CSIRO PUBLISHING

International Journal of Wildland Fire **2016**, *25*, 715–729
http://dx.doi.org/10.1071/WF15124

Projecting wildfire area burned in the south-eastern United States, 2011–60

Jeffrey P. Prestemon[A,F], *Uma Shankar*[B], *Aijun Xiu*[B], *K. Talgo*[B], *D. Yang*[B], *Ernest Dixon IV*[C], *Donald McKenzie*[D] *and Karen L. Abt*[E]

[A]USDA Forest Service, Southern Research Station, PO Box 12254, Research Triangle Park, North Carolina 27709, USA.
[B]Center for Environmental Modeling for Policy Development, University of North Carolina Institute for the Environment, Campus Box 1105, Room 4040 Suite 490 Europa Center, 100 Europa Drive, Chapel Hill, NC 27517, USA.
[C]Global Forest Partners LP, 67 Etna Road, Suite 500, Lebanon, NH 03766, USA.
[D]Pacific Wildland Fire Sciences Laboratory, USDA Forest Service, 400 North 34th Street, Suite 201, Seattle, WA 98103, USA.
[E]USDA Forest Service, Southern Research Station, PO Box 12254, Research Triangle Park, North Carolina 27709, USA.
[F]Corresponding author. Email: jprestemon@fs.fed.us

Abstract. Future changes in society and climate are expected to affect wildfire activity in the south-eastern United States. The objective of this research was to understand how changes in both climate and society may affect wildfire in the coming decades. We estimated a three-stage statistical model of wildfire area burned by ecoregion province for lightning and human causes (1992–2010) based on precipitation, temperature, potential evapotranspiration, forest land use, human population and personal income. Estimated parameters from the statistical models were used to project wildfire area burned from 2011 to 2060 under nine climate realisations, using a combination of three Intergovernmental Panel on Climate Change-based emissions scenarios (A1B, A2, B2) and three general circulation models. Monte Carlo simulation quantifies ranges in projected area burned by county by year, and in total for higher-level spatial aggregations. Projections indicated, overall in the Southeast, that median annual area burned by lightning-ignited wildfire increases by 34%, human-ignited wildfire decreases by 6%, and total wildfire increases by 4% by 2056–60 compared with 2016–20. Total wildfire changes vary widely by state (−47 to +30%) and ecoregion province (−73 to +79%). Our analyses could be used to generate projections of wildfire-generated air pollutant exposures, relevant to meeting the National Ambient Air Quality Standards.

Additional keywords: climate change, human-caused wildfire, land use, lightning-caused wildfire.

Received 13 July 2015, accepted 5 April 2016, published online 2 June 2016

Introduction

Wildfire activity in the south-eastern USA is determined by two major factors: climate and society. Climate change is expected to alter patterns of precipitation, temperature and the severity of droughts, which will impact on the accumulation of fuels and the occurrence of favourable wildfire conditions in the region (e.g. Liu *et al.* 2013). Simultaneously, it has long been recognised (e.g. Prestemon *et al.* 2002; Mercer and Prestemon 2005; Prestemon and Butry 2005; Mercer *et al.* 2007; Prestemon *et al.* 2013) that humans are a dominant factor affecting ignition rates and the arrangement of fuels on the landscape. Although humans start most wildfires in the region, they also devote substantial resources to suppressing wildfires to limit their areal extent and associated damages (e.g. Butry *et al.* 2001), and they manage fuels, in part, to limit wildfire severity and intensity. Therefore, to gain insight into how climate change is likely to affect wildfire (e.g. Liu *et al.* 2013), the role of humans must be considered.

In addition to uncertainty regarding capturing the direct roles of humans in influencing ignition sources and fuels distributions and, thereby, the overall expected extent of wildfires, there is uncertainty about how society and climate will jointly evolve over time. The Intergovernmental Panel on Climate Change (IPCC) has published several assessments describing how climate will change under different greenhouse gas (GHG) emission scenarios. Nakicenovic and Steward (2000) elaborate how these emissions scenarios emerge under varying assumptions regarding energy policies, economic output trends and human population growth. These varying assumptions, along with the projected changes in climate, define scenario-storylines that were the data inputs into descriptions of various possible land-use futures in the south-eastern US (e.g. Wear 2013), including those of forestland, a critical variable affecting wildfire activity. The scenario-storylines were also the inputs into projections of possible futures regarding population and income growth in the region (and worldwide). Nobody knows precisely how emissions or society will change, but capturing the variety of potential changes in all of these variables (as inferred by the approach used by Littell *et al.* (2009), for example) would be an important step towards understanding how they combine to produce a picture of a wildfire future for the Southeast.

The primary objective of the present research was to project area burned in 13 south-eastern states of the United States during the period 2011–60, while accounting for projected changes in climate and society, including land use. To do this, we built statistical models of wildfire annual area burned, estimated by county within ecoregion provinces (Bailey 1995) for the historical period 1992–2010. Models were specified as functions of temperature, precipitation, potential evapotranspiration, personal income, population and land use. The estimated statistical models were then used in projection mode, using exogenous projections to 2060 of the same variables used in the model-building phase of the analysis. Three different emission scenarios were used in each of three general circulation models to generate nine realisations of future climate. The three emissions scenarios were A1B and A2 from the IPCC's Third Assessment Report, and B2, from a somewhat earlier model (see Joyce *et al.* 2014). Projections of land use, population and personal income, provided by the 2010 Resources Planning Act (RPA) Assessment (USDA Forest Service 2014*a*, 2014*b*) for each of these three scenarios, were combined to enable modelling to generate envelopes of possible futures of wildfire area burned, annually to 2060, by county, which can then be reported by state and by ecoregion province. Because of our attention to land-use and societal changes, motivated by theory regarding wildfire management (e.g. Donovan and Rideout 2003) and empirical evidence (Prestemon *et al.* 2002; Mercer and Prestemon 2005; Mercer *et al.* 2007; Butry *et al.* 2010), the present effort differs from studies that focus only on wildfire changes, and due only to climate change. The wildfire projections resulting from this effort therefore provide a more comprehensive picture of the range of possible futures of wildfire in the region by capturing projected changes in human factors.

The modelling and projection effort we describe differs significantly from other efforts to project changes in area burned, and their consequences, in a changing climate. A comprehensive review of past efforts is in Flannigan *et al.* (2009), who enumerate and classify modelling studies that project future wildfire. Our study extends these earlier efforts in three ways. First, this is the first study we are aware of to project future area burned in the south-eastern US. Second, we estimate not only the effects of climate change, but also land-use changes and shifts in the societal factors that affect area burned. Third, we address issues of data quality that are commonly found in wildfire activity databases (Short 2015), using a technique (Heckman 1979) common in the econometrics literature but not, as far as we are aware, applied in models for projecting wildfire area burned.

Methods

Theoretical structure

Wildfire area burned (W) by either lightning or human causes is posited to be dependent on the availability of ignition sources, fuels, weather favourable for ignition and spread, and intentional and unintentional contributions by humans ('Society'):

$$W = f(Weather, Fuels, Society) \quad (1)$$

Mercer and Prestemon (2005) documented how wildfire is affected by these sets of variables, while Mercer *et al.* (2007) described the theoretical underpinning for managerial interventions in wildfire processes. Consistent with descriptions by Donovan and Rideout (2003) and Mercer *et al.* (2007), among others, we also describe a Cost plus Net Value Change (CNVC) model, whose objective function minimises expected CNVC with respect to wildfire management input quantities. Abstracting from Mercer *et al.* (2007) by ignoring terms associated with long-run discounting, we have:

$$min_{\mathbf{x}} E[CNVC] = \mathbf{w}'\mathbf{x} + VE[W(\mathbf{x})] \quad (2)$$

where $\mathbf{x}$ is a vector of wildfire management inputs; $\mathbf{w}$ is a conformable vector describing the costs of those inputs; V quantifies the value at risk per unit of wildfire area burned; W is area burned; and E is the expectations operator (because area burned is a stochastic variable). In solving Eqn 2, at the optimum, the last unit of wildfire management input deployed is valued equally to the cost of the last unit of loss averted owing to the unit of input applied. In other words, wildfire managers (or society, more broadly) act to deploy costly and scarce resources (labour, capital, materials) to wildfire management in order to avert even greater losses of social welfare (values at risk) due to wildfire. One implication of the wildfire manager's actions is that, as values at risk increase owing to increased prevalence of weather and fuel favourable to fire ignition and spread, the optimal amount of resources devoted to wildfire management will also increase. Generally, on private lands, particularly in the south-east US, where private lands dominate, changes in these values at risk are beyond the direct control of wildfire managers; they instead derive from the broader economy and society.[1] In empirical modelling, values at risk can be represented by the number of people living in fire-prone landscapes, the value of their structures, and the value of vulnerable natural resources such as timber (e.g. Butry *et al.* 2001). Therefore, as the density (number of structures per unit of fire-prone landscape) rises along with human population and wealth, so will V in Eqn 2. Similarly, as wealth increases, the (market) value of each additional structure is also expected to rise (e.g. Mankiw and Weil 1989), further increasing V. Because timber and most other natural resources have market values that are far less than those of structures, development that reduces the quantity of natural resources on the landscape would only slow the rate of increase with rising wealth for a V that is calculated solely using market values.

[1]Landowners and public land managers, however, can alter the quantity of values at risk by managing fuels, reducing unwanted human-caused wildfire ignitions through fire prevention efforts, and making structures more fire-resistant.

The quantity of the values at risk can change not only owing to the weather, but also because of the availability of fuels. Because fuels are more contiguous and abundant in forests compared with other land uses in the south-eastern US, greater forest area can be linked to greater wildfire area burned, holding physical variables affecting wildfire (e.g. slope, weather conditions) constant. Increased population, requiring more structures, is associated with increased road densities and other interruptions of fuel contiguity on landscapes, which can slow wildfire spread, and provide easier access for firefighting, leading to less area burned (e.g. Mercer and Prestemon 2005; Narayanaraj and Wimberly 2012; Syphard *et al.* 2012). So the forest loss associated with increased population and economic activity can decrease values at risk by limiting how much wildfire burns on the landscape – although the link between area burned and values at risk is not necessarily linear or constant across space, or even across wildfires, owing to variations in fuels, wildfire intensities and severities, and the valuable resources, structures, and humans that each fire encounters. Likewise, humans intentionally or unintentionally ignite most wildfires in the United States at large, and in the Southeast in particular (Prestemon *et al.* 2013). These ignitions can be reduced by wildfire prevention (Butry *et al.* 2010) and law enforcement efforts (Prestemon and Butry 2005). A greater number of humans, holding other variables constant, would imply a greater number of fire ignitions due to accidental and intentional actions, but greater values at risk and larger human populations imply greater wildfire prevention and law enforcement efforts. Therefore, the net effect of increases in population and values at risk on wildfire ignitions is ambiguous, owing to the competing influences of greater prevention efforts and law enforcement, and greater human contact with fuels.

In modelling annual area burned, analysts need to acknowledge the possibility that the wildfire area burned data may be incomplete (e.g. Malamud *et al.* 2005; Short 2015) and also that wildfire occurrence is highly variable when viewed across broad landscapes. In the statistical modelling for the present study, we accounted for both of these phenomena partly owing to suggestions by K. C. Short (unpubl. data) regarding data adequacy for every county and year in our historical dataset. Another reason was our conception that the processes involved in determining whether a wildfire occurs at all in a spatial and temporal unit of inference is different from that determining total area burned in the spatial-temporal unit (Mercer and Prestemon 2005). Not accounting for either data quality or the truncated distribution of wildfire area burned in estimating statistical models (e.g. simply regressing area burned in a spatial-temporal unit on a set of predictors) could lead to biased and inconsistent estimates of the area burned production process (e.g. Heckman 1979; Greene 1992; Short 2014; Short 2015). To account for variable data quality and area burned distribution truncation, we specify an empirical model structure that accounts for sample selection (Heckman 1979) at two stages in advance of estimating a final-stage equation of wildfire area burned by cause category. Stage 1 is on the existence of a 'valid' observation, where 'valid' has been determined by K. C. Short (unpubl. data) based on whether the observation on reported wildfire area burned (a valid observation could be 0 ha burned, for example) is likely to reflect actual wildfire area burned accurately for each spatial and temporal unit. This stage involves estimating a statistical model that predicts whether each spatial-temporal unit of observation is valid. A summary statistic of this equation estimate, the Inverse Mills Ratio (IMR), measures the likelihood of each observation's validity. Stages 2 and 3 are estimated using a two-step estimator. Stage 2 controls for the truncation of the dependent variable at zero, and quantifies the probability that a valid observation has zero reported wildfire, given the IMR from the first stage, as well as a set of additional exogenous predictor variables. Finally, Stage 3 is an ordinary least-squares equation relating the area burned, if non-zero, to a set of exogenous predictors and the IMRs from Stages 1 and 2.

More specifically, in the statistical modelling, we estimated equations at each stage sequentially: we estimated the first stage and calculated the first-stage IMR; estimated the second stage that included the first stage IMR as an additional predictor, and, with that equation, calculated the second-stage IMR; and estimated the third stage that included both IMRs as predictors. The first stage of the three-stage wildfire model controls for the selection effect of inadequate data. All observations for all counties and all years in our dataset were coded as 1 if not valid, and 0 otherwise. Although the value of the IMR varies from one observation to another, for all observations included in an equation with the IMR from an earlier stage, the coefficient measures the direction of the biasing effect of sample selection.

This first-stage equation, a probit model, controls for the potential biasing effects of non-randomness in the sample of observations making it into the next stage of estimation. Heckman's (1979) insight is that the sample selection (dropping of particular observations) is a form of omitted variables bias, which can lead to incorrect inference and lead to poor out-of-sample performance of the resulting estimated equation. In the Heckman (1979) approach, the probit model explaining the selection bias is based on a theory explaining the selection process. When the resulting IMR is included in the subsequent equation estimated on the remaining data, parameter estimates are, in the limit, unbiased. The first-stage probit model is (Greene 1992, p. 663):

$$Prob[Y_i = 1] = \Phi\left(\boldsymbol{\beta}_1' \mathbf{x}_{1,i}\right) \tag{3}$$

where $Y_i = [0,1]$ is a discrete variable identifying whether the observation on wildfire area burned (in our case) is valid (0) or not valid (1); $\Phi(\cdot)$ is the Standard Normal probability distribution function; $\mathbf{x}_{1,i}$ is a vector of variables for observation i that are hypothesised to be related to the validity of the observation; and $\boldsymbol{\beta}_1$ is a vector of estimation parameters conformable to $\mathbf{x}_1$. From an estimate of Eqn 1, the IMR for the ith observation ($\lambda_{1,i}$) can be calculated (Heckman 1979):

$$\begin{aligned} \lambda_{1,i} &= \frac{\phi(Z_{1,i})}{\Phi(-Z_{1,i})}, \\ Z_{1,i} &= -\boldsymbol{\beta}_1' \mathbf{x}_{1,i}/\sigma_1 \end{aligned} \tag{4}$$

where $\phi(\cdot)$ is the Standard Normal probability density and σ_1 is the standard error of the estimate of the residuals in Eqn 3. Variables in the set of predictors ($\mathbf{x}_1$) for observation validity could include indicators of the state from which the observation derives, because state and federal agencies in charge of data

Table 1. Groups of ecoregion provinces and the selection model stages in the final versions of statistical models

Ecoregion provinces included in the model	Lightning wildfire selection model stages	Human wildfire selection model stages
Eastern Broadleaf Forest (Oceanic) (221)	2	3
Eastern Broadleaf Forest (Continental) (222)	3	3
South-eastern Mixed Forest (231)	3	3
Coastal Plain mixed Forest (232), Lower Mississippi Riverine Forest (234), Everglades (411)	3	3
Prairie Parkland (Temperate) (251), Prairie Parkland (Subtropical) (255), Great Plains Steppe and Shrub (311), South-west Plateau and Plains Dry Steppe and Shrub (315), Chihuahuan Semidesert (321), Great Plains–Palouse Dry Steppe (331), Great Plains Steppe (332)	2	2

reporting may have variable success across space and time in reporting wildfires completely and in a statistically unbiased way (Short 2014; Short 2015). Variables in $\mathbf{x}_1$ could also include measures of weather, fuels and human factors that are themselves hypothesised to be related to wildfire processes.

The second stage evaluates whether an observation that is valid has recorded wildfire area burned that is greater than zero.[2] For the present study, this second stage seeks to explain why the area burned by a human- or lightning-ignited wildfire is greater than zero in a particular county in a particular year. This stage includes the IMR from Stage 1 as an additional explanatory variable:

$$Prob[A_i = 1|Y_i = 0] = \Phi\left(\boldsymbol{\beta}_2'\mathbf{x}_{2,i} + \gamma_1\lambda_{1,i}\right) \quad (5)$$

where $A_i = [0,1]$ indicates if the area burned for observation i is zero ($A_i = 0$) or positive ($A_i = 1$); $\mathbf{x}_{2,\mathrm{i}}$ is a vector of variables for observation i that are hypothesised to be related to whether wildfire area burned was greater than zero for observation i; $\boldsymbol{\beta}_2$ is a vector of estimation parameters conformable to $\mathbf{x}_2$; and γ_1 is a parameter measuring the effect of sample selection in the first stage on the probability of non-zero wildfire area burned, which controls for the potentially biasing effects (on the estimate of $\boldsymbol{\beta}_2$) of observation validity. Examples of variables in $\mathbf{x}_2$ affecting wildfire area burned could include summary measures of weather, fuels and societal factors. As in the first stage, the IMR for the ith observation of the second stage ($\lambda_{2,i}$) can also be calculated for each observation where $Y_i = 0$:

$$\lambda_{2,i} = \frac{\phi(Z_{2,i})}{\Phi(-Z_{2,i})}, \qquad Z_{2,i} = -\frac{\boldsymbol{\beta}_2'\mathbf{x}_{2,i} + \gamma_1\lambda_{1,i}}{\sigma_2} \quad (6)$$

where σ_2 is the standard error of the estimate of the residuals in Eqn 5.

The third and final stage of modelling of wildfire area burned is a least-squares equation relating non-zero area burned to a set of predictors $\lambda_{1,i}$, and $\lambda_{2,i}$ into the equation specification to control for both the validity of the observation and the likelihood of zero area burned:

$$E[W_i|Y_i = 0, A_i = 1] = \boldsymbol{\beta}_3'\mathbf{x}_{3,i} + \alpha_1\lambda_{1,i} + \alpha_2\lambda_{2,i} \quad (7)$$

where $W_i > 0$ is area burned; $\mathbf{x}_{3,i}$ is a vector of variables for observation i that are hypothesised to be related to area burned, given that an observation was both valid and greater than zero; $\boldsymbol{\beta}_3$ is a vector of estimation parameters conformable to $\mathbf{x}_3$; and α_1 and α_2 control for the potentially biasing effects (on the estimate of $\boldsymbol{\beta}_3$) of the probability that the observation is valid, and the probability that it is non-zero respectively. Variables contained in $\mathbf{x}_3$ could be the same variables (in $\mathbf{x}_2$) used to predict whether a unit of observation had zero wildfire.

Data and model development

Historical data

Historical wildfire data on annual area burned by lightning and annual area burned by human causes (which we call 'human-ignited') in acres by county for the years 1992–2010 were obtained from K. C. Short (unpubl. data) for estimating Eqns 1–5. Data were aggregated (i.e. the area burned was added up across all wildfires of (i) lightning cause, and (ii) all wildfires of other causes, including unknown cause, which we label 'human') to the county spatial unit and annually, based on the county of origin and the date of wildfire ignition. As recommended by K. C. Short (unpubl. data), observations of wildfire area burned by lightning and humans were flagged as not valid, but the indicator of observation validity and a set of predictor variables associated with each observation were used in the first of the three stages of the wildfire statistical model estimation process. Each model of lightning wildfire and of human-ignited wildfire was separately estimated for five spatial domains based on ecoregion provinces found in the south-eastern US (Bailey 1995), shown in Table 1. Counties were assigned to whole ecoregion provinces based on Butry (2003). Fig. 1 is a map of the ecoregion provinces for the Southeast.

[2]An alternative functional form at this stage would be the Tobit model rather than a Heckman sample selection model. Preliminary tests showed that the Tobit performed significantly worse in out-of-sample predictions compared with the Heckman sample selection model.

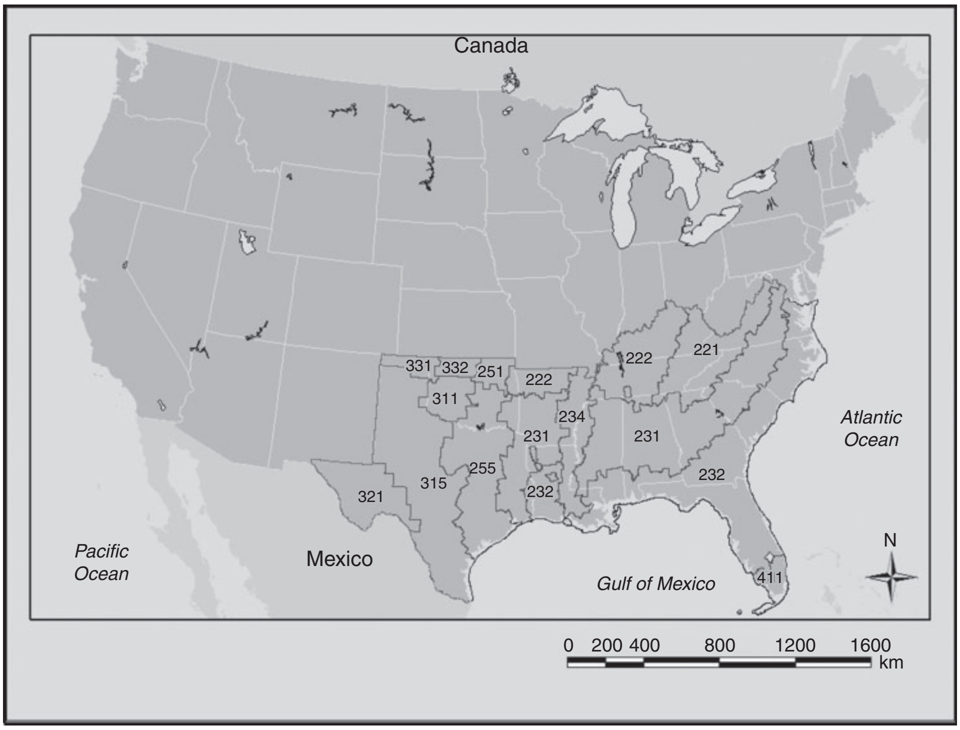

Fig. 1. Bailey's ecoregion provinces in the south-eastern US, with delineations drawn at county lines (source: Butry 2003).

The form of weather observations included as predictors in the three-stage statistical models of wildfire annual area burned was based on what was available from temporal and spatial downscaling emerging from general circulation model (GCM) projections. Weather data for the historical time series (1992–2010) were obtained from the historical data assembled for the Parameter-elevation Regressions on Independent Slopes Model (PRISM) developed by Daly *et al.* (2002). Potential evapotranspiration (PET) was calculated using a modified Linacre (1977) method described by Joyce *et al.* (2014). Daily data on these four measures for each county were summarised as monthly values (average maximum and minimum temperature, total precipitation, average PET) for each county. In equation specifications for Eqns 3–7, only the aggregated observations for January, March, May, July and September were used. The January–September range corresponded best with fire seasons across the south-eastern US. Initial data exploration indicated that intertemporal correlations of monthly meteorological variables fall with temporal distance; including all months runs the risk of multicollinearity. Hence, we omitted meteorological observations from the intervening months in the equation estimation.

Data on county land area and historical forest land use by county were obtained from the USDA Forest Service (2014*a*), with methods on land use described in Wear (2013). Data for the historical time series were reported only for 1997 and for 2010. Data for intervening years were obtained by linear interpolation (1998–2009) and extrapolation (1992–96).

Data on population by county in the historical time series were obtained from the US Census Bureau (2012). Data on annual personal income by county (US Bureau of Economic Analysis 2013*a*), 1992–2010, were converted to real values (in constant 2005 US dollars) with the US gross domestic product deflator (US Bureau of Economic Analysis 2013*b*).

Projected data

Climate inputs for the statistical models for the projection period, 2011–60, were statistically downscaled from three GCMs (Daly *et al.* 2002), namely the MIROC32, the CSIR-OMK35 and the CGCM31, for each of three IPCC emission scenarios (A1B, A2, B2), producing nine climate projections for the region. The statistically downscaled data for each of these nine projections were obtained from Joyce *et al.* (2014), who produced them for the 2010 RPA assessment using data from the IPCC, at 5-arcmin resolution in latitude–longitude coordinates. They were remapped to a Lambert Conformal Conic map projection grid at 12-km resolution for an air-quality modelling domain that includes 13 states from eastern Texas to the Carolinas east to west, and Kentucky to Florida north to south, for eventual use in regional air quality assessments. To generate county-level meteorological data for area burned projections, county-level averages of daily maximum and minimum temperature, total precipitation and PET were averaged across all of the model grid cells whose majority area was in the county.

Land area and land-use data by county and by year for each emissions scenario (A1B, A2, B2) in the projection years were obtained from Wear (2013). Projections of land use were by 10-year increments (2010, 2020, 2030, 2040, 2050, 2060) for each of the three scenarios. Intervening-year values for each county were calculated using linear interpolation. Data on

Table 2. Variables used in the initial empirical specification for each stage of the empirical modelling

	Stage 1 (data validity flag, probit)		Stage 2 (non-zero wildfire, probit)		Stage 3 (least-squares)	
	Lightning	Human	Lightning	Human	Lightning	Human
Monthly average daily maximum temperature (°C)	X	X	X	X	X	X
Monthly average daily PET	X	X	X	X	X	X
Monthly total precipitation (mm)	X	X	X	X	X	X
Land area (km^2)	X	X	X		X	X
Forest land-use area (km^2)	X	X		X	X	X
Population	X	X	X	X	X	X
Population density (population per land area)	X	X	X	X	X	X
Personal income per capita (real personal income, US$ per population)	X	X	X	X	X	X
State dummy variables	X	X	X	X	X	X

population and personal income (real, in constant 2005 US dollars) were obtained from USDA Forest Service (2014*b*) projections made at 5-year increments for each scenario. Intervening-year values were calculated using linear interpolation. Land areas by county were held at the initial totals for 1997 reported in Wear (2013).

We chose the older climate projection models that supported the IPCC Third Assessment Report, The Coupled Model Intercomparison Project Phase 3 (or CMIP3), rather than the analogous supporting models of the Fifth Assessment Report (CMIP5) for two main reasons. First, unlike the CMIP5 scenarios, the CMIP3 scenarios are directly, mechanistically linked to internally consistent projections of human population and economic growth that are tied to the emissions scenarios – that is, they form an internally consistent picture of societal development under alternative emissions pathways. Second, projections of county-level income, population and land use were produced in the 2010 RPA Assessment, providing a ready dataset of not just the evolution of climate at fine spatial scales in the United States, but also variables that we assert are connected to wildfire. Updating the projections based on CMIP5, including more recently available alternative representative concentration pathways (RCPs), would require identifying population and economic projections for the United States Southeast, and hence the region's land-use futures, that are consistent with the CMIP5 RCPs. The extra effort required is beyond the scope of this analysis. We note that the CMIP5 RCPs, if we were able to use them, could conceivably alter our conclusions, but Knutti and Sedláček (2013) showed that the CMIP5 projections *per se* are not significantly different from the CMIP3 (and fourth phase, CMIP4), and their uncertainty remains almost the same, even though climate projection methods themselves have been refined in later iterations. Additional evidence of their similarity is available from IPCC Working Group I Technical Support Unit (2015, p. 811).

Model development

Given the large number of potential predictor variables, we employed a model selection approach to final model identification in all three stages of the statistical modelling. The list of predictor variables used in the initial (full) equation specifications for each of the three stages differed slightly across stages, across lightning- vs human-caused wildfires, and across ecoregion provinces. Differences across ecoregion provinces involved dummy variable intercept shifters measuring the effects of individual states and constituent ecoregion provinces for cases where the data for counties in multiple ecoregion provinces were included together in a single ecological domain model (e.g. the models for aggregated ecoregion provinces 232, 234 and 411).

In model estimation, we began with the full list of variables shown in the first column of Table 2 and estimated models using a subset of the data – i.e. only covering 1992–2003, which we call the 'in-sample' data. We reduced the numbers of predictor variables to arrive at final specifications for each stage in every lightning- and human-caused wildfire in each ecological domain by dropping statistically insignificant variables until all included variables were at P values smaller than 0.10. Once these final specifications were found, we examined their out-of-sample forecast performances for the held-out sample (2004–10), i.e. the out-of-sample data. Fit statistics for the out-of-sample forecasts were recorded, and then the final specifications based on the in-sample data were re-estimated using the entire dataset (1992–2010). It was these whole-sample model estimates that were used in the projection modelling (2011–60). Statistics of wildfire area burned by lightning and by human causes are summarised for all of the valid observations of annual area burned in the counties within the ecoregion provinces included in model estimation (Table 3).

As Heckman (1979) explained, bias can be eliminated (in large samples) or reduced (in smaller samples) in the presence of sample selection as long as the first stage of the selection model accurately explains the selection process. We hypothesised that the validity of the observation of wildfire area burned in a county in a year (i.e. that $Y_i = 0$) was related in unknown ways to the efforts by the government fire-occurrence reporting agencies to record wildfire occurrences consistently, which were in turn related to factors affecting the overall extent of wildfire (e.g. large fires are more likely to be observed and recorded). For this reason, we included the same variables in the specification of Eqn 3 as we did in subsequent stages in the modelling. Model selection (the process of dropping insignificant variables to arrive at a specification) naturally led to a specification of Stage 1 equations that differed from those of subsequent stages.

One concern in employing the model selection approach was the introduction of biases in statistical models that could have led to poor out-of-sample forecasts of area burned. A comparison of the area burned predictions made in-sample (i.e. using models

Table 3. Summary statistics for wildfire area burned for annual area burned by county, in hectares, 1992–2010

		Ecoregion province number				
		221	222	231	232, 234, 411	251, 255, 311, 315, 321, 331, 332
Lightning wildfire	Valid observations	1601	1423	4687	4504	1132
	Mean area burned observed (ha)	12	5	15	127	24
	Standard deviation area burned observed (ha)	95	110	143	1599	182
	Minimum area burned observed (ha)	0	0	0	0	0
	Maximum area burned observed (ha)	2361	4047	5564	50226	3844
Human-caused wildfire	Valid observations	1601	1423	4687	4504	1132
	Mean area burned observed (ha)	244	113	154	271	403
	Standard deviation area burned observed (ha)	701	348	396	928	1804
	Minimum area burned observed (ha)	0	0	0	0	0
	Maximum area burned observed (ha)	10884	5404	10035	29278	32786

estimated with data for 1992–2003) with the area burned predicted out-of-sample (i.e. 2004–10) allowed us to evaluate whether the model selection procedure resulted in reduced forecast performance. Goodness-of-fit measures included bias (average prediction error in area burned), percentage bias, root mean-squared error, maximum absolute error and mean absolute error. Bias was a primary criterion used to evaluate forecast performance, whereas the other fitness measures were used secondarily. If bias was detected in out-of-sample performance, we returned to the full model specifications and did not drop some 'borderline' variables (i.e. those with P values in the 0.10 to 0.20 range). This process was continued until a final specification for each stage was identified.[3]

The final empirical specification of the three-stage selection models depended in part on whether the coefficient estimate for the IMR from the first stage was statistically significant in the second-stage model estimate. If not, the first stage was dropped for the final version of the second- and third-stage model estimates (Table 1), resulting in the standard two-stage selection model (i.e. Stages 2 and 3 in our framework) described by Heckman (1979).

Not included among the list of potential predictor variables (Table 2) were direct measures of wildfire suppression, prescribed fire, the fuel treatment effects of wildfire, or fire prevention. Firefighting organisations in the United States seek to comply with standards advanced by the National Fire Protection Association, which recommend that capacity levels adjust to changing densities of structures and other landscape features at risk of fire to maintain minimum staffing and average response times (e.g. National Fire Protection Association 2015). We therefore assume that firefighting agencies adjust their staffing and response times in step with changes in values at risk that were measured by the land-use and other socioeconomic variables included in the statistical equations. Omission of variables that would account for the fuel treatment effects of wildfire (e.g. Mercer *et al.* 2007), which tend to reduce area burned in subsequent periods, implies that such treatment effects are accounted for generally by the intercepts and variables accounting for states and ecoregions, which capture historical wildfire activity. A shortcoming of this assumption is that in places with large positive changes in projected wildfire, these projections would be positively biased, and the reverse would be true for locations with large negative changes in projected wildfire. Another underlying assumption is that the relationship of wildfire to the included variables takes into account the effects of prescribed fire on annual area burned. Prescribed fire currently is, and has been historically, an important part of public and private management of southern pine forests, especially in ecoregion provinces 231 and 232. The implication of explicitly omitting them in these two provinces in particular is that prescribed fire activity, including operational windows for burning, is assumed to be explained by the societal and climatic variables already appearing in the estimated statistical models. Finally, studies show that wildfire prevention efforts are effective in reducing occurrences of wildfire (e.g. Prestemon *et al.* 2010). An assumption of our modelling was therefore that the other included variables in the statistical models adequately accounted for the spatial and temporal variations in wildfire prevention efforts.

Monte Carlo simulations

Future states of nature are inherently uncertain; this applies to both climate and societal factors affecting wildfire. In our analyses, annual area burned for each county was predicted using three-stage statistical models using climate variables extracted from each of the nine climate model realisations.[4] All such models have uncertainties regarding their predictions, and this uncertainty can be used, in conjunction with the model uncertainty in the nine-member ensemble of GCM–emission scenario realisations, to describe the range of potential futures. Median annual area burned projections (where the median is the middle projected value in a Monte Carlo simulation) and probability bounds around the median can provide the analyst

[3]Goodness-of-fit measures were also used to compare the three-stage approach with a simple Tobit model, which predicted area burned as a truncated variable (continuous from zero), and ignored model selection. We found that the Tobit generally had larger bias than the three-stage selection model.

[4]The models estimated in the first-stage equations using historical data (Appendix S1, Tables S1 through S4) provided the probit model parameter estimates needed to generate IMRs for projection years for all counties.

with a sense of the range of future possible wildfire outcomes for any desired location or area of inference (down to the smallest modelling unit).

To describe the range of potential outcomes in our projections, we conducted Monte Carlo simulations of equation estimates. These simulations randomly selected samples, with replacement, from among the observations in the historical time series, and then used the resulting equation estimates in each iteration to project wildfire in each county in each of the years 2011–60; 2006 and 2010 wildfire burned areas were also predicted using historical data. Across all scenario–GCM combinations, these produced a total of 2250 iterations of projections for each county for each year. From these iterations, we summarised quantile distribution bounds of expected wildfire for every county, for both lightning- and human-ignited wildfire, and their sum. We also did these summaries for other higher-level spatial units of inference, including by state, and by ecosystem province-based ecological domain of counties in the south-eastern US.

Equation estimation results

The first-stage equations (*Stata/MP 13.1*, StataCorp LP, College Station, TX) were used in all statistical modelling; probit models predicting whether a county's observation of wildfire area burned for the calendar year was valid ($Y_i = 0$) or considered to be not valid ($Y_i = 1$) were the same for both lightning- and human-ignited wildfires, because K. C. Short (unpubl. data) judged validity for the entire county and year in question for all wildfires recorded, regardless of cause. Equation estimates (summarised in Tables 4 and 5, with detailed statistical results shown in Appendix S1, Tables S1–S4, available as online supplementary material) showed that non-valid data were often related to the state where the county was located, as well as to factors that affected the number of wildfire occurrences. The direction of the effect of many weather-related variables indicated that the probability of a non-valid observation was related in a complex way to weather and climate, with some months showing positive relationships, and others negative, for each of maximum temperature, PET and precipitation. For ecoregion province 221, there was a negative relationship of non-valid data with personal income per capita and a positive relationship with population density. For 222, a negative relationship to personal income per capita was also found, whereas population was positively related. Forest area in this equation was negatively related, which might be linked to the idea that wildfires occurring in heavily wooded areas often go unreported, or could be connected to other problems with reporting accuracy. For counties in 231, land area was negatively related, whereas forest area was positively related to data validity, a result that is difficult to explain. It should be noted that the first-stage probit model for counties in ecoregion provinces 251, 255, 311, 315, 321, 331 and 332 (whose counties were grouped together in a single ecological domain for modelling) had broad insignificance, including the intercept. Thus, for both lightning- and human-ignited wildfire models, the three-stage approach collapsed to a two-stage model in this ecological domain.

The results of second- and third-stage statistical models are summarised in Tables 4 and 5 (with detailed results available in Appendix S1, Tables S5–S14). For the lightning wildfires (Table 4), the IMR produced by the first-stage equation was found to be positively related to the probability that wildfire in the county was zero for the counties grouped together for the ecological domain defined by ecoregion provinces 232, 234 and 411, and in ecoregion province 222. For counties in ecoregion province 231, the relationship was negative. For the counties in the other ecological domain consisting of ecoregion provinces 251, 255, 311, 315, 321, 331 and 332, the first-stage probit model was not significant, so that no IMR was included; for 221, the IMR from the first stage was not statistically significant in initial second-stage model estimates and was therefore dropped from the final specification. For lightning wildfires, land area of the county, included in four out of five lightning model second-stage specifications, had a positive effect on the likelihood that a county had non-zero lightning wildfire; that is, larger counties were more likely to have had at least one lightning fire occurrence. Forest area was included in only two models: counties for the ecological domain consisting of ecoregion provinces 251, 255, 311, 315, 321, 331 and 332, where it was positively related, and counties in ecoregion province 231, where it was negative but not significant (but where land area was positive).[5] Personal income per capita was positively related to non-zero lightning fire in counties of the ecological domain containing ecoregion provinces 232, 234 and 411, and counties in ecoregion province 221, but negatively for those in 222. When population density entered the second-stage model, its effect was negative. Weather variables had effects that were generally expected based on theory: precipitation was negatively related to non-zero lightning fires in most cases, whereas temperature and PET were both positively related. In most cases, these models were substantially simpler than the second-stage (and first-stage) selection equations, limited to a few climate and societal variables, and land use. In one case, namely the model of lightning wildfire area burned in counties of ecoregion province 222, the area burned was a function of just a constant and a statistically insignificant IMR from the second-stage equation. Land area, when it appeared, was positively related. In ecoregion province 231, the positive sign on land area was accompanied by a negative sign on forest area, a result that was not necessarily expected. The only societal variable affecting area burned other than forest area, a land-use measure, was population density, which appeared in the final-stage model for counties in the ecological domain containing ecoregion provinces 251, 255, 311, 315, 321, 331 and 332. Here, its effect was negative. This appears reasonable because more densely populated counties would have more breaks in the landscape that slow wildfire spread, so that wildfire sizes should be smaller, all other things being equal. For area burned in the final stage of lightning wildfire, weather variables were also related as expected from theory – precipitation, negatively, and temperature and PET, positively.

[5]In model development using in-sample data, 1992–2003, forest area was statistically significant. But when the final model was estimated using all data (1992–2010), its significance disappeared. We included it in this specification based on its in-sample model significance.

Table 4. Lightning-ignited wildfire summary statistical results (first-, second-, third-stage models), by ecoregion province
For each variable in each column, the three indicators refer to the first, second and third stages of the selection models. '+' implies a positive and '–' a negative parameter estimate that was significantly different from zero at 10% or stronger; '0' implies no statistically significant effect (or not included in the model). T is temperature in degrees Celsius, PET is potential evapotranspiration

Predictor variable	221	222	231	232, 234, 411	251, 255, 311, 315, 321, 331, 332
Weather variables					
January average max. T	−, 0, 0	0, 0, 0	−, 0, 0	−, 0, 0	0, 0, 0
March average max. T	+, 0, 0	0, 0, 0	+,−, 0	+, 0, 0	0, 0, 0
May average max. T	+, 0, 0	0, 0, 0	−, 0, 0	−, 0, 0	0, +, 0
July average max. T	0, 0, 0	0, 0, 0	+, 0, 0	+, 0, 0	0, 0, 0
September average max. T	−, 0, 0	−, 0, 0	−, 0, 0	+, −, 0	0, +, 0
January average PET	+, +, 0	−, 0, 0	+, +, +	+, +, +	0, 0, 0
March average PET	−, 0, 0	0, 0, 0	−, 0, +	−, 0, 0	0, 0, 0
May average PET	+, 0, 0	−, 0, 0	+, 0, 0	+, +, 0	0, 0, 0
July average PET	+, 0, 0	+, +, 0	−, 0, 0	−, 0, 0	0, 0, 0
September average PET	+, 0, 0	+, 0, 0	−, 0, 0	−, 0, 0	0, 0, 0
January average precipitation	−, 0, 0	0, 0, 0	+, 0, 0	+, 0, −	0, 0, 0
March average precipitation	+, 0, −	0, 0, 0	0, −, 0	+, 0, 0	0, 0, 0
May average precipitation	+, −, 0	0, 0, 0	−, −, 0	+, 0, −	0, 0, 0
July average precipitation	0, 0, 0	0, 0, 0	+, −, 0	+, 0, 0	0, 0, 0
September average precipitation	+, −, 0	+, 0, 0	0, +, 0	0, 0, 0	0, 0, 0
Biophysical variables					
Land area	+, +, 0	0, +, 0	−, +, +	0, +, +	0, 0, +
Forest area	0, 0, 0	−, 0, 0	+, −, −	0, 0, 0	0, +, 0
Socioeconomic variables					
Population	0, 0, 0	+, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0
Population density	+, −, 0	0, 0, 0	0, −, 0	0, 0, 0	0, −, −
Personal income per capita	−, +, 0	−, −, 0	0, 0, 0	0, +, 0	0, 0, 0
Categorical variables					
Alabama	0, 0, 0	−, 0, 0	+, −, 0	0, 0, 0	0, 0, 0
Arkansas	0, 0, 0	−, 0, 0	+, +, 0	0, 0, 0	0, 0, 0
Florida	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0
Georgia	+, 0, 0	0, 0, 0	+, 0, 0	+, 0, 0	0, 0, 0
Kentucky	+, 0, 0	−, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0
Louisiana	0, 0, 0	0, 0, 0	+, −, 0	+, 0, 0	0, 0, 0
Mississippi	0, 0, 0	0, 0, 0	+, 0, 0	−, 0, 0	0, 0, 0
North Carolina	−, 0, 0	0, 0, 0	0, 0, 0	−, 0, 0	0, 0, 0
Oklahoma	0, 0, 0	−, 0, 0	+, 0, 0	0, 0, 0	0, 0, 0
South Carolina	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0
Tennessee	+, −, 0	0, 0, 0	+, 0, 0	0, 0, 0	0, 0, 0
Texas	0, 0, 0	0, 0, 0	+, 0, 0	−, 0, 0	0, 0, 0
Virginia	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0
Province 234	0, 0, 0	0, 0, 0	0, 0, 0	−, −, 0	0, 0, 0
Province 411	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0
Province 255	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0	0, −, 0
Inverse Mills ratios					
Second stage	0	+	–	+	0
Third stage	–	0	0	+	–

For human-ignited wildfires (Table 5), only one second-stage model included land area, consisting of the counties of the ecological domain containing ecoregion provinces 232, 234 and 411, and its effect was positive. Forest area was included in the other second-stage models, and its effect was positive. Personal income per capita had varying effects, including a positive effect for ecoregion provinces 232, 234 and 411, and a negative effect for counties in 221 and 231. Population density had a positive effect on the probability of non-zero human-ignited wildfire for counties in 251, 255, 311, 315, 321, 331 and 332. Population itself, however, had a generally positive effect for counties whose models included it: ecoregion provinces 222 and 231. This makes sense because higher population density would mean that there are fewer people in close proximity to wildlands, while higher populations should be connected to more chances for wildfires to be accidentally ignited through any number of specific causes (e.g. debris-burning, children, campfires, equipment). In nearly all cases, weather variables had the expected

Table 5. Human-ignited wildfire summary statistical results province (first-, second-, third-stage models), by ecoregion province
For each variable in each column, the three indicators refer to the first, second and third stages of the selection models. '+' implies a positive and '–' a negative parameter estimate that was significantly different from zero at 10% or stronger; '0' implies no statistically significant effect (or not included in the model). T is temperature in degrees Celsius, PET is potential evapotranspiration

Predictor variable	221	222	231	232, 234, 411	251, 255, 311, 315, 321, 331, 332
Weather variables					
January average max. T	−, 0, 0	0, 0, 0	−, 0, 0	−, 0, 0	0, 0, 0
March average max. T	+, 0, 0	0, 0, 0	+, 0, 0	+, 0, 0	0, 0, 0
May average max. T	+, 0, 0	0, 0, 0	−, 0, 0	−, 0, 0	0, 0, 0
July average max. T	0, 0, 0	0, 0, +	+, 0, 0	+, 0, 0	0, 0, 0
September average max. T	−, 0, 0	−, 0, 0	−, 0, 0	+, 0, −	0, +, 0
January average PET	+, 0, 0	−, +, 0	+, 0, +	+, 0, +	0, 0, 0
March average PET	−, 0, 0	0, 0, +	−, 0, 0	−, 0, 0	0, 0, 0
May average PET	+, 0, 0	−, 0, 0	+, 0, 0	+, 0, 0	0, 0, 0
July average PET	+, 0, 0	+, 0, 0	−, 0, 0	−, 0, 0	0, 0, 0
September average PET	+, 0, 0	+, 0, 0	−, 0, 0	−, 0, 0	0, 0, 0
January average precipitation	−, 0, 0	0, 0, 0	+, 0, 0	+, 0, 0	0, 0, 0
March average precipitation	+, 0, 0	0, 0, −	0, −, −	+, 0, 0	0, 0, −
May average precipitation	+, −, 0	0, 0, 0	−, 0, 0	+, 0, 0	0, 0, 0
July average precipitation	0, 0, 0	0, 0, 0	+, 0, 0	+, 0, 0	0, 0, 0
September average precipitation	+, 0, 0	+, −, 0	0, 0, 0	0, −, 0	0, 0, 0
Biophysical variables					
Land area	+, 0, 0	0, 0, 0	−, 0, 0	0, 0, 0	0, 0, 0
Forest area	0, +, +	−, +, +	+, +, 0	0, +, 0	0, +, +
Socioeconomic variables					
Population	0, 0, 0	+, −, 0	0, +, +	0, +, +	0, +, 0
Population density	+, 0, 0	0, 0, 0	0, 0, −	0, 0, −	0, −, 0
Personal income per capita	−, −, 0	−, 0, −	0, −, −	0, 0, 0	0, 0, 0
Categorical variables					
Alabama	0, 0, 0	−, 0, 0	+, +, +	0, 0, 0	0, 0, 0
Arkansas	0, 0, 0	−, 0, 0	+, 0, 0	0, 0, 0	0, 0, 0
Florida	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0
Georgia	+, 0, 0	0, 0, 0	+, 0, 0	+, 0, 0	0, 0, 0
Kentucky	+, 0, +	−, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0
Louisiana	0, 0, 0	0, 0, 0	+, 0, 0	+, 0, 0	0, 0, 0
Mississippi	0, 0, 0	0, 0, 0	+, 0, +	−, 0, 0	0, 0, 0
North Carolina	−, 0, 0	0, 0, 0	0, 0, 0	−, 0, 0	0, 0, 0
Oklahoma	0, 0, 0	−, 0, 0	+, 0, +	0, 0, 0	0, −, 0
South Carolina	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0
Tennessee	+, 0, 0	0, 0, 0	+, 0, 0	0, 0, 0	0, 0, 0
Texas	0, 0, 0	0, 0, 0	+, 0, +	−, 0, 0	0, 0, 0
Virginia	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0
Province 234	0, 0, 0	0, 0, 0	0, 0, 0	−, −, 0	0, 0, 0
Province 411	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, +	0, 0, 0
Province 255	0, 0, 0	0, 0, 0	0, 0, 0	0, 0,	0, 0, 0
Inverse Mills ratios					
Second stage	+	−	+	+	0
Third stage	−	+	+	0	+

signs in these second-stage model estimates. Final-stage area burned models for human-ignited wildfires, like lightning wildfires, also demonstrated a tendency (except for ecoregion province 231) to be simpler, including fewer predictor variables.

Land area and forest area were both positively related to area burned. This indicates that a prediction of forest area loss would be connected to a decline in the area burned by human-ignited wildfires. In these final-stage equations, population density had a negative effect on the area burned by human-caused wildfires in counties of the ecological domain containing ecoregion provinces 251, 255, 311, 315, 321, 331 and 332, and those of ecoregion province 231. The finding on population density is consistent with the idea that a greater density of humans is linked to more fuel breaks in the landscape and greater opportunities for wildfire suppression resources to access and extinguish active fires. In contrast, higher overall populations in counties of ecoregion province 222 were linked to more area burned by human-caused wildfires. This is logical because more people are in contact with the wildland, creating more opportunities to ignite wildfires, but also have them recorded by fire management agencies. Personal income per capita was found to be negatively related to area burned for counties of ecoregion province 222 and 231. The negative relationship makes sense from an economics perspective: locations with higher wealth generally have greater financial resources available for fire suppression and prevention, leading to smaller overall wildfires. Such locations also typically have greater values at risk, which would compel greater investments in suppression and prevention. Finally, the IMRs from the first and second stage were usually statistically significant and positively related to area burned, indicating that places with more likely non-zero wildfire tend to also have larger wildfires, all other things being equal, which, in turn, indicates the importance of accounting for this tendency.

Monte Carlo simulation of projected wildfire futures

While individual iteration data for each county and each of the 2250 iterations performed in this projection of the annual area burned future of wildfire in the Southeast are available from the authors on request,[6] Appendix S2, available as online supplementary material, describes these graphically. In Appendix S2, charts show projected annual area burned over the projection period for each state and each ecosystem province-based ecological domain. For a concise description, Figs 2–4 display the Southeast-wide summary for lightning and human causes and for total wildfire respectively.

The results of Monte Carlo simulations are probability distributions of projected wildfire in the Southeast, given our modelling and projection assumptions, and it is from these that we extract median and quantile information in our discussion of the projection results. Our results show that, for the projection period after 2010, there is a wide range in potential outcomes consistent with the historical data and our chosen scenarios of projected climate and societal futures. This is apparent for the extent of both lightning- and human-ignited wildfire area burned in the south-eastern US. However, from 2011 on, it is apparent that the expected annual area burned by lightning rises over time.[7] This increase is projected in percentage terms for every state and ecosystem province (Table 6). The average of the Monte Carlo median area burned by lightning rises by an average of ~616 ha per year across the entire region from the 2016–20 average of the Monte Carlo median to the 2056–60 average of the Monte Carlo median, a 34% rise.[8] For human-ignited wildfires, the picture is different. The average of the Monte Carlo median area burned declines 373 ha per year across the entire region, a 6% fall. Human-ignited wildfires represent the majority of annual area burned in the region. Projections indicate that their human share will drop from 76% of all area burned in the 2016–20 period to ~69% by 2056–60. The net effect of these changes across the Southeast in total is that expected annual area burned by wildfires

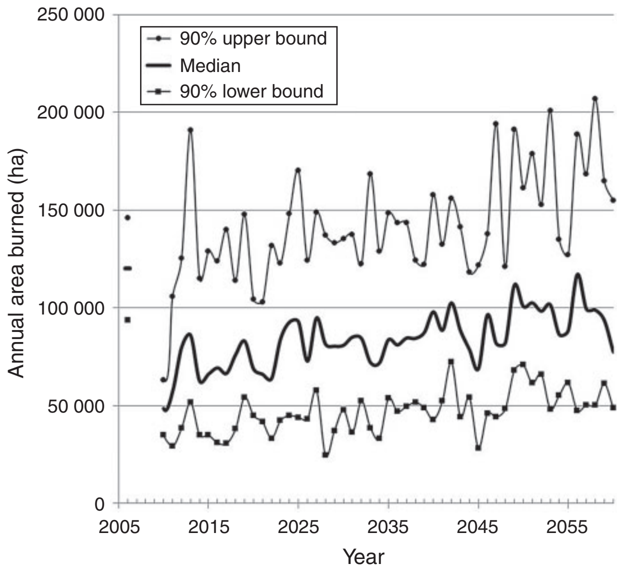

Fig. 2. Projections of lightning-ignited wildfires, 2006 and 2010–60, for the south-eastern US, in aggregate (i.e. the sum of wildfire for all counties in the region), including upper and lower 90% bounds of 2250 Monte Carlo iterations of models under all scenario and general circulation model (GCM) realisations. (Note: No projections were made for 2005, 2007, 2008 or 2009.)

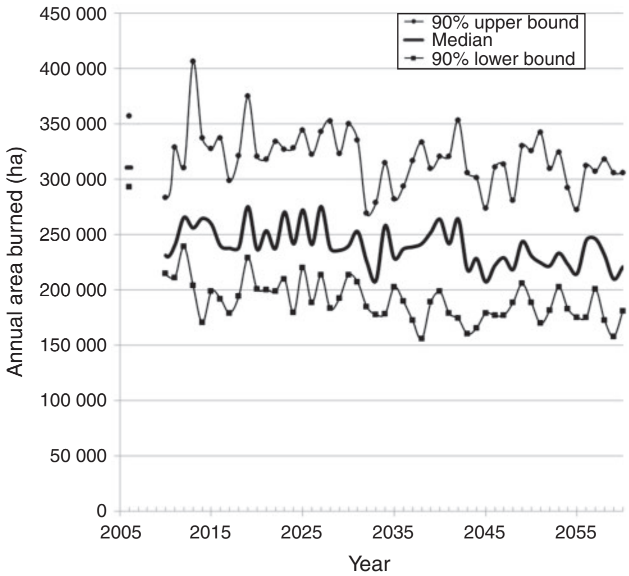

Fig. 3. Projections of human-ignited wildfires, 2006 and 2010–60, for the south-eastern US, in aggregate (i.e. the sum of area burned for all counties in the region), including upper and lower 90% bounds of 2250 Monte Carlo iterations of models under all scenario and general circulation model (GCM) realisations. (Note: No projections were made for 2005, 2007, 2008 or 2009.)

[6]That is, data can be summarised at the county level for any of the nine emission scenario–GCM combinations. Results generally indicate substantial ranges in potential futures, with the A2 scenario combinations generating higher levels of wildfire and the B2 the lowest amounts, although these trends vary across space.

[7]The higher level of predicted area burned for 2006 in these figures is due to the unusually fire-prone conditions found in many parts of the Southeast that year, especially from lightning and in states along the Atlantic coast and much of the Gulf of Mexico coast, compared with 2011–15 projections.

[8]We compared the average of the Monte Carlo median area burned for 2056–60 with the average of the Monte Carlo median area burned projected for 2016–20 rather than with a historical (pre-2011) period because of missing data in the pre-2011 period.

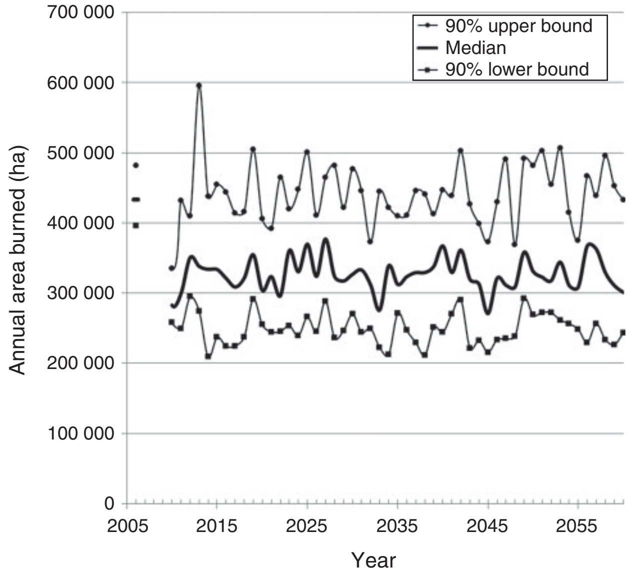

Fig. 4. Projections of all wildfires combined, 2006 and 2010–60, for the south-eastern US, in aggregate (i.e. the sum of area burned for all counties in the region), including upper and lower 90% bounds of 2250 Monte Carlo iterations of models under all scenario and general circulation model (GCM) realisations. (Note: No projections were made for 2005, 2007, 2008 or 2009.)

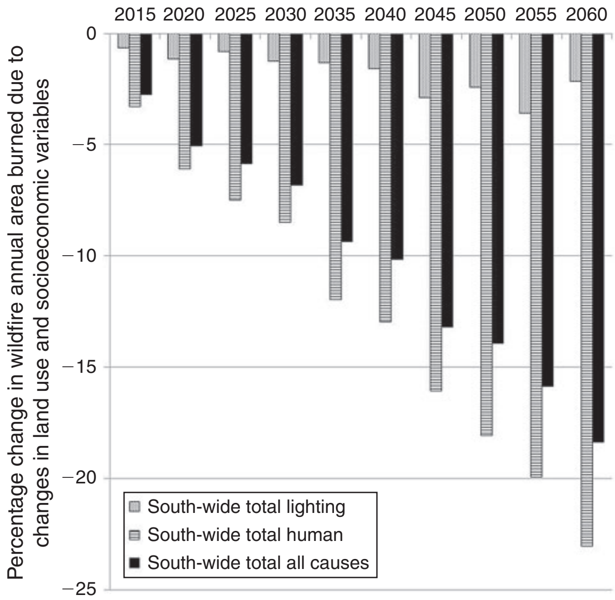

Fig. 5. Projected effects of changes, in percentage, in wildfire annual area burned in the Southeast owing to changes in land use and other socioeconomic variables, mean effect across all scenario–GCM combinations, in 5-year increments, in aggregate (i.e. the sum of area burned for all counties in the region), 2015–60.

of all causes is projected to rise by ~4% by 2056–60 compared with 2016–20.

Such changes at the regional level mask more substantial, and sometimes contrasting projected changes for geographical subsets of the region. For example, Florida and Louisiana are projected to see median wildfire area burned rise by 15 and 30% respectively. Other states with smaller expected rises include Georgia, Mississippi, North Carolina, Oklahoma, South Carolina and Texas. States with projected double-digit percentage declines include Arkansas, Kentucky, Tennessee and Virginia. Ecoregion provinces 221, 222 and 231 are each projected to have double-digit percentage declines in median annual area burned, but double-digit percentage increases in expected annual area burned are projected for ecoregion provinces 232, 234 and 411. It should be emphasised that great uncertainties exist in the likelihood of projected changes, even in the context of our modelling assumptions. The 90% upper and lower bounds, documented in Figs 2–4, and those for individual states and ecoregion provinces available in Appendix S2, are indicators that definitive conclusions cannot be made about the ultimate changes in wildfire in any state or ecoregion province.

The results suggest that changes in society have a large impact on human-ignited wildfires and a smaller impact on lightning-ignited wildfires. We summarise the net effects of the projected net reductions in forestland, and increases in income and human population across the region in Fig. 5. This figure was constructed by: (1) projecting wildfire allowing only changes in the climate variables from all nine climate realisations while holding land use, income and population constant at levels observed in 2006; and (2) subtracting the mean projection from the mean projections reported in Figs 2–4 that do allow land use, income and population to change. Fig. 5 therefore demonstrates the net effects of the land-use, income and population variables on our projections. The figure shows that land-use and socioeconomic variables have a minimal effect on projected lightning-ignited fire areas; changes in these variables reduce projected lightning wildfire area burned by only ~2% by 2060 compared with the simulation with no change in land-use and socioeconomic variables. For human-ignited wildfire, the effects are substantially greater: annual area burned by human-ignited wildfires is projected to be 23% lower in 2060 when land-use and socioeconomic variables are allowed to change, compared with holding those variables constant; conversely, ignoring land-use and socioeconomic variable changes results in a projected median that is 30% higher than our median projection. In total, the effect of land use, income and population changes is to reduce wildfire by all causes by ~18% in 2060 compared with what it would be if land-use and socioeconomic changes did not occur. In summary, climate change, on average, is projected to push up annual area burned in the Southeast, but changes in the distribution of fuels and other factors that are related to income and population growth serve to counteract most of the effects of climate change on area burned, largely through reductions in the annual area burned by human-ignited wildfires.

Discussion

Modelling of wildfire in the south-eastern US, to be broadly useful, requires information on whether historical data are valid. Selection models can help compensate for (a lack of) valid data. Such projections should also account for the uncertainties in both the statistical models as well as future states of nature. Finally, wildfire, particularly in the south-eastern US, is dominated by humans, who ignite the majority of wildfires and suppress nearly all of them. Humans break up landscapes, thus limiting fire spread, and, with their roads and other infrastructure, facilitate the

Table 6. Summary of projected expected percentage changes in annual areas burned by wildfire by lightning and human causes and in total, 2056–60 compared with 2016–20

Provinces 315, 321, 331 and 332 were represented by very few counties in the spatial domain of inference, and so their wildfire changes were not separately tracked in the modelling

Geographic unit	Lightning wildfire annual area burn change	Human wildfire annual area burn change	Total wildfire annual area burn change
	(% change, 2056–60 average compared with 2016–20 average)		
Alabama	39	−12	−3
Arkansas	60	−59	−31
Florida	11	16	16
Georgia	33	−7	6
Kentucky	189	−28	−22
Louisiana	57	8	30
Mississippi	85	−2	8
North Carolina	74	−10	4
Oklahoma	43	−1	3
South Carolina	34	−7	5
Tennessee	173	−62	−52
Texas	26	−5	4
Virginia	164	−50	−26
Ecoregion province 221	218	−41	−29
Ecoregion province 222	8	−88	−78
Ecoregion province 231	55	−20	−10
Ecoregion province 232	22	10	20
Ecoregion province 234	151	34	82
Ecoregion province 251	37	3	4
Ecoregion province 255	31	−5	−1
Ecoregion province 311	52	2	6
Ecoregion province 411	5	22	17
Ecoregion provinces 315, 321, 331, 332	27	−3	3
Southeast	34	−6	4

suppression of wildfires. Our modelling results found that counties in the Southeast tend to have less area burned in places with higher personal income per capita. Furthermore, dense populations and lower areas of forest are related to less overall wildfire area burned. However, larger populations may be connected to more fires ignited, and therefore the net effect of rising population density and rising overall population numbers is an empirical question.

The scenarios defined in the 2010 RPA Assessment projected rising populations, rising population densities, rising incomes and falling forest area for the region. These effects would tend to favour less wildfire in the region, at least for fires ignited by humans. However, based on our model estimation results (Table 4 and Table 5), warming temperatures tend to favour more wildfire. Additionally, precipitation patterns could also change. The net effect of these climate variables, our results show, is therefore especially to force up lightning-caused wildfire. In our simulations, we would expect more lightning-caused wildfires in locations with already plentiful lightning-ignited wildfires – especially along the Gulf of Mexico and Atlantic coasts. Human-ignited wildfires, especially in interior portions of the Southeast, are projected to decline in most states owing to loss of forest, and in relation to rising incomes. Overall, the median of our wildfire area projections indicates that the Southeast would not experience a large net increase in annual area burned. However, considerable uncertainty remains in how both society and nature and, therefore, wildfire will evolve in the coming decades. Our results show that there is ample scope for either increases or decreases in annual wildfire area burned when viewed Southeast-wide. Hence, wildland fire managers and environmental policy-makers concerned with how wildfire may affect the costs and losses accruing from wildfire and its management, as well as the air quality implications of such changes, would do well to plan for either eventuality. Our simulations provide context for the possible ranges of these changes.

Although state-level policies regarding landowner liability have shifted to favour increased use of prescribed fire (Yoder *et al.* 2004), concerns about air quality and other factors (e.g. Liu *et al.* 2009; Quinn-Davidson and Varner 2012) have led to the possibility of growing restrictions on its use. It is enlightening to consider how the net effect of such changes, not modelled in our study, may affect wildfire in the Southeast. Mercer *et al.* (2007) found that a 1% increase in prescribed fire led to ~0.23% long-run decrease in the annual area burned by wildfire in Florida, which lies primarily in ecoregion province 232, a state with large amounts of prescribed fire. The median annual area burned for ecoregion province 232 is projected in our models to rise from ~99 000 ha in 2016–20 to 120 000 ha in 2056–60, or by ~21.6% (Table 6). If prescribed fire in all of ecoregion province 232 were to decline by 25%, over and above any changes forced by a changing climate, an application of the estimate of Mercer *et al.* (2007) would project the ecoregion province annual area burned to rise by an additional 6900 ha above our median projection for the ecoregion province, to ~126 900 ha. This represents a rise of an additional 7 percentage points (to ~28.6%) compared to the 2016–20 average.

A sensitivity analysis that examined the impacts of changes in land use and other socioeconomic factors on wildfire showed that these trends in the south-eastern US should not be ignored when seeking to understand how wildfire could change into the future. Economic and population growth in the Southeast leads directly to forest loss because humans require land to build houses (Wear 2013), requiring new investments in fire suppression resources by states and municipalities, and involving the construction of new roads, which create breaks in fuels and allow easier fire suppression access. For example, from 2006 to 2060, under scenario A1B, population is projected to increase by 54% and forestland use to decrease by 10% in the counties covered in the current study. Similarly, increasing personal income, projected to rise by 252% under the A1B scenario, and rising population imply greater densities of values at risk on the landscape, which are a focus of fire suppression. In summary, we find that ignoring changes in land-use and socioeconomic factors that have a statistically significant effect on wildfire would overestimate area burned by nearly one-fifth, and the effect is particularly acute for human-ignited wildfires.

In contrast, rising population in the face of nearly stable wildfire area burned implies rising human exposure to wildfire when viewed in total across the region. These changes in the exposure are projected to vary widely across the counties of the region, in step with widely varying changes in wildfire and population. Nevertheless, the consequences of greater exposure to emissions of ozone and particulate species from wildfire include an expected rise in their human morbidity and mortality impacts (Rappold *et al.* 2011; Fann *et al.* 2013). This result implies that there could be significant social and economic benefits to management approaches that are shown to reduce the occurrence of wildfires in the region, including wildfire prevention and fuels management (e.g. Mercer *et al.* 2007; Butry *et al.* 2010).

Finally, these annual burned area projections could be a valuable dataset for informing broader assessments, from air quality to exposure to health burden, over the next few decades in the Southeast. There could be significant consequences for the region's future air quality and the human morbidity and mortality associated with wildfire activity changes (e.g. Viswanathan *et al.* 2006; Johnston *et al.* 2011). With projections of both climate and wildfire activity, models for studying the air quality impacts of anthropogenic and natural emissions that can include those from wildfire (e.g. the Community Multiscale Air Quality model (Byun and Schere 2006)) can provide additional insights into how a changing climate and society may in turn have health consequences for humans.

Conclusions

Flannigan *et al.* (2009) described the need for new studies that account for the complex, non-linear interactions between climate, fuels and humans if we hope to understand the implications of climate on future wildfire. These authors also highlighted the importance of improved data on wildfire occurrence so that prospective analyses can be done for many parts of the world. Our study addresses the first need, bringing in variables that are surrogates for the many roles that humans have in igniting, suppressing and preventing wildfire, managing fuels and changing aggregate fuels through land-use shifts; in the process, it also addresses the second need by compensating for the obvious data-quality problems that face analysts focussed on the US.

This study has also directly addressed some of the uncertainties that arise when considering the future of wildfire. By combining nine alternative views of both climate and society, generating probability distributions of area burned, we can begin to appreciate the extent of wildfire uncertainty facing society in the south-eastern US. Such uncertainty implies that wildfire managers and other decision-makers need to prepare for a possibly wide-ranging direction of change in wildfire processes in the region. Although some of this uncertainty is connected to bounded confidence of statistical models of wildfire production, much is due to uncertainty about how driving climate and societal variables will change over the coming decades in this region. In future studies, uncertainties could be reduced first by the development of more accurate wildfire production models than the ones we estimated. As newer climate projections are produced with improved GCMs and RCPs for greenhouse gas climate forcing, and as scientists and policymakers gain a better picture of our emissions, economic, population and policy futures, greater prediction accuracy as well as precision could be achieved. This enhanced projection environment will enable land managers and policymakers to take more concrete actions to minimise some of the negative effects of altered wildfire patterns.

Acknowledgements

This research was prepared in part through funding provided to the University of North Carolina–Chapel Hill through joint venture agreement 11-JV-11330143–080. The authors would like to acknowledge the contributions of Scott Goodrick, Linda Joyce, Karen Short, David N. Wear, Ana Rappold and unnamed participants in the 13th International Wildland Fire Safety Summit & 4th Human Dimensions of Wildland Fire Conference, Boise, ID, 20–24 April 2015.

References

Bailey RG (1995) Description of the ecoregions of the United States (2nd edn). USDA Forest Service Miscellaneous Publication Number 1391. Washington, DC.

Butry DT (2003) Ecocnty dataset. Available at http://www.srs.fs.usda.gov/sustain/data/ecocnty/ [Verified 3 February 2015].

Butry DT, Mercer DE, Prestemon JP, Pye JM, Holmes TP (2001) What is the price of catastrophic wildfire? *Journal of Forestry* **99**(11), 9–17.

Butry DT, Prestemon JP, Abt KL, Sutphen R (2010) Economic optimisation of wildfire intervention activities. *International Journal of Wildland Fire* **19**, 659–672. doi:10.1071/WF09090

Byun D, Schere KL (2006) Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system. *Applied Mechanics Reviews* **59**, 51–77. doi:10.1115/1.2128636

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

Donovan GH, Rideout DB (2003) A reformulation of the Cost Plus Net Value Change (C+NVC) model of wildfire economics. *Forest Science* **49**, 318–323.

Fann N, Fulcher CM, Baker K (2013) The recent and future health burden of air pollution apportioned across US sectors. *Environmental Science & Technology* **47**, 3580–3589. doi:10.1021/ES304831Q

Flannigan MD, Krawchuk MA, de Groot WJ, Wotton BM, Gowman LM (2009) Implications of changing climate for global wildland fire. *International Journal of Wildland Fire* **18**, 483–507. doi:10.1071/WF08187

Greene WH (1992) 'Econometric analysis.' (Macmillan Publishing Company: New York, NY)

Heckman J (1979) Sample selection bias as a specification error. *Econometrica* **47**, 153–161. doi:10.2307/1912352

IPCC Working Group I Technical Support Unit (2015) 'Climate change 2013: the physical science basis; Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.' (Cambridge University Press: New York, NY). Available at www.ipcc.ch/report/ar5/wg1/ [Verified 7 December 2015].

Johnston F, Hanigan I, Henderson S, Morgan G, Bowman D (2011) Extreme air pollution events from bushfires and dust storms and their association with mortality in Sydney, Australia 1994–2007. *Environmental Research* **111**, 811–816. doi:10.1016/J.ENVRES.2011.05.007

Joyce LA, Price DT, Coulson DP, McKenney DW, Siltanen RM, Papadopol P, Lawrence K (2014) Projecting climate change in the United States: a technical document supporting the Forest Service RPA 2010 Assessment. USDA Forest Service, Rocky Mountain Research Station, General Technical Report RMRS-GTR-320. (Fort Collins, CO).

Knutti R, Sedláček J (2013) Robustness and uncertainties in the new CMIP5 climate model projections. *Nature Climate Change* **3**, 369–373. doi:10.1038/NCLIMATE1716

Linacre ET (1977) A simple formula for estimating evaporation rates in various climates using temperature data alone. *Agricultural Meteorology* **18**, 409–424. doi:10.1016/0002-1571(77)90007-3

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 Y, Goodrick S, Achtemeier G, Jackson WA, Qu JJ, Wang W (2009) Smoke incursions into urban areas: simulation of a Georgia prescribed burn. *International Journal of Wildland Fire* **18**, 336–348. doi:10.1071/WF08082

Liu Y, Prestemon JP, Goodrick S, Holmes TP, Stanturf JA, Vose JM, Sun G (2013) Future wildfire trends, impacts, and mitigation options in the southern United States. In 'Climate change adaptation and mitigation management options'. (Eds JM Vose, KD Klepzig) pp. 85–125. (CRC Press: New York, NY)

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

Mankiw NG, Weil DN (1989) The baby boom, the baby bust, and the housing market. *Regional Science and Urban Economics* **19**, 235–258. doi:10.1016/0166-0462(89)90005-7

Mercer DE, Prestemon JP (2005) Comparing production function models for wildfire risk analysis in the wildland–urban interface. *Forest Policy and Economics* **7**, 782–795. doi:10.1016/J.FORPOL.2005.03.003

Mercer DE, Prestemon JP, Butry DT, Pye JM (2007) Evaluating alternative prescribed burning policies to reduce net economic damages from wildfire. *American Journal of Agricultural Economics* **89**, 63–77. doi:10.1111/J.1467-8276.2007.00963.X

Nakicenovic N, Steward R (Eds) (2000) 'Special report on emissions scenarios: a special report of Working Group III of the Intergovernmental Panel on Climate Change.' (Cambridge University Press: Cambridge, UK). Available at http://www.grida.no/climate/ipcc/emission/index.htm [Verified 18 June 2015].

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

National Fire Protection Association (2015) NFPA 1720: Standard for the organization and deployment of fire suppression operations, emergency medical operations, and special operations to the public by volunteer fire departments, 2010 edn. Available at http://www.nfpa.org/codes-and-standards/standards-development-process/safer-act-grant/nfpa-1720 [Verified 11 May 2015].

Prestemon JP, Butry DT (2005) Time to burn: modeling wildland arson as an autoregressive crime function. *American Journal of Agricultural Economics* **87**, 756–770. doi:10.1111/J.1467-8276.2005.00760.X

Prestemon JP, Pye JM, Butry DT, Holmes TP, Mercer DE (2002) Understanding broad-scale wildfire risks in a human-dominated landscape. *Forest Science* **48**, 685–693.

Prestemon JP, Butry DT, Abt KL, Sutphen R (2010) Net benefits of wildfire prevention education efforts. *Forest Science* **56**, 181–192.

Prestemon JP, Hawbaker TJ, Bowden M, Carpenter J, Scranton S, Brooks MT, Sutphen R, Abt KL (2013) Wildfire ignitions: a review of the science and recommendations for empirical modeling. USDA Forest Service, Southern Research Station, General Technical Report SRS-171. (Asheville, NC)

Quinn-Davidson LN, Varner JM (2012) Impediments to prescribed fire across agency, landscape and manager: an example from northern California. *International Journal of Wildland Fire* **21**, 210–218. doi:10.1071/WF11017

Rappold A, Stone SL, Cascio WE, Neas LM, Kilaru VJ, Carraway MS, Szykman JJ, Ising A, Cleve WE, Meredith JT, Vaughan-Batten H, Deyneka L, Devlin RB (2011) Peat bog wildfire smoke exposure in rural North Carolina is associated with cardiopulmonary emergency department visits assessed through syndromic surveillance. *Environmental Health Perspectives* **119**, 1415–1420. doi:10.1289/EHP.1003206

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 (2015) Sources and implications of bias and uncertainty in a century of US wildfire activity data. *International Journal of Wildland Fire* **24**, 883–891.

Syphard AD, Keeley JE, Massada AB, Brennan TJ, Radeloff VC (2012) Housing arrangement and location determine the likelihood of housing loss due to wildfire. *PLoS One* **7**(3), e33954. doi:10.1371/JOURNAL.PONE.0033954

US Bureau of Economic Analysis (2013*a*) Interactive data. Available at http://www.bea.gov/itable/ [Verified 15 May 2015]

US Bureau of Economic Analysis (2013*b*) Current-dollar and 'real' GDP. Available at http://www.bea.gov/national/index.htm [Verified 4 November 2013]

US Census Bureau (2012) Population estimates. Available at http://www.census.gov/popest/data/historical/index.html [Verified 4 April 2012]

USDA Forest Service (2014*a*) 2010 RPA Assessment land use data (xlsx). Available at http://www.fs.fed.us/research/rpa/assessment/ [Verified 31 March 2014]

USDA Forest Service (2014*b*) 2010 RPA Assessment population and income data (xlsx). Available at http://www.fs.fed.us/research/rpa/assessment/ [Verified 31 March 2014]

Viswanathan S, Eria L, Diunugala N, Johnson J, McClean C (2006) An analysis of effects of San Diego wildfire on ambient air quality. *Journal of the Air & Waste Management Association* **56**, 56–67. doi:10.1080/10473289.2006.10464439

Wear DN (2013) Forecasts of land uses. In 'The Southern Forest Futures project: technical report.' (Eds DN Wear, JG Greis) USDA Forest Service, Southern Research Station, General Technical Report SRS-GTR-178, pp. 45–71. (Asheville, NC)

Yoder J, Engel D, Fuhlendorf S (2004) Liability, incentives, and prescribed fire for ecosystem management. *Frontiers in Ecology and the Environment* **2**, 361–366. doi:10.1890/1540-9295(2004)002[0361:LIAPFF]2.0.CO;2