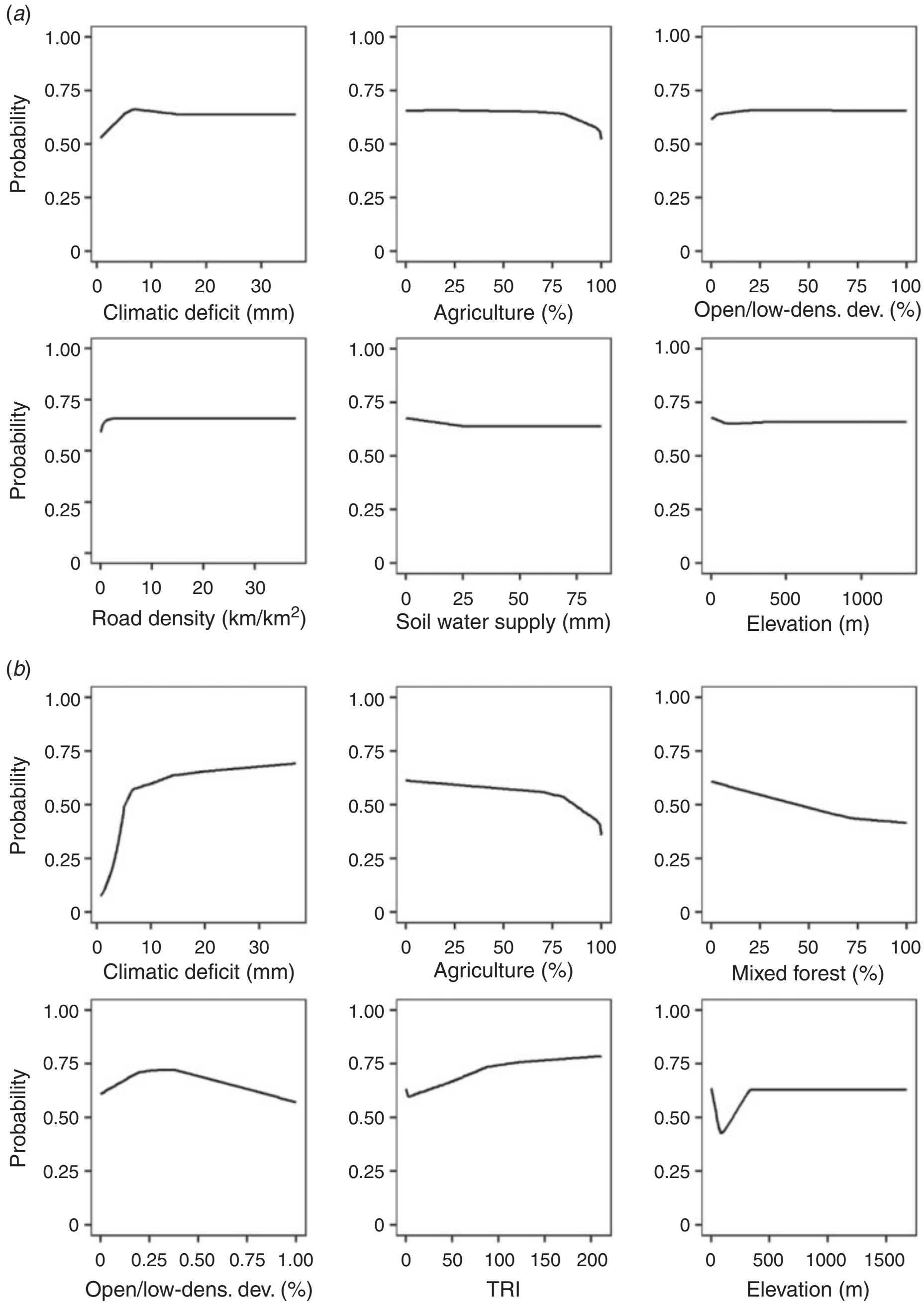

Fig. 8. Response curves for the six variables with the highest percentage contributions to the MaxEnt model predicting occurrence of (*a*) all wildfires; and (*b*) large fires (>4 ha) only for the Northeast region.

proportion of fire occurrences. Estimations of burn probability that only consider occurrences of large fires (e.g. >4 ha in our study) do not adequately capture wildfire hazards in small vegetation patches intermixed with developed land cover, as seen in the differences between our all-fires and large-fires-only models. We demonstrated that small wildfire occurrence probabilities can be accurately assessed using point-based species distribution models, which may be better-suited to characterising these types of events than simulation-based fire spread models (Finney *et al.* 2011). Our small-fire occurrence probability model can therefore be a valuable supplement to the burn probability component of national wildfire risk assessments based on large-fire hazard simulation models, such as the USFS's Wildfire Risk to Communities project (https://wildfirerisk.org/). Because these products are used to inform resource allocation for firefighting and fuel treatment, considering small-fire hazard in the WUI can have major implications for effective wildfire planning.

Accounting for hazards of small wildfires has substantial implications for fire suppression in the Northeast in terms of planning and resource allocation. Maps of fire hazard near urban areas are therefore useful for determining where firefighting resources are most needed, allocating those resources efficiently and identifying where resources are currently lacking (Thompson *et al.* 2020). Understanding seasonal differences in wildfire hazard will additionally aid in effective resource allocation. Although wildfire danger is relatively low in the Northeast, generally, warming temperatures and more extreme droughts in the next several decades have the potential to create conditions where fires are more difficult to suppress and resources may be strained by multiple simultaneous events (Butler-Leopold *et al.* 2018). During the summer of 2020, for example, multiple large fires in Massachusetts occurred owing to a combination of factors including extreme drought, an increase in use of illegal fireworks and strain on firefighting departments associated with the COVID-19 pandemic (Lisinski 2020). While fire has many positive effects in Northeast ecosystems (Pausas and Keeley 2019), effective management of ignitions near where people live will be increasingly important for minimising fire danger.

Developing more reliable assessments of small fire occurrence probabilities will depend on collecting reliable data on small wildfire occurrence patterns, requiring consistent definitions of what constitutes a significant 'wildland fire' in human-modified landscapes. Federal guidance on defining wildland fires in the USA is provided by the National Wildfire Coordinating Group, which defines a wildland fire as 'any non-structure fire that occurs in vegetation or natural fuels' (Fire Management Board 2009). This highly inclusive definition can include a wide range of event sizes in varying levels of developed land cover types. Effort is therefore needed to (1) use consistent methodology across agencies to report all fires that meet this definition; and (2) improve the accuracy of spatial information included with fire reports. Spatially complete datasets are vital for reliable quantitative wildfire hazard assessments, while fire location data that match actual burned areas can allow data users to refine wildfire definitions based on detailed fuel types or vegetation patch size. Given that frequent, small fires are important to overall wildfire hazard in regions with high WUI exposure, improved efforts to quantify these events can be of great benefit for informing fire management.

Conflicts of interest

The authors declare no conflicts of interest.

Declaration of funding

This study was funded by a grant from the Wisconsin Department of Natural Resources and the USFS Eastern Region State and Private Forestry.

Acknowledgements

We thank the National Cohesive Wildland Fire Strategy Northeast Regional Steering Committee for helping us formulate the goals of this project. Additional thanks to K. Short, M. Cook and C. Purpura for answering questions related to fire incident reporting. We also thank T. Hawbaker, M. Mockrin, D. Helmers, I. La Puma and J. Meunier for their feedback and suggestions. Finally, we thank an Associate Editor and three anonymous reviewers whose suggestions greatly improved the manuscript.

References

Abatzoglou JT, Williams AP (2016) Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 11770–11775. doi:10.1073/PNAS.1607171113

Abatzoglou JT, Dobrowski SZ, Parks SA, Hegewisch KC (2018) Terraclimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Available at http://www.climatology-lab.org/terraclimate.html [Verified 12 February 2020]

Ager AA, Palaiologou P, Evers CR, Day MA, Ringo C, Short K (2019) Wildfire exposure to the wildland urban interface in the western US. *Applied Geography* **111**, 102059. doi:10.1016/J.APGEOG.2019.102059

Balch JK, Bradley BA, Abatzoglou JT, Nagy RC, Fusco EJ, Mahood AL (2017) Human-started wildfires expand the fire niche across the United States. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 2946–2951. doi:10.1073/PNAS.1617394114

Bar Massada A, Syphard AD, Stewart SI, Radeloff VC (2013) Wildfire ignition-distribution modelling: a comparative study in the Huron–Manistee National Forest, Michigan, USA. *International Journal of Wildland Fire* **22**, 174–183. doi:10.1071/WF11178

Bowman DMJS, Williamson GJ, Abatzoglou JT, Kolden CA, Cochrane MA, Smith AMS (2017) Human exposure and sensitivity to globally extreme wildfire events. *Nature Ecology & Evolution* **1**, 0058. doi:10.1038/S41559-016-0058

Butler-Leopold P, Iverson L, Thompson FR, III, Brandt LA, Handler SD, Shannon D, Kelly M, *et al.* (2018) Mid-Atlantic forest ecosystem vulnerability assessment and synthesis: a report from the Mid-Atlantic Climate Change Response Framework project. USDA Forest Service, Northern Research Station, Technical report NRS-GTR-181. (Delaware, OH). doi:10.2737/NRS-GTR-181

Cardille JA, Ventura SJ (2001) Occurrence of wildfire in the northern Great Lakes Region: effects of land cover and land ownership assessed at multiple scales. *International Journal of Wildland Fire* **10**, 145–154. doi:10.1071/WF01010

Cary GJ, Flannigan MD, Keane RE, Bradstock RA, Davies ID, Lenihan JM, Li C, Logan KA, Parsons RA (2009) Relative importance of fuel management, ignition management and weather for area burned: evidence from five landscape fire succession models. *International Journal of Wildland Fire* **18**, 147–156. doi:10.1071/WF07085

Dillon G, Menakis J, Fay F (2015) Wildland fire potential: a tool for assessing wildfire risk and fuels management needs. In 'Proceedings of the large wildland fires conference, Missoula, MT'. (Eds RE Keane, M Jolly, R Parsons, K Riley). USDA Forest Service, Rocky Mountain Research Station, Proceedings RMRS-P-73, pp. 60–76. (Fort Collins, CO, USA)

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

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

Fire Management Board (2009) Federal wildland fire policy terms and definitions. Available at https://www.nwcg.gov/sites/default/files/docs/eb-fmb-m-19-004a.pdf [Verified 12 February 2021]

Frost CC (1995) Presettlement fire regimes in southeastern marshes, peatlands, and swamps. In 'Proceedings of the Tall Timbers Fire Ecology Conference, No. 19'. (Eds SI Cerulean, RT Engstrom) pp. 39–60. (Tall Timbers Research Station: Tallahassee, FL, USA)

Gesch DB, Evans GA, Oimoen MJ, Arundel S (2018) The national elevation dataset. In 'Digital Elevation Model Technologies and Applications: DEM Users Manual, 3rd edn'. (Eds DF Maune, A Nayegandhi) pp. 83–110. (US Geological Survey, Earth Resources Observation and Science Center, American Society for Photogrammetry and Remote Sensing: Washington DC, USA)

Grimm EC (1984) Fire and other factors controlling the Big Woods vegetation of Minnesota in the mid-nineteenth century. *Ecological Monographs* **54**, 291–311. doi:10.2307/1942499

Guyette RP, Dey DC, Stambaugh MC, Muzika RM (2005) Fire scars reveal variability and dynamics of eastern fire regimes. In 'Fire in Eastern Oak Forests: Delivering Science to Land Managers'. (Ed. MB Dickinson) USDA Forest Service, Northern Research Station, Proceedings GTR-NRS-P-1, pp. 20–39. (Columbus, OH)

Haines DA, Main WA, Frost JS, Simard AJ (1983) Fire-danger rating and wildfire occurrence in the northeastern United States. *Forest Science* **29**, 679–696.

Hawbaker TJ, Radeloff VC, Syphard AD, Zhu Z, Stewart SI (2008) Detection rates of the MODIS active fire product in the United States. *Remote Sensing of Environment* **112**, 2656–2664. doi:10.1016/J.RSE.2007.12.008

Hawbaker TJ, Radeloff VC, Stewart SI, Hammer RB, Keuler NS, Clayton MK (2013) Human and biophysical influences on fire occurrence in the United States. *Ecological Applications* **23**, 565–582. doi:10.1890/12-1816.1

Hawbaker TJ, Vanderhoof MK, Schmidt GL, Beal Y-J, Picotte JJ, Takacs JD, Falgout JT, Dwyer JL (2020) The Landsat Burned Area algorithm and products for the conterminous United States. *Remote Sensing of Environment* **244**, 111801. doi:10.1016/J.RSE.2020.111801

Hijmans R (2017) dismo: Species distribution modeling. Available at http://rspatial.org/sdm/ [Verified 4 March 2021]

Jain P, Wang X, Flannigan MD (2017) Trend analysis of fire season length and extreme fire weather in North America between 1979 and 2015. *International Journal of Wildland Fire* **26**, 1009–1020. doi:10.1071/WF17008

Kolden CA (2020) Wildfires: count lives and homes, not hectares burnt. *Nature* **586**, 9. doi:10.1038/D41586-020-02740-4

Kramer HA, Mockrin MH, Alexandre PM, Stewart SI, Radeloff VC (2018) Where wildfires destroy buildings in the US relative to the wildland–urban interface and national fire outreach programs. *International Journal of Wildland Fire* **27**, 329–341. doi:10.1071/WF17135

Leitner LA, Dunn CP, Guntenspergen GR, Stearns F, Sharpe DM (1991) Effects of site, landscape features, and fire regime on vegetation patterns in presettlement southern Wisconsin. *Landscape Ecology* **5**, 203–217. doi:10.1007/BF00141435

Lisinski C (2020) Heat and drought cause 'not so normal' fire season in Mass. Available at https://www.wbur.org/earthwhile/2020/09/10/heat-drought-fire-season-mass [Verified 10 February 2021].

Little S (1974) 'Effects of fire on temperate forests: Northeastern United States.' (Academic Press: New York, NY, USA)

Liu Y, Stanturf J, Goodrick S (2010) Trends in global wildfire potential in a changing climate. *Forest Ecology and Management* **259**, 685–697. doi:10.1016/J.FORECO.2009.09.002

Malamud BD, Millington JDA, Perry GLW (2005) Characterizing wildfire regimes in the United States. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 4694–4699. doi:10.1073/PNAS.0500880102

Martinuzzi S, Stewart SI, Helmers DP, Mockrin MH, Radeloff VC (2015) The 2010 wildland–urban interface of the conterminous United States. USDA Forest Service, Northern Research Station, Research Map NRS-8. (Newtown Square, PA)

McLauchlan KK, Higuera PE, Miesel J, Rogers BM, Schweitzer J, Shuman JK, Tepley AJ, Varner JM, Veblen TT, Adalsteinsson SA, Balch JK, Baker P, Batllori E, Bigio E, Brando P, Cattau M, Chipman ML, Coen J, Crandall R, Daniels L, Enright N, Gross WS, Harvey BJ, Hatten JA, Hermann S, Hewitt RE, Kobziar LN, Landesmann JB, Loranty MM, Maezumi SY, Mearns L, Moritz M, Myers JA, Pausas JG, Pellegrini AFA, Platt WJ, Roozeboom J, Safford H, Santos F, Scheller RM, Sherriff RL, Smith KG, Smith MD, Watts AC (2020) Fire as a fundamental ecological process: research advances and frontiers. *Journal of Ecology* **108**, 2047–2069. doi:10.1111/1365-2745.13403

Mietkiewicz N, Balch JK, Schoennagel T, Leyk S, St. Denis LA, Bradley BA (2020) In the line of fire: consequences of human-ignited wildfires to homes in the US (1992–2015). *Fire* **3**, 50. doi:10.3390/FIRE3030050

Miranda BR, Sturtevant BR, Stewart SI, Hammer RB (2012) Spatial and temporal drivers of wildfire occurrence in the context of rural development in northern Wisconsin, USA. *International Journal of Wildland Fire* **21**, 141–154. doi:10.1071/WF10133

Moritz MA, Batllori E, Bradstock RA, Gill AM, Handmer J, Hessburg PF, Leonard J, McCaffrey S, Odion DC, Schoennagel T, Syphard AD (2014) Learning to coexist with wildfire. *Nature* **515**, 58–66. doi:10.1038/NATURE13946

Motzkin G, Patterson WA, III, Foster DR (1999) A historical perspective on pitch pine–scrub oak communities in the Connecticut Valley of Massachusetts. *Ecosystems* **2**, 255–273. doi:10.1007/S100219900073

Nagy RC, Fusco E, Bradley B, Abatzoglou JT, Balch J (2018) Human-related ignitions increase the number of large wildfires across US ecoregions. *Fire* **1**, 4. doi:10.3390/FIRE1010004

NIFC (2020) Wildland fire summaries (source NICC). (National Interagency Fire Council). Available at https://www.nifc.gov/fireInfo/fireInfo_statistics.html [Verified 10 February 2021]

Nowacki GJ, Abrams MD (2008) The demise of fire and 'mesophication' of forests in the eastern United States. *Bioscience* **58**, 123–138. doi:10.1641/B580207

Omernik JM (1987) Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* **77**, 118–125. doi:10.1111/J.1467-8306.1987.TB00149.X

Parisien M-A, Miller C, Parks SA, DeLancey ER, Robinne F-N, Flannigan MD (2016) The spatially varying influence of humans on fire probability in North America. *Environmental Research Letters* **11**, 075005. doi:10.1088/1748-9326/11/7/075005

Pausas JG, Keeley JE (2019) Wildfires as an ecosystem service. *Frontiers in Ecology and the Environment* **17**, 289–295. doi:10.1002/FEE.2044

Peters MP, Iverson LR, Matthews SN, Prasad AM (2013) Wildfire hazard mapping: exploring site conditions in eastern US wildland–urban interfaces. *International Journal of Wildland Fire* **22**, 567–578. doi:10.1071/WF12177

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

Plucinski MP (2019) Contain and control: wildfire suppression effectiveness at incidents and across landscapes. *Current Forestry Reports* **5**, 20–40. doi:10.1007/S40725-019-00085-4

Pyne SJ (2019) 'The Northeast: a fire survey.' (University of Arizona Press: Tucson, AZ)

R Core Team (2019) R: A language and environment for statistical computing. (R Foundation for Statistical Computing: Vienna, Austria) Available at https://www.R-project.org/ [Verified 4 March 2021]

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, Helmers DP, Kramer HA, Mockrin MH, Alexandre PM, Bar-Massada A, Butsic V, Hawbaker TJ, Martinuzzi S, Syphard AD, Stewart SI (2018) Rapid growth of the US wildland–urban interface raises wildfire risk. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 3314–3319. doi:10.1073/PNAS.1718850115

Radosavljevic A, Anderson RP (2014) Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of Biogeography* **41**, 629–643. doi:10.1111/JBI.12227

Rhemtulla JM, Mladenoff DJ, Clayton MK (2007) Regional land-cover conversion in the US upper Midwest: magnitude of change and limited recovery (1850–1935–1993). *Landscape Ecology* **22**, 57–75. doi:10.1007/S10980-007-9117-3

Riley SJ, DeGloria SD, Elliot R (1999) Index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences* **5**, 23–27.

Rollins MG (2009) LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment. *International Journal of Wildland Fire* **18**, 235–249. doi:10.1071/WF08088

Scott JH, Thompson MP, Calkin DE (2013) A wildfire risk assessment framework for land and resource management. USDA Forest Service, Rocky Mountain Research Station, General Technical Report RMRS-GTR-315. (Fort Collins, CO)

Short KC (2014) A spatial database of wildfires in the United States, 1992–2011. *Earth System Science Data* **6**, 1–27. doi:10.5194/ESSD-6-1-2014

Short KC (2018) Spatial wildfire occurrence data for the United States, 1992–2017. Available at doi:10.2737/RDS-2013-0009.4 [verified 20 February 2020]

Soil Survey Staff (2020) Gridded Soil Survey (gSSURGO) database for the conterminous United States. Available at https://gdg.sc.egov.usda.gov/ [Verified 4 March 2021]

Spyratos V, Bourgeron PS, Ghil M (2007) Development at the wildland–urban interface and the mitigation of forest-fire risk. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 14272–14276. doi:10.1073/PNAS.0704488104

St Denis LA, Mietkiewicz NP, Short KC, Buckland M, Balch JK (2020) All-hazards dataset mined from the US National Incident Management System 1999–2014. *Scientific Data* **7**, 64. doi:10.1038/S41597-020-0403-0

Stephens SL, Collins BM, Biber E, Fulé PZ (2016) US Federal fire and forest policy: emphasizing resilience in dry forests. *Ecosphere* **7**, e01584. doi:10.1002/ECS2.1584

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

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

Syphard AD, Radeloff VC, Hawbaker TJ, Stewart SI (2009) Conservation threats due to human-caused increases in fire frequency in Mediterranean-climate ecosystems. *Conservation Biology* **23**, 758–769. doi:10.1111/J.1523-1739.2009.01223.X

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**, e71708. doi:10.1371/JOURNAL.PONE.0071708

Thompson MP, Calkin DE, Finney MA, Ager AA, Gilbertson-Day JW (2011) Integrated national-scale assessment of wildfire risk to human and ecological values. *Stochastic Environmental Research and Risk Assessment* **25**, 761–780. doi:10.1007/S00477-011-0461-0

Thompson MP, Gannon BM, Caggiano MD, O'Connor CD, Brough A, Gilbertson-Day JW, Scott JH (2020) Prototyping a geospatial atlas for wildfire planning and management. *Forests* **11**, 909. doi:10.3390/F11090909

US Geological Survey (2017) USGS National Transportation Dataset (NTD) downloadable data collection. Available at https://www.sciencebase.gov/catalog/item/4f70b1f4e4b058caae3f8e16 [Verified 6 April 2020].

Whitney GG (1987) An ecological history of the Great Lakes forest of Michigan. *Journal of Ecology* **75**, 667–684. doi:10.2307/2260198

Yang J, He HS, Shifley SR (2008) Spatial controls of occurrence and spread of wildfires in the Missouri Ozark highlands. *Ecological Applications* **18**, 1212–1225. doi:10.1890/07-0825.1

Yang L, Jin S, Danielson P, Homer C, Gass L, Bender SM, Case A, Costello C, Dewitz J, Fry J, Funk M, Granneman B, Liknes GC, Rigge M, Xian G (2018) A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. *ISPRS Journal of Photogrammetry and Remote Sensing* **146**, 108–123. doi:10.1016/J.ISPRSJPRS.2018.09.006