Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

ARTICLE TYPE

Anthropogenic and meteorological effects on the counts and sizes of moderate and extreme wildfires

Elizabeth S. Lawler | **Benjamin A. Shaby**

Department of Statistics, Colorado State University, Colorado, USA

Correspondence
Elizabeth S. Lawler,
Email: liz.lawler@colostate.edu

Funding Information

This research was supported by the National Science Foundation, Grant Number: DMS 2001433

Abstract

The growing frequency and size of wildfires across the US necessitates accurate quantitative assessment of evolving wildfire behavior to predict risk from future extreme wildfires. We build a joint model of wildfire counts and burned areas, regressing key model parameters on climate and demographic covariates. We use extended generalized Pareto distributions to model the full distribution of burned areas, capturing both moderate and extreme sizes, while leveraging extreme value theory to focus particularly on the right tail. We model wildfire counts with a zero-inflated negative binomial model, and join the wildfire counts and burned areas sub-models using a temporally-varying shared random effect. Our model successfully captures the trends of wildfire counts and burned areas. By investigating the predictive power of different sets of covariates, we find that fire indices are better predictors of wildfire burned area behavior than individual climate covariates, whereas climate covariates are influential drivers of wildfire occurrence behavior.

KEYWORDS
Extended generalized Pareto, Extreme value, Hierarchical Bayesian, Spline

1 | INTRODUCTION

Wildfires have increased in frequency and size across the United States (US) (Dennison et al., 2014; Goss et al., 2020; Iglesias et al., 2022). In the last few years alone, the world has seen an unprecedented speed, intensity, and breadth in wildfires (Duane et al., 2021), with the likelihood of large wildfires expected to increase over the long term (Stavros et al., 2014). As regions continue to develop risk mitigation, land management, and resource allocation policies, it is imperative to quantitatively assess the changing wildfire risk factors and be able to predict future extreme wildfire behavior at a national (and global) level (Goss et al., 2020). To address this problem, we build a joint model for wildfire counts and burned areas in the contiguous United States (CONUS), conditional on meteorological and demographic predictors. We model the entire distribution of burned areas to simultaneously capture extreme and moderate events, leveraging extreme value theory to focus on the right tail, since extreme burn events are the most critical for risk assessment. Our model borrows strength spatially using the concept of ecoregions,

which are geographic areas with common topographic features and vegetation types. We link the sub-models for wildfire counts and burned areas with shared random effects.

The proportion of total wildfire burned area in the Western US that can be attributed to the top 1% of the largest wildfires has been shown to range between 80–96% (Strauss et al., 1989), and as large fires increase in frequency across parts of the world, so too will the large proportion of burned area that they cause (Barbero et al., 2014). Climate models have demonstrated that climate change has and will continue to worsen wildfire weather and lengthen the wildfire season and is anticipated to cause a large increase in the number of expected wildfires through the end of the century (Liu et al., 2010; Martinuzzi et al., 2019; Pimont et al., 2022; Jones et al., 2022). With temperature extremes becoming more common due to climate change, there is a higher risk of seeing extreme wildfire seasons (Parente et al., 2018), particularly like those seen with the Australian wildfires of 2019–2020 (Van Oldenborgh et al., 2021).

When considering the effects of climate change, climate scientists have stressed the importance of studying and understanding the relationships among climate, climate extremes, humans, and wildfire activity to appropriately predict wildfire activity and protect against fires' risks (Jones et al., 2022). Early studies suggest that human factors have led to more than half of the observed increase in wildfire weather and danger (Abatzoglou & Williams, 2016; Williams et al., 2019; Barbero et al., 2020), suggesting that anthroprogenic covariates should be incorporated in wildfire modeling. And while humans have impacted climate change, their presence has also modified wildfire ignition patterns and landscapes so that, when coupled with extreme weather conditions in areas with wildfire fuel sources, there is an increased probability of wildfire ignitions (Pausas & Keeley, 2021; Hawbaker et al., 2013).

Like some previous studies assessing wildfire risk, we use ecoregions to pool information across space according to characteristics that are directly relevant to wildfire behavior (unlike, e.g. political boundaries). Ecoregions provide a convenient way to incorporate differences in covariates along with differences in soil, vegetation, and land-use (Hawbaker et al., 2013; Littell et al., 2009). The feedback process between wildfires and the vegetation they influence further demonstrates the need to look at wildfire patterns by ecoregions (Malamud et al., 2005).

Given this multifaceted nature of the occurrence and burned area of wildfires, we endeavor to tackle this problem by incorporating climate, anthropogenic, and wildfire risk covariates in a joint model that captures zero-inflated counts along with moderate and extreme sized burned areas.

Many previous works have studied the distribution of wildfires. Joseph et al. (2019) also use ecoregions as their areal units, independently modeling wildfire counts using zero-inflated models and burned areas using a variety of response distributions. However, while their interpretation focuses on extremes, they eschew established best practices by fitting a generalized Pareto distribution to the entire burn size dataset, rather than just the tail, and conversely draw inferences about the tail based on distributions that are not suitable for extremes. Opitz et al. (2020) incorporate a log-Gaussian Cox process model into a Bayesian

framework to regress wildfires on a set of land use and weather covariates using Laplace approximations. Pimont et al. (2021) develop a two-component flexible Bayesian model called 'firelihood' to fit the full distribution of wildfire burned areas and wildfire counts. Their approach requires categorizing the wildfire burned areas into four size categories through logistic regression, then fitting Pareto distributions to each. Drawbacks to this approach are that the piecewise specification is discontinuous, requires many parameters, and does not allow sharing of covariate and wildfire information across model pieces. Castel-Clavera et al. (2022) extend 'firelihood' by adding a spatio-temporal component. Becker et al. (2022) use a joint model to look at the counts and sizes of wildfires in a Canadian province, restricting their study period to the months of April–November across four decades, and linking count and burned area sub-models using a shared random effect similar to our model. Koh et al. (2023) builds upon Opitz et al. (2020) and Pimont et al. (2021) by developing a two-component mixture model for wildfire sizes that allows sharing of information between the moderate and extreme components. Their approach requires selecting a threshold selection prior to model fitting. Cisneros et al. (2023) develop a complex four-stage approach using stochastic partial differential equations to model wildfires in the US, which does not allow for sharing of information across stages. Also somewhat related is Richards & Huser (2022), who propose a neural-network model with extreme quantile regression to examine burned fractions of spatial grid boxes in the CONUS. However, their model focuses only on predicting extreme wildfires and does not capture the full distribution of wildfires.

Our approach is novel in three main respects. First, it models the full distribution of wildfire burned areas with a continuous distribution that leverages extreme value theory, considering moderate wildfires while emphasizing inference on the tails, without requiring threshold selection. These models are taken from the literature but, to our knowledge, have not been previously applied in a hierarchical model for wildfire data. Second, we share information across space using an easily interpretable random effect that partitions the spatial domain by ecological features rather than the more typical political boundaries. This random effects structure reflects the geographical hierarchical nesting of smaller ecoregions within larger ecoregions within even larger ecoregions. Finally, we demonstrate the utility of two fire weather indices, which together capture somewhat orthogonal aspects of fire risk, as effective substitutes for more complicated climate and weather covariates for predicting wildfire burned areas.

2 | OVERVIEW OF UNIVARIATE EXTREME VALUE THEORY

We present a brief overview of threshold models in univariate extreme value theory (EVT) (Coles, 2001). In the context of wildfire burned area, denoted as the random variable Y below, the asymptotic characterization of the threshold exceedance model states that the amount of area burned, given that it exceeds a high threshold u, is well-approximated by generalized Pareto

distribution (GPD), as the threshold u gets large. That is,

$$P(Y > y \mid Y > u) \approx \bar{H}_{\sigma,\xi}(y),$$

where $\bar{H}_{\sigma,\xi}$ is the survival function of the GPD and is defined as

$$\bar{H}_{\sigma,\xi}(y) = \begin{cases} (1 + \xi(y-u)/\sigma)_+^{(-1/\xi)}, & \xi \neq 0 \\ \exp(-(y-u)/\sigma), & \xi = 0 \end{cases} \tag{1}$$

where $a_+ = \max(a, 0)$, $\xi \in \mathbb{R}$ is the shape parameter which controls the thickness of the right tail, and $\sigma > 0$ is the scale parameter. The GPD has support $\{y \geq u : 1 + \xi(x-u)/\sigma \geq 0\}$, so that if $\xi < 0$, the upper tail of the distribution is bounded. If $\xi = 0$, the GPD reduces to an exponential distribution, which has finite moments. If $\xi > 0$, the distribution has a heavy upper tail. Empirically, the distribution of wildfire burned area over the CONUS appears to have a heavy tail (see Figure 1), which is consistent with observations in the literature that a small number of fires cause most of the damage (e.g. Strauss et al., 1989). Therefore, we limit the scope of this research to the latter two cases of the GPD, where $\xi \geq 0$.

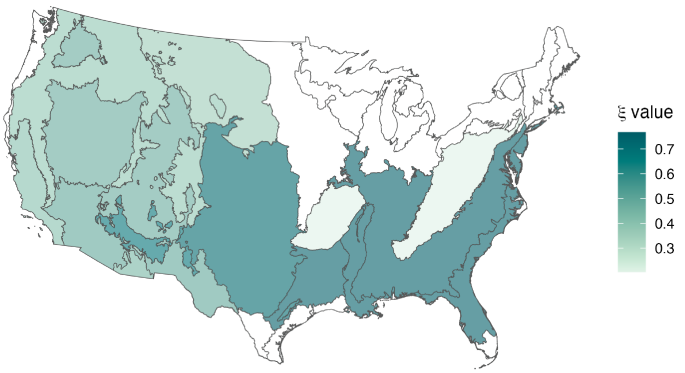

FIGURE 1 Map of maximum likelihood estimates (MLE) of ξ over the contiguous US, split by level 2 (L2) ecoregion. A univariate GPD was fit to every L2 ecoregion that had 20 or more burns at or above its respective 90th quantile. Shades correspond to the estimated shape (ξ) parameter. Higher values of ξ correspond to heavier tails.

In practice, implementation of the GPD model requires choosing a threshold u, subsequently fitting the GPD model only to those values in excess of u. In a standard extreme value analysis (EVA), where interest lies solely in the most extreme events, discarding observations below u is desirable; this "lets the tail speak for itself" and avoids bias that would occur by including moderate and small values in the fitting procedure. In the current analysis, however, moderate and small values are important, even if the extreme values are of primary interest—wildfire burned areas that would fall in the bulk of a distribution still cause

damage, socially and ecologically. We therefore seek a model that can accurately model the full distribution of fire sizes, with particular care taken to capture key features of the tail.

2.1 | Recent extensions of univariate EVT

Several recent works have proposed models for continuous random variables that preserve the properties of the GPD in the tail, while also modeling the full distribution. One approach is to "stitch" one or more non-extreme distributions to a GPD tail. A subset of these methods require threshold selection prior to model fitting (Castro-Camilo et al., 2019; MacDonald et al., 2011), while others allow for the threshold to be treated as an unknown parameter within the model (Frigessi et al., 2002; Behrens et al., 2004; Tancredi et al., 2006; Carreau & Bengio, 2009; do Nascimento et al., 2012). Another series of approaches models the bulk of the distribution while achieving GPD-like tails through the composition of cumulative distribution functions (cdfs), an idea introduced by Papastathopoulos & Tawn (2013). One flavor of this general strategy involves the composition of only two cdfs and is intended to model positive data where the right tail is of primary interest (Naveau et al., 2016; Gamet & Jalbert, 2022; Tencaliec et al., 2020). A more flexible version allows for the composition of more than two cdfs and can accommodate distributions with more general support and with GPD-like behavior in both the left and right tails, permitting detailed inference on both the very large and very small values (Stein, 2021b,a; Krock et al., 2022). Since our wildfire size data is strictly positive, and primary interest lies in the right tail, we use extended generalized Pareto distribution (EGPD) models from Naveau et al. (2016).

The EGPD (Papastathopoulos & Tawn, 2013; Naveau et al., 2016; Gamet & Jalbert, 2022) is motivated by the probability integral transform and takes the general form

$$F(y) = G\{H_{\sigma,\xi}(y)\}, \tag{2}$$

where $H_{\sigma,\xi}(\cdot)$ is the cdf corresponding to the survival function defined in (1), and $G(\cdot)$ is a continuous cdf with support on $[0, 1]$, which Naveau et al. (2016) call a *carrier function*. Naveau et al. (2016) propose a set of constraints on G to preserve the upper tail behavior of $H_{\sigma,\xi}$ and give the distribution power law behavior near 0.

These conditions result in a Pareto-like right tail and ensure that the lower portion of the distribution is driven by the carrier function $G(\cdot)$, which can depend on parameters that are estimated from the data. Naveau et al. (2016) propose four parametric

families of carrier functions that satisfy the specified constraints:

$$G_1(v) = v^{\kappa},\ \kappa > 0$$

$$G_2(v) = pv^{\kappa_1} + (1-p)v^{\kappa_2},\ \kappa_1, \kappa_2 > 0,\ p \in [0,1]$$

$$G_3(v) = 1 - Q_{\delta}\{(1-v)^{\delta}\},\ \delta > 0$$

$$G_4(v) = [G_3(v)]^{\kappa/2},\ \kappa > 0,$$

where Q_δ in the definition of $G_3(v)$ is a Beta cdf, with $\alpha = 1/\delta$ and $\beta = 2$. In our model, we follow de Carvalho et al. (2022) and build upon the EGPD by adding a regression link between covariates and the parameters of the carrier functions.

3 | UNITED STATES WILDFIRE DATA

3.1 | Wildfire counts and burned area data

We use wildfire data from the Monitory Trends in Burn Severity (MTBS) program from 1990–2020 (Eidenshink et al., 2007), yielding T=372 monthly timesteps. The MTBS provides a robust dataset including wildfire discovery date, wildfire occurrence location, and final burned area of the wildfire, across the entire CONUS. Fire occurrence location is determined by the geographic centroid of the delineated burned area boundary for each wildfire. Fires are included in the MTBS dataset if they are over 1,000 acres and 500 acres in the Western and Eastern United States, respectively. For consistency, we include only fires that are larger than 1,000 acres in our model. This necessitates truncated versions of EGPDs, which are easily accommodated in our model (see §4.3). To allow for better model mixing, we scale down the burned areas by a factor of 1,000, working in units of thousands of acres.

3.2 | Meteorological and anthroprogenic data

We include four meteorological covariates and one anthropogenic covariate in our model, broadly following the approach of Joseph et al. (2019). We use the surface meteorological data aggregated by gridMET (Abatzoglou, 2013), which is provided in daily timesteps at 4km resolution. The gridMET dataset contains daily relative minimum humidity, maximum air temperature, wind velocity, and precipitation accumulation, which we then aggregate to monthly timesteps, by taking the monthly mean for the first three variables and the monthly sum for precipitation. We then further aggregate precipitation on an annual scale to provide a measure of the previous year's accumulated precipitation, giving us five total meteorological variables. We use

housing density estimates captured from the 1990, 2000, 2010, and 2020 census surveys as a proxy for anthroprogenic drivers, provided by the SILVIS lab at the block group level (Mockrin et al., 2023).

3.3 | Regional data

The Environmental Protection Agency (EPA) provides a hierarchical framework of ecological regions (ecoregions) in North America, based on characteristics like vegetation, soil type, and land use (Omernik & Griffith, 2014). While the EPA provides four ecoregion levels (level one being the coarsest, level four the finest), we only use the first three levels due to computational limitations and data sparsity at the finest level. Within the contiguous US, there are S=84 total level three (L3) ecoregions, 20 L2, and 10 L1 (see Figure 2).

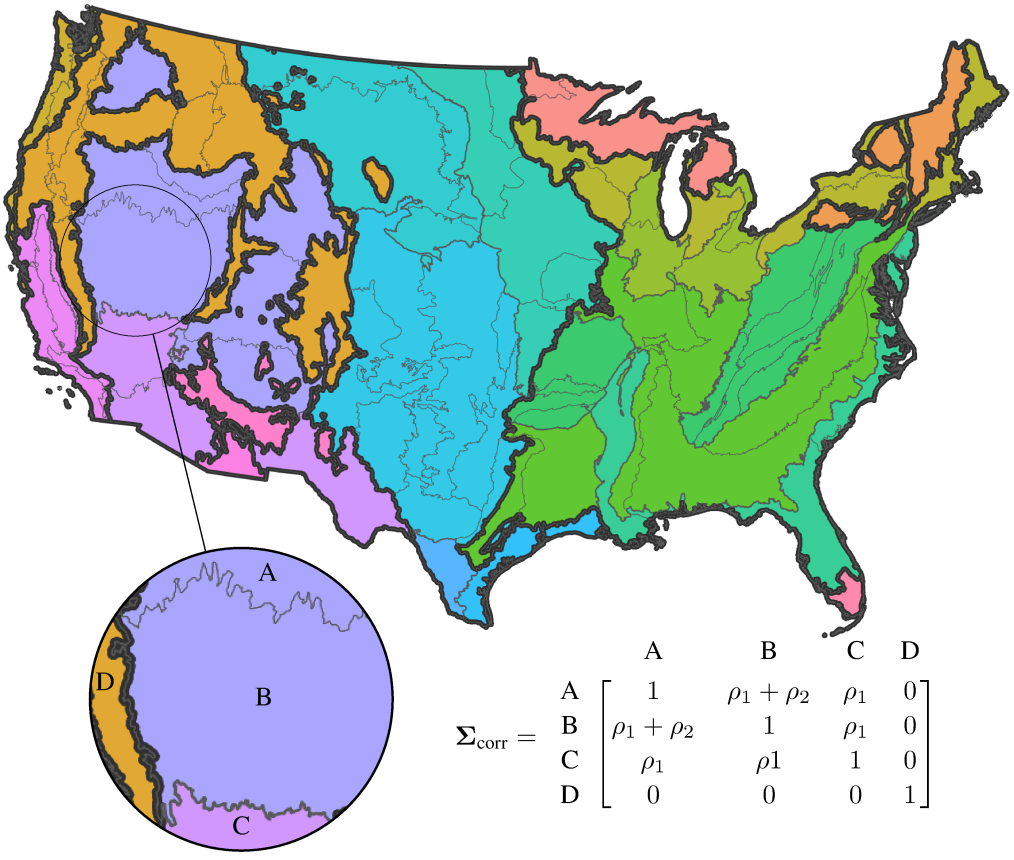

FIGURE 2 Hierarchical nesting of ecoregions in the contiguous United States. The EPA's L3 ecoregions are delineated by the thin light grey lines; L2 by the different colors; and L1 by the thick dark grey lines. A small section is magnified to illustrate the structure of the correlation matrix in the prior, as detailed in §4.1.

3.4 | Fire indices

There is a tradition in wildfire research to combine several (usually weather-related) variables into a single index that is intended to summarize wildfire risk (e.g. Fosberg, 1978). We were curious whether a univariate wildfire index had similar explanatory power in a counts model or an EGPD burned area model, relative to the combination of several weather variables. We therefore experimented with replacing the weather covariates with some combination of two indices, which are designed to capture different aspects of wildfire risk. Fire weather and fuel indices, like the Energy Release Component (ERC, Bradshaw et al., 1984) and Canadian Forest Fire Weather Index (FWI, Turner & Lawson, 1978) have shown conflicting results in their ability to predict extreme wildfires. Fire indices may capture the relationships and interactions between meteorological variables better than the variables alone and therefore be better able to predict their subsequent effects on wildfire activity (H. K. Preisler et al., 2008; Abatzoglou & Kolden, 2013). Abatzoglou & Kolden (2013) demonstrated that, when compared to individual weather components like temperature and precipitation, ERC had a more direct link to fuel flammability and conditions habitable to fires burning large areas. This suggests that ERC may provide more valuable information than the four weather covariates we initially fit our model on, particularly regarding burned area size. However, some wildfire indices (FWI specifically) consistently under-represent the occurrence of large wildfires and their associated burned areas (Pimont et al., 2021).

ERC is a calculated fuel moisture index from the National Fire Danger Rating System, which considers the cumulative drying effect of temperature, precipitation, humidity, and solar radiation as an overall measure of wildfire fuel (Abatzoglou & Kolden, 2013). The ERC is a relative index, meaning that an ERC value of 30 represents a potential heat release that is twice that of an ERC value of 15. We use the daily ERC data which is derived by gridMET under a conifer forest model (Abatzoglou, 2013). We then aggregate these data to a monthly timestep.

FWI is a wildfire danger index that incorporates the effects of fuel moisture and weather conditions into the projected behavior of a wildfire. This index is often used as the main covariate in wildfire models (see Abatzoglou & Williams, 2016; Castel-Clavera et al., 2022; Pimont et al., 2021). We use the `R` package `cffdrs` (Wang et al., 2017) to calculate the FWI on a daily timestep using the meteorological raster data described in §3.2. We then aggregate these data to a monthly timestep.

4 | MODEL

4.1 | Ecoregion nesting

In building our models for wildfire counts and burned areas, we leverage the concept of ecoregions to borrow strength across space. Rather than using generic spatially autoregressive areal models, we observe that drivers of spatial correlation in wildfire behavior might include things like common vegetation types, common topographical features, and so forth. Furthermore, the

ecoregions defined by EPA have an exploitable hierarchical structure, where the coarsest division is further divided into successively finer partitions, according to levels of ecological similarity. To account for and incorporate the hierarchical nature of the three levels of ecoregions, we build a 84×84 correlation matrix $\mathbf{\Sigma}_{\text{corr}}$, with each row and column corresponding to an L3 ecoregion. This matrix contains ones along the main diagonal (each L3 ecoregion is perfectly correlated with itself) and values between 0 and 1 elsewhere. Any two L3 ecoregions $\{i,j\}$ that are in the same L2 ecoregion will have the correlation coefficient ρ_2. This same pair of L3 ecoregions will also be in the same L1 ecoregion, therefore adding a correlation coefficient ρ_1, for a total correlation of $\rho_1 + \rho_2$ assigned to this pair of L3 ecoregions. Any two L3 ecoregions $\{i,j\}$ that are in the same L1, but not the same L2 ecoregion, will have the correlation coefficient ρ_1. L3 ecoregion pairs that do not share any nesting will have a value of 0. See Figure 2 for an illustration of a subset of this matrix. We use this spatial nested structure throughout our model-building process.

4.2 | Wildfire counts

We model wildfire counts as a zero-inflated (ZI) negative binomial (NB) random variables. This model defines the following likelihood for $n_{s,t}$, the number of wildfires in spatial unit s at time t, as

$$n_{s,t} \mid \lambda_{s,t}, \pi_{s,t}, \delta_s \sim \begin{cases} \pi_{s,t} + (1-\pi_{s,t}) \times \text{NB}(0 \mid \lambda_{s,t}, \delta_s) & \text{if } n_{s,t} = 0 \\ (1-\pi_{s,t}) \times \text{NB}(n_{s,t} \mid \lambda_{s,t}, \delta_s) & \text{if } n_{s,t} > 0, \end{cases}$$

where $s = 1, \ldots, S$ is the spatial unit (L3 ecoregion) and $t = 1, \ldots, T$ is the monthly timestep. Different levels of complexity are possible within this model. We tested versions of the model (see §5.1) where different parameters depended on covariates or not, where parameters varied spatially or not, and so forth. In the most general version, the rate parameter $\lambda_{s,t}$ and the probability of extra zeroes $\pi_{s,t}$ depend on spatially and temporally varying covariates, while the overdispersion parameter δ_s varies spatially according to the structured spatial prior described in §4.1. We also tried fixing $\delta_s = 0$ for all $s = 1, \ldots, S$, which yields a zero-inflated Poisson model. To ensure that $\delta_s, \lambda_{s,t} > 0$ and $\pi_{s,t} \in (0,1)$, we modeled these parameters with log and logit links, respectively, as

$$\begin{aligned} \log(\boldsymbol{\lambda}_s) &= \mathbf{X}_s \boldsymbol{\beta}_s^{(\lambda)} + \boldsymbol{\Phi}_s^{(\lambda)} + a_s \\ \text{logit}(\boldsymbol{\pi}_s) &= \mathbf{X}_s \boldsymbol{\beta}_s^{(\pi)} + \boldsymbol{\Phi}_s^{(\pi)} \\ \log(\boldsymbol{\delta}) &\sim \text{Normal}(\mathbf{0}, \mathbf{\Sigma}_{\text{corr}}^{(\delta)}), \end{aligned}$$

where $\boldsymbol{\lambda}_s$ and $\boldsymbol{\pi}_s$ are vectors of length T for L3 ecoregion s, $\boldsymbol{\delta}$ is a vector of length S, and a_s is a known offset equal to the area of the L3 ecoregion s. $\mathbf{X}_s$ is a $T \times (p+1)$ design matrix, $\boldsymbol{\beta}_s^{(\lambda)}$ and $\boldsymbol{\beta}_s^{(\pi)}$ are vectors of length $(p+1)$ with the priors discussed in §4.5, and $\boldsymbol{\Sigma}_{\text{corr}}^{(\delta)}$ is a structured correlation matrix described in §4.1. To account for residual structured variability, we include the spatio-temporal random effects $\boldsymbol{\Phi}^{(\lambda)}$ and $\boldsymbol{\Phi}^{(\pi)}$, which are $(T \times S)$ matrices, each composed of a spatial intrinsic conditional autoregressive component and a temporal autoregressive component, as

$$\begin{aligned} \boldsymbol{\phi}_{t=1}^{(\lambda)} &\sim \text{N}(\mathbf{0}, [\tau^{(\lambda)}(\mathbf{D}-\mathbf{W})]^{-1}) \\ \boldsymbol{\phi}_t^{(\lambda)} \mid \boldsymbol{\phi}_{t-1}^{(\lambda)} &\sim \text{N}(\eta^{(\lambda)}\boldsymbol{\phi}_{t-1}^{(\lambda)}, [\tau(\mathbf{D}-\mathbf{W})]^{-1}), \end{aligned} \tag{3}$$

and analogously for $\boldsymbol{\Phi}^{(\pi)}$, where $\boldsymbol{\phi}_t^{(\lambda)}$ is a row vector of length S. $\mathbf{D}$ is a fixed $(S \times S)$ diagonal matrix, with $D_{i,i}$ equal to the number of neighbors that region i has. The matrix $\mathbf{W}$ is a fixed $(S \times S)$ adjacency matrix, where $W_{i,j} = 1$ if regions i and j are neighbors and 0 otherwise. The parameters η and τ are given weakly informative priors

$$\begin{aligned} \eta^{(\lambda)} &\sim \text{Beta}(2, 8) \\ \tau^{(\lambda)} &\sim \text{Exp}(1). \end{aligned} \tag{4}$$

4.3 | Wildfire burned area

We model the extreme and non-extreme wildfire burned areas simultaneously using the family of EGPD models defined in §2.1. Taking, for example, the carrier function $G_1(v) = v^\kappa$, yields the cdf $F_1(y)$ and density $f_1(y)$

$$\begin{aligned} F_1(y) &= \{1 - [1 + \xi(\frac{y}{\sigma})]^{-1/\xi}\}^\kappa \\ f_1(y) &= \frac{\kappa}{\sigma}[1 + \xi(\frac{y}{\sigma})]^{-(1/\xi+1)}\{1 - [1 + \xi(\frac{y}{\sigma})]^{-1/\xi}\}^{\kappa-1} \end{aligned} \tag{5}$$

Since all of the fires in our dataset are greater than 1,000 acres, we construct a truncated density from (5), as

$$\text{t-EGPD}_1(x) := \frac{f_1(x)}{1 - F_1(1)}, \tag{6}$$

where we normalize by $1 - F_1(1)$ because x has units of thousands of acres. Equation (6) defines the following likelihood for y_{s_i,t_i}, the burned area of the ith wildfire event in spatial unit s at timestep t, where s and t are defined as in §4.2:

$$y_{s_i,t_i} \mid \kappa_{s_i,t_i}, \sigma_{s_i,t_i}, \xi_s \sim \text{t-EGPD}_1(y_{s_i,t_i} \mid \kappa_{s_i,t_i}, \sigma_{s_i,t_i}, \xi_s).$$

In most cases, $i = 1$, except when more than one wildfire occurs in the same spatial unit s in the same timestep t, therefore leading to multiple burned areas with otherwise identical s and t. Again, several levels of complexity are possible by allowing the EGPD parameters to depend on spatially and temporally varying covariates, to be spatially varying, to be constant across space, etc. For this example of the carrier function $G_1(\cdot)$, the most general model we considered allows κ, σ, and ξ to depend on spatially and temporally varying covariates, according to the same nested spatial prior described in §4.1. To ensure that all three parameters remained positive as required, we again use a log link, analogous to the links used in the count model (see §4.2), as

$$\log(\boldsymbol{\kappa}_s) = \mathbf{X}_s\boldsymbol{\beta}_s^{(\kappa)} + \boldsymbol{\Phi}_s^{(\kappa)}$$
$$\log(\boldsymbol{\sigma}_s) = \mathbf{X}_s\boldsymbol{\beta}_s^{(\sigma)} + \boldsymbol{\Phi}_s^{(\sigma)}$$
$$\log(\boldsymbol{\xi}_s) = \mathbf{X}_s\boldsymbol{\beta}_s^{(\xi)} + \boldsymbol{\Phi}_s^{(\xi)}$$

The priors on $\boldsymbol{\Phi}^{(\kappa)}$, $\boldsymbol{\Phi}^{(\sigma)}$, and $\boldsymbol{\Phi}^{(\xi)}$ are analogous to (3). The priors on the hyperparameters $\eta^{(\kappa)}$, $\eta^{(\sigma)}$, $\eta^{(\xi)}$, $\tau^{(\kappa)}$, $\tau^{(\sigma)}$, and $\tau^{(\xi)}$ are analogous to those in (4).

While every timepoint and spatial unit in the count model contains a value in the dataset (either with $n = 0$ or $n > 0$), this is not the case for the burned area model. To account for this, we treat any spatial unit and timepoint pair that does not have a wildfire as "missing" and model these missing data as latent parameters within the overall Bayesian model.

4.4 | Joint model

To share information between the wildfire counts and burned area sub-models, we link the two using a shared random effect. Coupling the two models has the potential to increase performance, but it moreover has interesting interpretive potential. For example, one conjecture that could be investigated using such a coupled model is that a region experiencing few fires may be more likely to experience very large fires. Similar coupled count and magnitude models for marked point processes have appeared recently (see Koh et al., 2023; Becker et al., 2022; Yadav et al., 2023).

We include a temporally varying shared random effect $\boldsymbol{\theta}$, a vector of length T, in the regression components of the model parameters as

$$\log(\boldsymbol{\lambda}_s) = \mathbf{X}_s\boldsymbol{\beta}_s^{(\lambda)} + \boldsymbol{\Phi}_s^{(\lambda)} + a_s + \boldsymbol{\theta}$$
$$\log(\boldsymbol{\kappa}_s) = \mathbf{X}_s\boldsymbol{\beta}_s^{(\kappa)} + \boldsymbol{\Phi}_s^{(\kappa)} + \gamma_s\boldsymbol{\theta}. \tag{7}$$

We include a spatially-varying link parameter γ_s to spatially modulate the relationship between the rate of counts and the distribution of wildfire sizes. To accommodate temporal correlation, the shared parameter $\boldsymbol{\theta}$ has an AR(1) prior, while the

modulating parameter $\boldsymbol{\gamma}$ is modeled as a spatial random effect, as

$$\theta_t \,|\, \theta_{t-1} \sim \mathrm{N}(\eta^{(\theta)}\theta_{t-1}, \epsilon^2)$$
$$\boldsymbol{\gamma} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}^{(\gamma)}_{\text{corr}}),$$

along with the weakly informative hyperprior $\epsilon \sim \mathrm{t}_4(0, 1)$.

4.5 | Nonlinear effects

To capture nonlinear relationships between our covariates and wildfire counts and burned areas across the large spatial domain (Brillinger et al., 2003; H. Preisler et al., 2004; H. K. Preisler & Westerling, 2007; Woolford et al., 2011, 2021), we introduce cubic B-splines with five basis functions for each of our covariates. To smooth the relationship among the basis functions within each covariate, we included a Gaussian auto-regressive of order 1 (AR(1)) prior for each coefficient vector (Lang & Brezger, 2004). This prior induces a penalty that is the Bayesian analogue of the classical penalized-spline approach of Eilers & Marx (1996).

In addition, it is advantageous to borrow strength across space for the covariate effects, which we accomplish using the spatial nesting scheme described in §4.1. This results in a Normal prior for each parameter vector $\boldsymbol{\beta}$ in the models defined above, centered around zero, with a covariance matrix constructed as the Kronecker product of an AR(1) covariance matrix and a nested spatial correlation matrix, as

$$\boldsymbol{\beta}^{(\lambda)}_s \sim \mathrm{N}(\mathbf{0}, \Sigma^{(\lambda)}_{\text{corr}} \otimes \Sigma^{(\lambda)}_{\text{AR(1)}}), \tag{8}$$

with analogous specifications for $\boldsymbol{\beta}^{(\pi)}_s$, $\boldsymbol{\beta}^{(\kappa)}_s$, $\boldsymbol{\beta}^{(\sigma)}_s$, and $\boldsymbol{\beta}^{(\xi)}_s$.

4.6 | Training and validation

We first assessed multiple candidate models for both the wildfire counts and wildfire burned areas, separately, on a dataset that included five climate covariates and housing density, to determine the best model for each component. Then, we determined the best collection of covariates (either the meteorological covariates or one or both of the wildfire indices) for each model component. Each model was trained on 21 years of data (1995–2015, inclusive) and tested on 10 years of data (1990–1994, 2016–2020), yielding 252 monthly timesteps in the training dataset and 120 monthly timesteps in the test dataset.

We estimate all parameters through Monte Carlo Markov Chain (MCMC) sampling in `Stan`, using the default NUTS sampler (Stan Development Team, 2023; Gabry et al., 2023). We run three chains for each model, with 2000 iterations per chain (1000 for warm-up, 1000 for sampling).

5 | RESULTS

5.1 | Model comparison

We developed multiple candidate models for both the wildfire counts and the wildfire burned area models. We then compared the various models on the test dataset using log-likelihood scores (log scores) for the counts and the burned areas, along with threshold-weighted continuous ranked probability scores (twCRPS) for the burned area models. We first give a brief description of these scores, and then we present the scores for the various candidate models.

Log scores are proper scoring rules based on the model's predictive density (Gneiting & Raftery, 2007). We calculate the log score at every MCMC iteration for each observation in the test dataset, and then sum all of the scores within each iteration. We then compare the median log scores across the iterations among each model. The CRPS (Matheson & Winkler, 1976) is another proper scoring rule, defined in terms of the probabilistic forecast (as opposed to the predictive density), and is sensitive to the distance between predictions and the true realizations. With $F(z)$ the predictive cdf associated with the probabilistic forecast and y the realization, the CRPS is defined as:

$$\text{CRPS}(F,y) = \int_{-\infty}^{\infty} (F(z) - \mathbf{1}\{z \geq y\})^2 \, dz. \tag{9}$$

Gneiting & Ranjan (2011) propose the twCRPS, which incorporates a weighting function to emphasize a specific region of interest while still being a proper scoring rule. With the same notation as (9) and weighting function $w(z)$, the twCRPS is defined as:

$$\text{twCRPS}(F,y) = \int_{-\infty}^{\infty} (F(z) - \mathbf{1}\{z \geq y\})^2 w(z) \, dz. \tag{10}$$

Since we are most interested in accurately capturing the upper tails of wildfire burned areas while still maintaining good predictive accuracy in the bulk of the burned area distribution, we use a Normal cdf as our weight function, with a mean of 21 thousand acres and standard deviation of nine thousand acres. This weight function is centered around the 90th quantile of the burned area sizes, thereby emphasizing fires above the 90th quantile while still giving some weight to fires below the 90th quantile. We denote this weight function as $\Phi_{21,9}(z)$. We use Monte Carlo approximation to evaluate the twCRPS through the discretization

$$\begin{aligned} \text{twCRPS}(F,y) &\approx \frac{\max\{y\} - \min\{y\}}{I-1} \sum_{i=1}^{I} \Phi_{21,9}(z_i)(F(z_i) - \mathbf{1}\{z_i \geq y\})^2, \text{ where} \\ z_i &= \min\{y\} + i\frac{\max\{y\} - \min\{y\}}{I} \end{aligned} \tag{11}$$

(Gneiting & Ranjan, 2011). We generate I=5,000 equally spaced points z_i at which we evaluate the predictive cdf, $F(z_i)$, corresponding to the predictive density in (6). We perform this approximation for each burned area, y, in the test dataset for every MCMC iteration. The score for each iteration is the mean of the scores calculated for each y. We then compare the means across iterations for each model, with values closer to zero indicating better performance.

Ideally, we would compare all joint combinations of the counts and burned area sub-models in a full factorial design. Unfortunately, doing so would be computationally prohibitive due to the large number of potential models under consideration. Therefore, we used a multi-stage approach to select the best-performing sub-models before creating the joint model. We developed six candidate models for wildfire counts. We fit a ZINB model with overdispersion parameter δ held constant throughout space, as well as a version where δ is spatially-varying according to the nesting structure in §4.1. We also fit a zero-inflated Poisson (ZIP) model, which fixes $\delta = 0$. For these three models, we allowed the probability of inflated zeroes π to either be spatially-varying or have a logit regression link with the covariates, yielding six candidate models in total. We did not consider parametrizations of these models with a regression link on δ or π out of concern for overparametrizing the models. Table 1 shows the median difference of each model's scores, along with the standard deviation (SD) of that difference, when compared with the best performing model, which was the ZINB model with spatially-varying δ parameter and spatially-varying π.

[Table 1 about here.]

We then ran two additional versions of our best performing model by replacing the five climate covariates with either the ERC or the FWI (see §3.4). The best performing set of covariates was the original set with all five climate covariates (see Table 2).

[Table 2 about here.]

We compare burned area models composed of the four EGPD carrier functions listed in §2.1, along with a lognormal model. We include the lognormal model, despite its inflexibility with respect to tail characteristics, because it was found to be the best-performing model in Joseph et al. (2019). With the five model families (four EGPD carrier functions and lognormal), various permutations of which parameter(s) contains a regression link versus a nested spatial structure, and four different sets of covariates, we have 68 candidate models for wildfire burned area. To pare down these potential candidate models, we locally optimized in three stages. We first determined which parameter(s) would contain a regression link by running the four parameter permutations with carrier function G_1 on the climate covariate dataset. Table 3 shows the mean and median difference, along with the SD of that difference, of each permutation's twCRPS and log scores, respectively, on the test dataset. Despite the large SD, relative to the median and mean differences of the scores, we chose to proceed with the most parsimonious and easily interpretable model, which is the model with regression on κ only.

[Table 3 about here.]

We then ran this version of the G_1 model on three additional sets of covariates to determine the best combination. Table 4 shows that while all four sets of covariates perform similarly on the test dataset, the combination of ERC-FWI outperforms them all (although the differences have very large standard errors) and lends itself nicely to interpretation by capturing different aspects of wildfire behavior.

[Table 4 about here.]

Finally, we assessed the best model family from among the EGPD models with different carrier functions and the lognormal model by running the analogous versions of each EGPD carrier function and the lognormal model. We show the performance of all five models on the "best" dataset in Table 5. The additional 'Time' column shows that while ~~G_4 may slightly outperform the other four models~~G_3 and G_4 may perform similarly to G_1, ~~it takes~~ they take approximately ~~27~~ 13 and 19 times longer, respectively, than G_1 to complete 1,000 iterations in the sampling phase. We also note that some of the MCMC chains in G_3 and G_4 did not mix well, and we did not encounter this issue with G_1, G_2, and the lognormal models. We therefore proceed with G_1 for the joint model.

[Table 5 about here.]

5.2 | Model checking

With the best model and covariates in hand, we turn to assessing model adequacy for capturing key features of the wildfire data. To assess overall model performance, we combine the best-performing wildfire counts model with the best-performing burned area model using the shared random effects scheme in (7), and compare the predictive distribution to the observed wildfire counts and burned areas.

To assess the predictive ability of the wildfire counts sub-model, we first calculate the expected number of wildfires at each timepoint t for each L3 ecoregion s:

$$E[n_{s,t} \mid \pi_s, \lambda_{s,t}] = (1 - \pi_s) \times (\lambda_{s,t}). \tag{12}$$

We do this calculation for every iteration of each chain, for a total of 3,000 values of $E[n_{s,t} \mid \pi_s, \lambda_{s,t}]$, for each $t = 1, \ldots, T$ and $s = 1, \ldots, S$. For ease of visual representation, we then aggregate these values by summing over each year and across L1 ecoregion. Figure 3 displays the boxplots of these predictions, along with the observed wildfire counts.

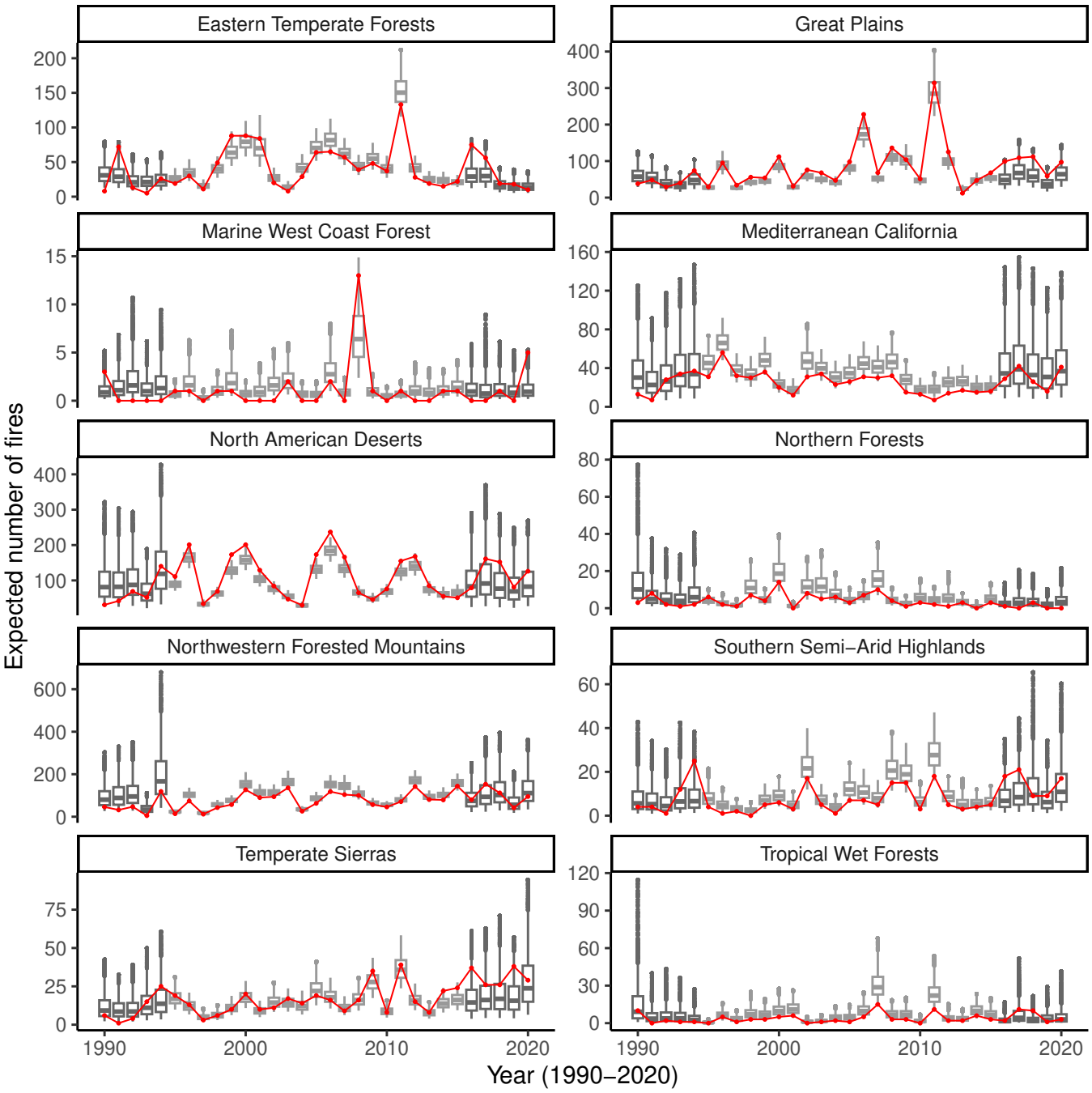

FIGURE 3 Expected wildfire counts compared with the truth. The red points represent the observed wildfire counts in a particular year for each L1 ecoregion. The dark grey boxplots indicate the test dataset years (1990–1994 and 2016–2020). The values shown are after applying a 95% winsorization to exclude extreme outliers and better visualize the trends across time and by region. Note the different scalings of the y-axes.

We next look at the predictive ability of the wildfire burned areas sub-model. We estimate the expected burned area at each timepoint t for each L3 ecoregion s through the following procedure. At each MCMC iteration, given the current parameter values,

1. Randomly draw $n_{s,t}$ from the ZINB model.
2. If $n_{s,t} = 0$, then let the expected burned area $E[y_{s,t} \mid \kappa_{s,t}, \sigma_s, \xi_s] = 0$. If $n_{s,t} > 0$, randomly draw $n_{s,t}$ observations from the EGPD model and sum these observations to get $w_{s_{i=1},t_{i=1}}$, the total expected burned area at time t for region s for the first repetition.
3. Repeat steps one and two for $i = 2, \ldots, 500$ times for each unique pair of $\{s, t\}$. Then take the average of the $i = 1, \ldots, 500$ values w_{s_i,t_i} to yield an estimate of $E[y_{s,t} \mid \kappa_{s,t}, \sigma_s, \xi_s]$, the expected burned area at timepoint t in L3 ecoregion s.

We repeat this procedure for each chain and every iteration, for a total of 3,000 values of $E[y_{s,t} \mid \kappa_{s,t}, \sigma_s, \xi_s]$, for each t and s. We then aggregate the these values by year and across L1 ecoregion. Figure 4 displays the boxplots of these predictions, along with the observed total burned areas. While our model captures temporal trends in total burned areas well, it overestimates total burned areas in some of the L1 ecoregions. We have investigated this issue and conjecture that the tail parameter ξ may be poorly estimated due to a lack of flexibility in the EGPD model. We thought we may have overestimated ξ due to the standard lognormal prior shrinking towards a value of one. To investigate this, we tried placing a logit-normal prior on ξ with support $(0, 1.5)$, which yielded indistinguishable results. This remains a topic for further follow-up.

5.3 | Covariate partial effects

We plot the covariate partial effects for both the wildfire count and burned area components of the joint model to determine which covariates have the strongest effect in the parameter with the regression link. The colors in each partial effects plot correspond to unique L2 ecoregions (see Figure 2, which has the same coloring scheme).

The two covariates that exhibited the greatest effect on the rate parameter λ within the negative binomial component of the joint model were minimum relative humidity and maximum air temperature (Figure 5). Wind speed also has a prominent effect in the Marine West Coast (L1 and L2 ecoregion of the same name). Interestingly, precipitation in the same month has a strong negative relationship with wildfire counts in this same region. However, the Marine West Coast did have many years with zero fires, so the lack of data may be contributing to these perceived effects. Precipitation in the same month has a positive relationship with wildfire counts in the Everglades (L2 ecoregion within the Tropical Wet Forests L1 ecoregion), which potentially reflects ignitions due to lightning strikes, which are common in this region.

The ERC partial effects plot for κ suggests that the importance of ERC in wildfire burned area begins at values of around 20 or above (Figure 6). Recall from §3.4 that the ERC is a relative index, so an ERC value of 20 indicates a heat potential double that of 10, similarly for an ERC value of 60 compared to ERC of 30. In the South Central Semi-Arid Prairies (part of the Great Plains L1 ecoregion), the Western Cordillera (part of the Northwestern Forested Mountains), and the Mediterranean California, ERC has a strong prominent effect on wildfire burned area. In contrast, the FWI partial effects on burned area are minimal in these

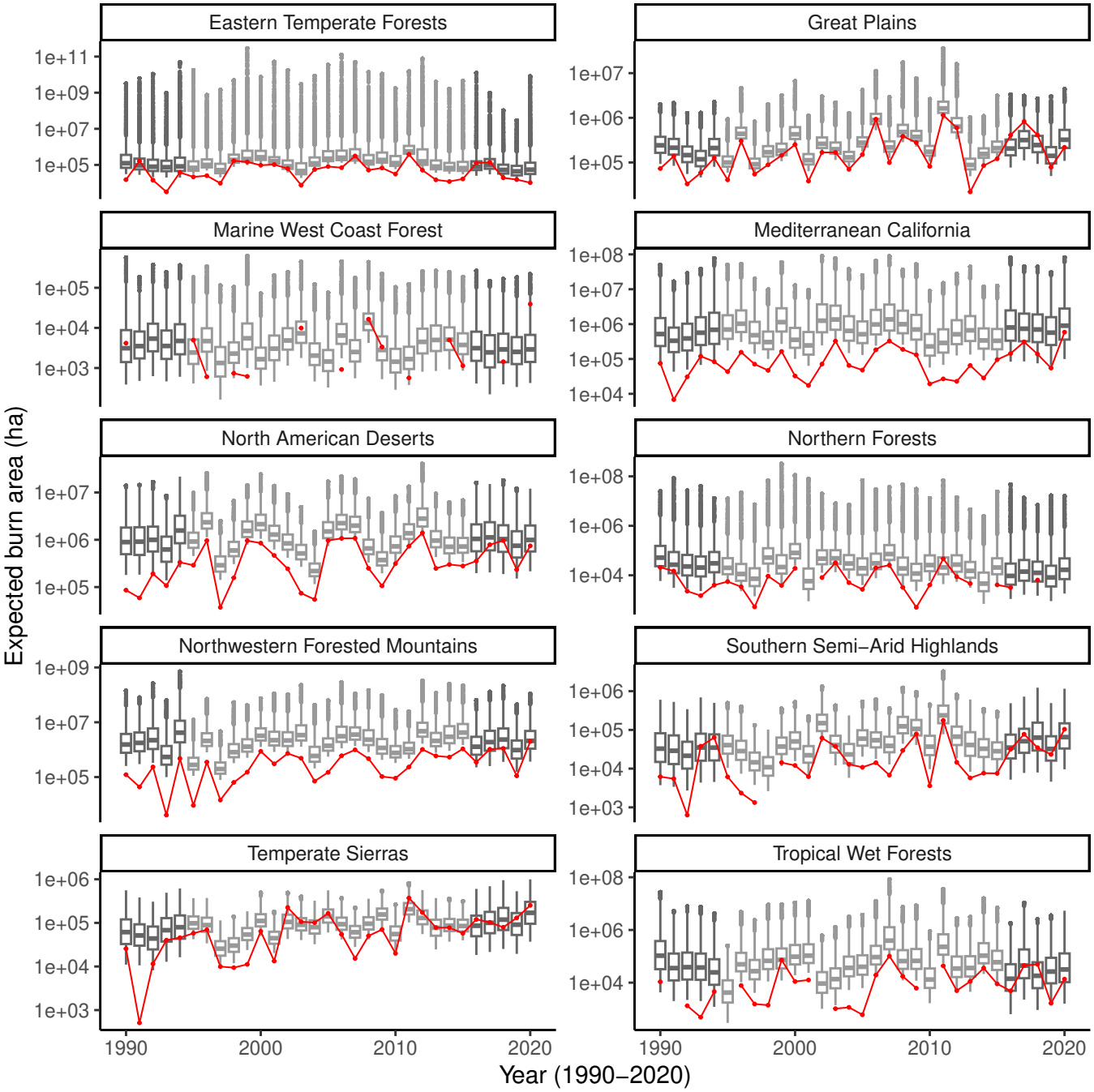

FIGURE 4 Predicted expected total wildfire burned area compared with the observed wildfire burned areas. Analogous to Figure 3, the red points are the observed wildfire burned area and the dark grey boxplots indicate the test dataset years. The values shown are on the log scale and are post 95% winsorization.

~~two~~three ecoregions, suggesting that heat potential is a more important driver of larger burned areas than wildfire weather. In the Cold and Warm Deserts (both part of the North American Deserts L1 ecoregion), FWI has a strong effect on wildfire burned area, while the ERC partial effects are much smaller in ~~this~~these regions. This suggests that wildfire weather is a more important determinant of burned area in the North American Deserts than the heat potential of a wildfire. In the Marine West Coast and

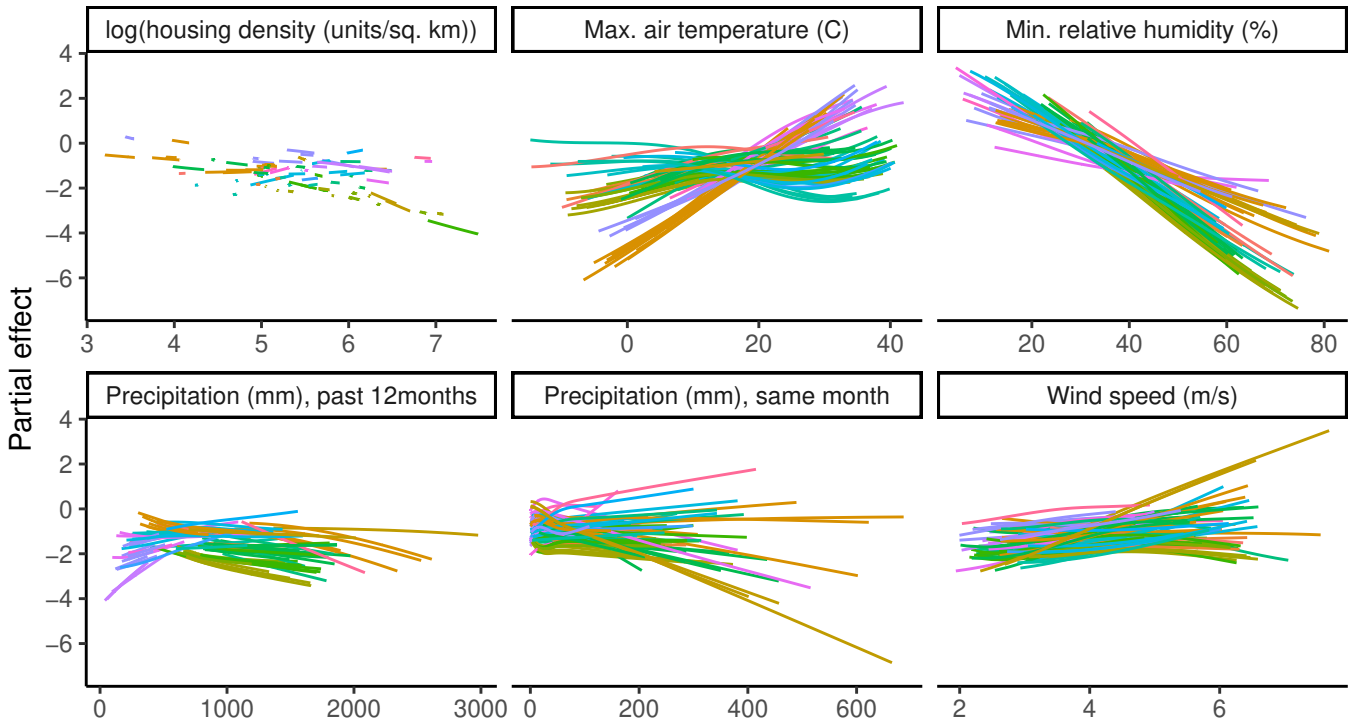

FIGURE 5 Partial effects on log-transformed rate parameter λ of negative binomial piece of overall model. These partial effects are solely the effects from the linear and basis functions of the covariates. These plots do not incorporate the spatio-temporal ICAR parameter $\mathbf{\Phi}$, the shared random effect θ_t, nor the area offset. Lines shown are posterior medians for each L3 ecoregion, colored by L2 ecoregion.

in the Northern Forests (corresponding to the two L2 ecoregions Atlantic Highlands and Boreal Plain), changes in ERC and FWI have minimal effects on the κ parameter. Again, this is likely due to lower numbers of wildfires in these ecoregions.

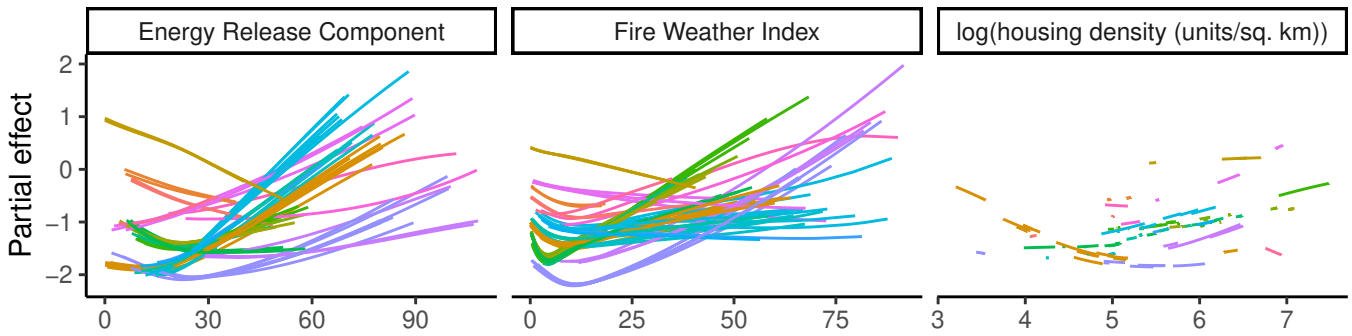

FIGURE 6 Partial effects on log-transformed parameter κ of the EGPD piece of overall model. Lines shown are posterior medians for each L3 ecoregion, colored by L2 ecoregion.

5.4 | Burned area return levels and high quantiles

We now turn to investigating predicted high quantiles of burned areas in L1 ecoregions in two ways. These plots are colored by L1 ecoregion, matching the map of the US in Figure 7.

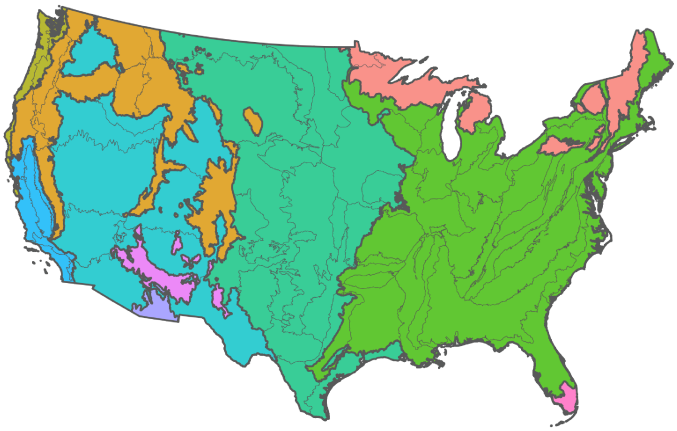

FIGURE 7 Map of the contiguous United States colored by L1 ecoregion.

First, we look at the 50-year return levels of wildfire burned areas by L1 ecoregion. We determine these burned area levels u through the following equation:

$$F_{s,t}^{-1}\{q^{1/(12\eta_{s,t})}\} = u_{s,t}, \tag{13}$$

where $F_{s,t}^{-1}$ is the quantile function corresponding to the predictive density of (6) for L3 ecoregion s at time t. The variable $\eta_{s,t}$ is the expected number of wildfires, as calculated from (12). We set $q = \frac{49}{50}$ to correspond to the 50-year return levels. From Figure 8, we see that Mediterranean California, North American Deserts, and Northwestern Forested Mountains have the most extreme wildfire burned areas. These predictions align with what we have experienced in the Western US over the past decade.

We then look at the 98th quantile of wildfire burned areas by L1 ecoregions. We calculated this by setting $q = \frac{49}{50}$ in the following equation:

$$F_{s,t}^{-1}\{q\} = u_{s,t}. \tag{14}$$

While (14) describes the 98th quantile of the distribution of wildfire burned area for any individual wildfire, given that such a wildfire occurs, (13) describes the 98th quantile of the distribution of total burned areas, marginalized over the distribution of wildfire counts. From Figure 9, we see that Marine West Coast Forest and Eastern Temperate Forests have the largest burned areas at the 98th quantile, but their credible intervals are also quite wide. Northwestern Forested Mountains and North American deserts have the least variability in their predictions.

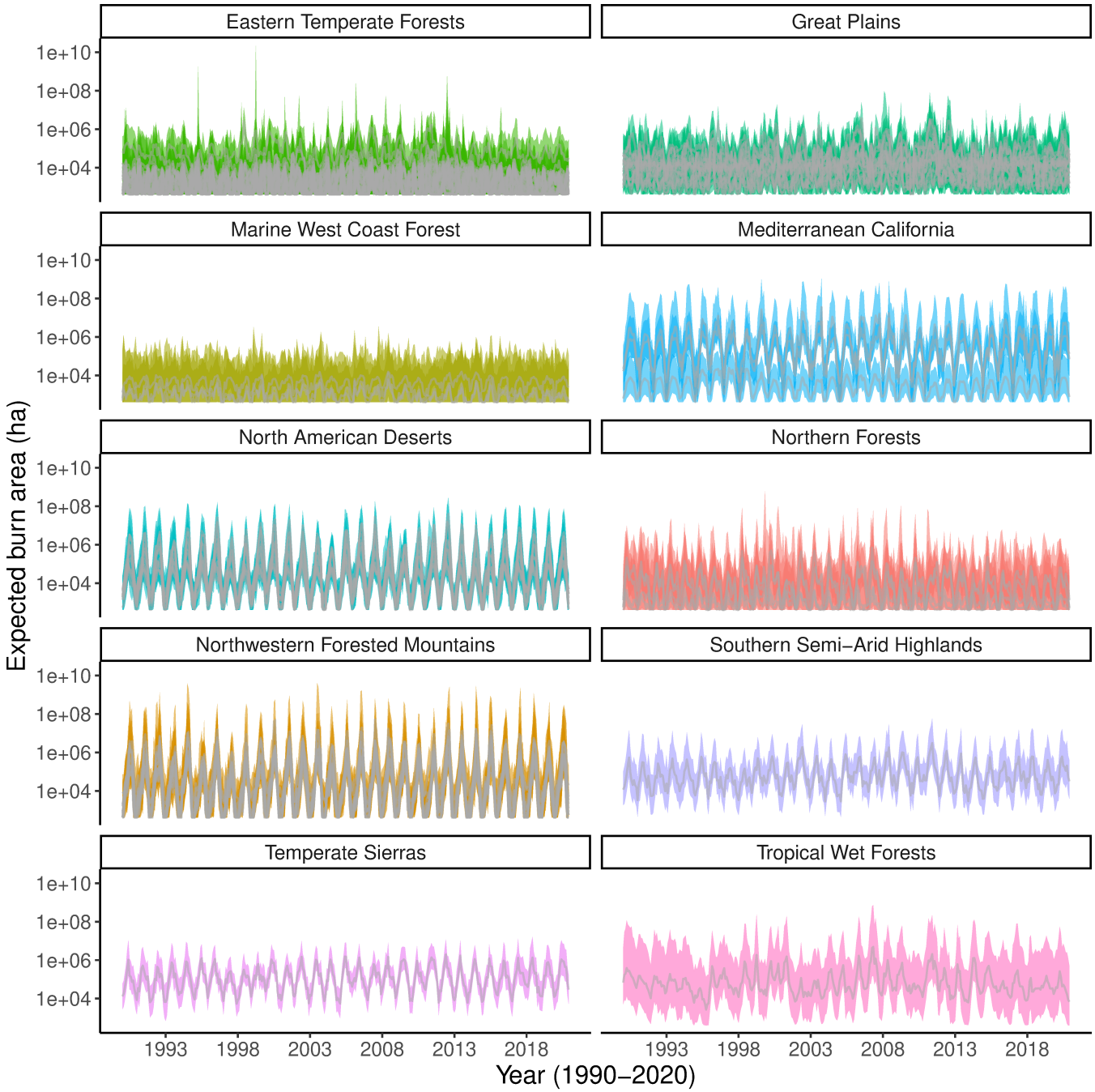

FIGURE 8 50-year return levels for wildfire burned areas by L1 ecoregion. Each grey line represents an L3 ecoregion and the colored ribbons represent that 95% credible interval for each line. The values shown are on the log scale.

5.5 | Maps of posterior parameter estimates

We compare the univariate GPD estimates of ξ with the posterior medians of ξ from our model (see Figure 10). Estimates of ξ from our model are larger in the Western US than in the Eastern US, which is consistent with the proliferation of high-impact

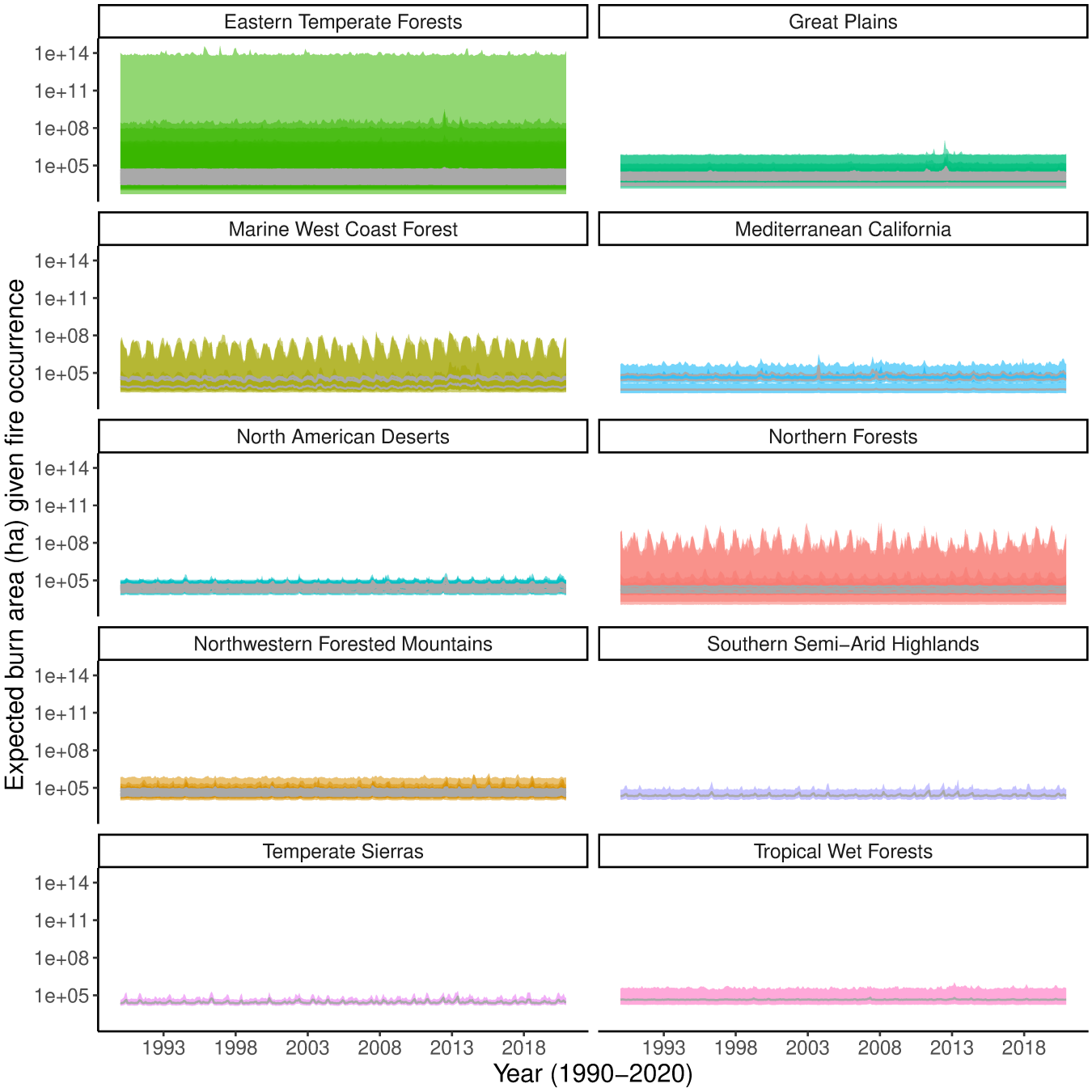

FIGURE 9 98th quantile of wildfire burned areas by L1 ecoregion. The lines and ribbons are analogous to those in Figure 8.

wildfire events in the West. In addition, the spatial borrowing of strength allows us to obtain estimates of ξ in areas where there were not enough observations to fit a GPD in the exploratory stage.

Figure 11 visualizes the spatially-varying modulating parameter γ of the joint model across the US by L3 ecoregion. Marine West Coast Forest ecoregions have a strong negative correlation between counts and burned areas, so the relationship between wildfire counts and burned area will follow the patterns in the first row of Table 6, while ecoregions in the Great Plains have a strong positive correlation (second row of the same table).

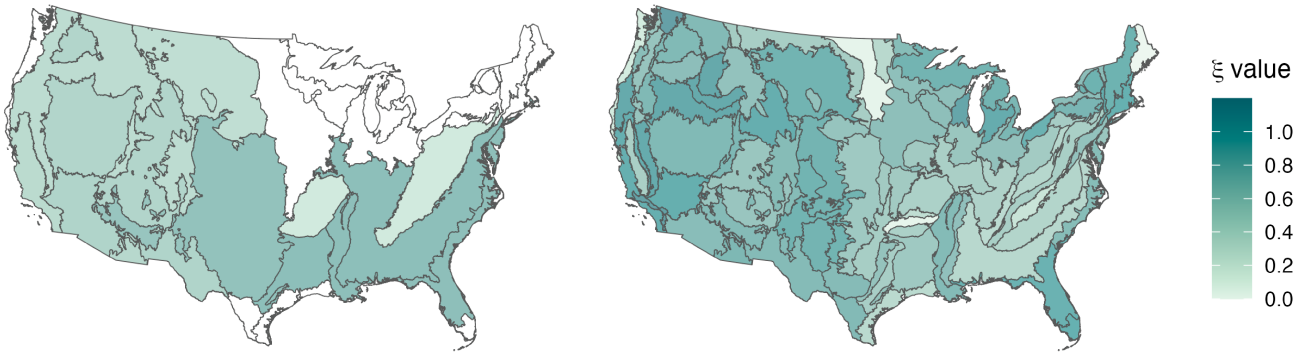

FIGURE 10 Posterior median of ξ by L3 ecoregion (right) compared to estimates ξ obtained from the MLE using the GPD by L2 ecoregion, with a the threshold set at the 90th percentile (left). The map on the left is the same as in Figure 1 but with an adjusted scale to be comparable to the posterior values of ξ on the right.

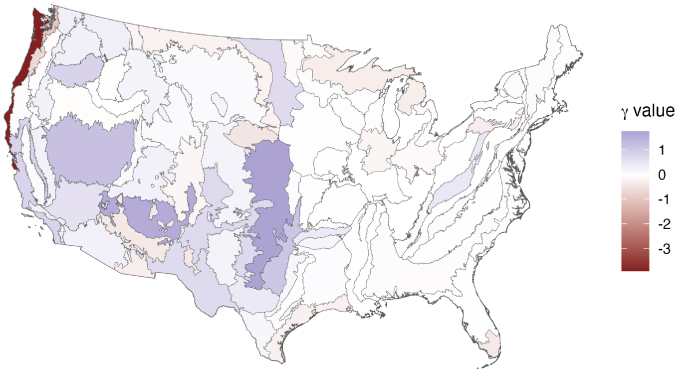

FIGURE 11 Posterior medians of γ across the United States by L3 ecoregion. Red regions have a negative γ value, indicating an inverse relationship between counts and burned areas. Blue regions have a positive γ value, indicating that counts and burned areas have a positive relationship. Darker shades of these two colors represent a stronger relationship between the count and burn parameters.

[Table 6 about here.]

6 | DISCUSSION

We have implemented a joint Bayesian hierarchical model featuring a truncated EGPD sub-model, incorporating nonlinear effects in the covariates and accounting for any residual spatio-temporal variation with a Gaussian random effect. The flexibility of the model allows us to incorporate all of the wildfire burned area data in the estimation of our spatially-varying shape and scale parameters, ξ and σ. The EGPD allows us to model the moderate and extreme wildfire burned areas, while allowing us to exploit EVT in extrapolating burned areas at high quantiles.

Our investigation into three datasets revealed a distinction in the covariates influencing wildfire occurrences compared to those affecting burned sizes. Notably, individual climate covariates excel in predicting wildfire counts, while fire indices, specifically ERC and FWI, outperform in predicting burned areas. This distinction underscores the nuanced nature of wildfire behavior and necessitates a tailored approach for accurate prediction. In our burned areas sub-model, the combined use of ERC and FWI surpasses individual fire indices, revealing regional variations in their effects. ERC significantly influences burned areas in the Great Plains, ~~and~~ the Northwestern Forested Mountains, and the Mediterranean California, while FWI emerges as more crucial in the North American Deserts. These regional differences highlight the need for a nuanced understanding of fire indices in different ecoregions.

While our model adeptly captures temporal trends of wildfire burned areas, potential overestimation in certain ecoregions warrants attention. We suspect that the tail parameter ξ may be poorly estimated due to the EGPD model structure, suggesting a need for further exploration. Additionally, variations in vegetation and forestry across the CONUS raise concerns about potential overestimation stemming from the use of a coniferous forest fuel model in the calculation of ERC. Future research should consider obtaining ERC values calculated from diverse fuel models that best represent each ecoregion.

As the threat of more extensive wildfires looms with changing climate patterns, our study contributes vital insights for improved wildfire management within the CONUS. Recognizing the nonlinear drivers of wildfire occurrences and burned areas is essential for developing effective mitigation strategies and resource allocation policies. Our multifaceted approach, considering individual climate covariates, fire weather risk, and fire burn potential, provides an improved understanding of the drivers of wildfire dynamics.

FINANCIAL DISCLOSURE

None reported.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

REFERENCES

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *Int. J. Climatol.*, *33*(1), 121–131. doi: 10.1002/joc.3413

Abatzoglou, J. T., & Kolden, C. A. (2013). Relationships between climate and macroscale area burned in the Western United States. *Int. J. Wildland Fire*, *22*, 1003–1020. doi: 10.1071/WF13019

Abatzoglou, J. T., & Williams, A. P. (2016). Impact of anthropogenic climate change on wildfire across western US forests. *Proc. Natl. Acad. Sci. U.S.A.*, *113*(42), 11770–11775. doi: 10.1073/pnas.1607171113

Barbero, R., Abatzoglou, J. T., Pimont, F., Ruffault, J., & Curt, T. (2020). Attributing increases in fire weather to anthropogenic climate change over France. *Front. Earth Sci.*, *8*. doi: 10.3389/feart.2020.00104

Barbero, R., Abatzoglou, J. T., Steel, E. A., & Larkin, N. K. (2014). Modeling very large-fire occurrences over the continental United States from weather and climate forcing. *Environ. Res. Lett.*, *9*(12). doi: 10.1088/1748-9326/9/12/124009

Becker, D. G., Woolford, D. G., & Dean, C. B. (2022). Assessing dependence between frequency and severity through shared random effects. *PLoS ONE*, *17*(8), e0271904. doi: 10.1371/journal.pone.0271904

Behrens, C. N., Lopes, H. F., & Gamerman, D. (2004). Bayesian analysis of extreme events with threshold estimation. *Stat. Modell.*, *4*(3), 227–244. doi: 10.1191/1471082X04st075oa

Bradshaw, L. S., Deeming, J. E., Burgan, R. E., & Cohen, J. D. (1984). *The 1978 National Fire-Danger Rating System: technical documentation* . Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. doi: 10.2737/INT-GTR-169

Brillinger, D. R., Preisler, H. K., & Benoit, J. W. (2003). Risk assessment: A forest fire example. In *Institute of Mathematical Statistics Lecture Notes - Monograph Series* (pp. 177–196). Beachwood, OH: Institute of Mathematical Statistics. doi: 10.1214/lnms/1215091142

Carreau, J., & Bengio, Y. (2009). A hybrid Pareto model for asymmetric fat-tailed data: the univariate case. *Extremes*, *12*(1), 53–76. doi: 10.1007/s10687-008-0068-0

Castel-Clavera, J., Pimont, F., Opitz, T., Ruffault, J., Rivière, M., Dupuy, J.-L., … Dupuy, J.-L. (2022). Disentangling the factors of spatio-temporal patterns of wildfire activity in south-eastern France. *Int. J. Wildland Fire*, *32*(1), 15–28. doi: 10.1071/WF22086

Castro-Camilo, D., Huser, R., & Rue, H. (2019). A spliced gamma-generalized Pareto model for short-term extreme wind speed probabilistic forecasting. *J. Agric. Biol. Environ. Stat.*, *24*(3), 517–534. doi: 10.1007/s13253-019-00369-z

Cisneros, D., Gong, Y., Yadav, R., Hazra, A., & Huser, R. (2023). A combined statistical and machine learning approach for spatial prediction of extreme wildfire frequencies and sizes. *Extremes*, *26*(2), 301–330. doi: 10.1007/s10687-022-00460-8

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. London: Springer. doi: 10.1007/978-1-4471-3675-0

de Carvalho, M., Pereira, S., Pereira, P., & de Zea Bermudez, P. (2022). An extreme value Bayesian lasso for the conditional left and right tails. *J. Agric. Biol. Environ. Stat.*, *27*(2), 222–239. doi: 10.1007/s13253-021-00469-9

Dennison, P. E., Brewer, S. C., Arnold, J. D., & Moritz, M. A. (2014). Large wildfire trends in the western United States, 1984–2011. *Geophys. Res. Lett.*, *41*(8), 2928–2933. doi: 10.1002/2014GL059576

do Nascimento, F. F., Gamerman, D., & Lopes, H. F. (2012). A semiparametric Bayesian approach to extreme value estimation. *Stat. Comput.*, *22*(2), 661–675. doi: 10.1007/s11222-011-9270-z

Duane, A., Castellnou, M., & Brotons, L. (2021). Towards a comprehensive look at global drivers of novel extreme wildfire

events. *Clim. Change*, *165*(3), 43. doi: 10.1007/s10584-021-03066-4

Eidenshink, J., Schwind, B., Brewer, K., Zhu, Z.-L., Quayle, B., & Howard, S. (2007). A project for monitoring trends in burn severity. *Fire Ecol.*, *3*(1), 3–21. doi: 10.4996/fireecology.0301003

Eilers, P. H. C., & Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Stat. Sci.*, *11*(2), 89–121. doi: 10.1214/ss/1038425655

Fosberg, M. A. (1978). Weather in wildland fire management: the fire weather index. *US For Serv Reprints of articles by FS employees*.

Frigessi, A., Haug, O., & Rue, H. (2002). A dynamic mixture model for unsupervised tail estimation without threshold selection. *Extremes*, *5*(3), 219–235. doi: 10.1023/A:1024072610684

Gabry, J., Češnovar, R., & Johnson, A. (2023). cmdstanr: R interface to 'cmdstan' [Computer software manual]. Retrieved from `https://mc-stan.org/cmdstanr/,https://discourse.mc-stan.org`

Gamet, P., & Jalbert, J. (2022). A flexible extended generalized Pareto distribution for tail estimation. *Environmetrics*, *33*(6). doi: 10.1002/env.2744

Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *J. Am. Stat. Assoc.*, *102*(477), 359–378. doi: 10.1198/016214506000001437

Gneiting, T., & Ranjan, R. (2011). Comparing density forecasts using threshold- and quantile-weighted scoring rules. *J. Bus. Econom. Statist.*, *29*(3), 411–422. doi: 10.1198/jbes.2010.08110

Goss, M., Swain, D. L., Abatzoglou, J. T., Sarhadi, A., Kolden, C. A., Williams, A. P., & Diffenbaugh, N. S. (2020). Climate change is increasing the likelihood of extreme autumn wildfire conditions across California. *Environ. Res. Lett.*, *15*(9). doi: 10.1088/1748-9326/ab83a7

Hawbaker, T. J., Radeloff, V. C., Stewart, S. I., Hammer, R. B., Keuler, N. S., & Clayton, M. K. (2013). Human and biophysical influences on fire occurrence in the United States. *Ecol. Appl.*, *23*(3), 565–582. doi: 10.1890/12-1816.1

Iglesias, V., Stavros, N., Balch, J. K., Barrett, K., Cobian-Iñiguez, J., Hester, C., … Travis, W. R. (2022). Fires that matter: Reconceptualizing fire risk to include interactions between humans and the natural environment. *Environ. Res. Lett.*, *17*(4). doi: 10.1088/1748-9326/ac5c0c

Jones, M. W., Abatzoglou, J. T., Veraverbeke, S., Andela, N., Lasslop, G., Forkel, M., … Le Quéré, C. (2022). Global and regional trends and drivers of fire under climate change. *Rev. Geophys.*, *60*(3). doi: 10.1029/2020RG000726

Joseph, M. B., Rossi, M. W., Mietkiewicz, N. P., Mahood, A. L., Cattau, M. E., St. Denis, L. A., … Balch, J. K. (2019). Spatiotemporal prediction of wildfire size extremes with Bayesian finite sample maxima. *Ecol. Appl.*, *29*(6). doi: 10.1002/eap.1898

Koh, J., Pimont, F., Dupuy, J.-L., & Opitz, T. (2023). Spatiotemporal wildfire modeling through point processes with moderate

and extreme marks. *Ann. Appl. Stat.*, *17*(1), 560–582. doi: 10.1214/22-AOAS1642

Krock, M., Bessac, J., Stein, M. L., & Monahan, A. H. (2022). Nonstationary seasonal model for daily mean temperature distribution bridging bulk and tails. *Weather Clim. Extremes*, *36*. doi: 10.1016/j.wace.2022.100438

Lang, S., & Brezger, A. (2004). Bayesian P-splines. *J. Comput. Graph. Statist.*, *13*(1), 183–212. doi: 10.1198/1061860043010

Littell, J. S., McKenzie, D., Peterson, D. L., & Westerling, A. L. (2009). Climate and wildfire area burned in western U.S. ecoprovinces, 1916–2003. *Ecol. Appl.*, *19*(4), 1003–1021. doi: 10.1890/07-1183.1

Liu, Y., Stanturf, J., & Goodrick, S. (2010). Trends in global wildfire potential in a changing climate. *For. Ecol. Manage.*, *259*(4), 685–697. doi: 10.1016/j.foreco.2009.09.002

MacDonald, A., Scarrott, C., Lee, D., Darlow, B., Reale, M., & Russell, G. (2011). A flexible extreme value mixture model. *Comput. Stat. Data Anal.*, *55*(6), 2137–2157. doi: 10.1016/j.csda.2011.01.005

Malamud, B. D., Millington, J. D. A., & Perry, G. L. W. (2005). Characterizing wildfire regimes in the United States. *Proc. Natl. Acad. Sci. U.S.A.*, *102*(13), 4694–4699. doi: 10.1073/pnas.0500880102

Martinuzzi, S., Allstadt, A. J., Pidgeon, A. M., Flather, C. H., Jolly, W. M., & Radeloff, V. C. (2019). Future changes in fire weather, spring droughts, and false springs across U.S. National Forests and Grasslands. *Ecol. Appl.*, *29*(5). doi: 10.1002/eap.1904

Matheson, J. E., & Winkler, R. L. (1976). Scoring rules for continuous probability distributions. *Manage. Sci.*, *22*(10), 1087–1096. doi: 10.1287/mnsc.22.10.1087

Mockrin, M. H., Radeloff, V. C., & Helmers, D. P. (2023). *Census block level housing change 1990 - 2020 for the conterminous United States* [vector]. SILVIS Lab, Dept of Forest & Wildlife Ecology, University of Wisconsin-Madison.

Naveau, P., Huser, R., Ribereau, P., & Hannart, A. (2016). Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. *Water Resour. Res.*, *52*(4), 2753–2769. doi: 10.1002/2015WR018552

Omernik, J. M., & Griffith, G. E. (2014). Ecoregions of the conterminous United States: Evolution of a hierarchical spatial framework. *Environ. Manage.*, *54*(6), 1249–1266. doi: 10.1007/s00267-014-0364-1

Opitz, T., Bonneu, F., & Gabriel, E. (2020). Point-process based Bayesian modeling of space–time structures of forest fire occurrences in Mediterranean France. *Spatial Stat.*, *40*. doi: 10.1016/j.spasta.2020.100429

Papastathopoulos, I., & Tawn, J. A. (2013). Extended generalised Pareto models for tail estimation. *J. Statist. Plann. Inference*, *143*(1), 131–143. doi: 10.1016/j.jspi.2012.07.001

Parente, J., Pereira, M. G., Amraoui, M., & Fischer, E. M. (2018). Heat waves in Portugal: Current regime, changes in future climate and impacts on extreme wildfires. *Sci. Total Environ*, *631–632*, 534–549. doi: 10.1016/j.scitotenv.2018.03.044

Pausas, J. G., & Keeley, J. E. (2021). Wildfires and global change. *Front. Ecol. Environ.*, *19*(7), 387–395. doi: 10.1002/fee.2359

Pimont, F., Fargeon, H., Opitz, T., Ruffault, J., Barbero, R., Martin-StPaul, N., … Dupuy, J.-L. (2021). Prediction of regional

wildfire activity in the probabilistic Bayesian framework of Firelihood. *Ecol. Appl.*, *31*(5). doi: 10.1002/eap.2316

Pimont, F., Ruffault, J., Opitz, T., Fargeon, H., Barbero, R., Castel-Clavera, J., … Dupuy, J.-L. (2022). Future expansion, seasonal lengthening and intensification of fire activity under climate change in southeastern France. *Int. J. Wildland Fire*, *32*(1), 4–14. doi: 10.1071/WF22103

Preisler, H., Brillinger, D. R., Burgan, R. E., & Benoit, J. (2004). Probability based models for estimation of wildfire risk. *Int. J. Wildland Fire*, *13*(2), 133–142. doi: 10.1071/WF02061

Preisler, H. K., Chen, S.-C., Fujioka, F., Benoit, J. W., & Westerling, A. L. (2008). Wildland fire probabilities estimated from weather model-deduced monthly mean fire danger indices. *Int. J. Wildland Fire*, *17*(3), 305–316. doi: 10.1071/wf06162

Preisler, H. K., & Westerling, A. L. (2007). Statistical model for forecasting monthly large wildfire events in western United States. *J. Appl. Meteorol. Climatol.*, *46*(7), 1020–1030. doi: 10.1175/JAM2513.1

Richards, J., & Huser, R. (2022). *Regression modelling of spatiotemporal extreme U.S. wildfires via partially-interpretable neural networks* (arXiv:2208.07581). arXiv. doi: 10.48550/arXiv.2208.07581

Stan Development Team. (2023). Stan modeling language users guide and reference manual [Computer software manual]. Retrieved from `https://mc-stan.org/`

Stavros, E. N., Abatzoglou, J. T., McKenzie, D., & Larkin, N. K. (2014). Regional projections of the likelihood of very large wildland fires under a changing climate in the contiguous western United States. *Clim. Change*, *126*(3), 455–468. doi: 10.1007/s10584-014-1229-6

Stein, M. L. (2021a). A parametric model for distributions with flexible behavior in both tails. *Environmetrics*, *32*(2). doi: 10.1002/env.2658

Stein, M. L. (2021b). Parametric models for distributions when interest is in extremes with an application to daily temperature. *Extremes*, *24*(2), 293–323. doi: 10.1007/s10687-020-00378-z

Strauss, D., Bednar, L., & Mees, R. (1989). Do one percent of the forest fires cause ninety-nine percent of the damage? *For. Sci.*, *35*(2), 319–328.

Tancredi, A., Anderson, C., & O'Hagan, A. (2006). Accounting for threshold uncertainty in extreme value estimation. *Extremes*, *9*(2), 87–106. doi: 10.1007/s10687-006-0009-8

Tencaliec, P., Favre, A.-C., Naveau, P., Prieur, C., & Nicolet, G. (2020). Flexible semiparametric generalized Pareto modeling of the entire range of rainfall amount. *Environmetrics*, *31*(2). doi: 10.1002/env.2582

Turner, J., & Lawson, B. (1978). *Weather in the Canadian Forest Fire Danger Rating System: A user guide to national standards and practices*. Victoria, British Columbia, Canada: Fisheries and Environment Canada, Canadian Forest Service, Pacific Forest Research Centre.

Van Oldenborgh, G. J., Krikken, F., Lewis, S., Leach, N. J., Lehner, F., Saunders, K. R., … Otto, F. E. L. (2021). Attribution of the Australian bushfire risk to anthropogenic climate change. *Nat. Hazards Earth Syst. Sci.*, *21*(3), 941–960. doi: 10.5194/nhess-21-941-2021

Wang, X., Wotton, B. M., Cantin, A. S., Parisien, M.-A., Anderson, K., Moore, B., & Flannigan, M. D. (2017). cffdrs: an r package for the canadian forest fire danger rating system. *Ecological Processes*, *6*(1), 1–11.

Williams, A. P., Abatzoglou, J. T., Gershunov, A., Guzman-Morales, J., Bishop, D. A., Balch, J. K., & Lettenmaier, D. P. (2019). Observed impacts of anthropogenic climate change on wildfire in California. *Earth's Future*, *7*(8), 892–910. doi: 10.1029/2019EF001210

Woolford, D. G., Bellhouse, D. R., Braun, W. J., Dean, C. B., Martell, D. L., & Sun, J. (2011). A spatio-temporal model for people-caused forest fire occurrence in the Romeo Malette Forest. *J. Environ. Stat.*, *2*(1), 2–25.

Woolford, D. G., Martell, D. L., McFayden, C. B., Evens, J., Stacey, A., Wotton, B. M., & Boychuk, D. (2021). The development and implementation of a human-caused wildland fire occurrence prediction system for the province of Ontario, Canada. *Can. J. For. Res.*, *51*(2), 303–325. doi: 10.1139/cjfr-2020-0313

Yadav, R., Huser, R., Opitz, T., & Lombardo, L. (2023). Joint modelling of landslide counts and sizes using spatial marked point processes with sub-asymptotic mark distributions. *J. R. Stat. Soc. Ser. C Appl. Stat.*, qlad077. doi: 10.1093/jrsssc/qlad077

SUPPORTING INFORMATION

To fully replicate this research, all code and datasets may be found at `https://doi.org/10.5281/zenodo.10520994`.

Model	Regression			Difference	
	λ	π	δ	Median	SD
ZI-Negative Binomial	✓	s	s	0.00	0.00
	✓	✓	s	-45.93	274.42
	✓	s	c	-502.91	314.63
	✓	✓	c	-544.13	307.83
ZI-Poisson	✓	✓		-1140.01	320.45
	✓	s		-1145.58	349.93

TABLE 1 Comparison of log scores on six candidate wildfire counts models. A '✓' in the 'Regression' column indicates that the respective parameter had a regression link with the covariates. An 's' indicates the parameter was spatially-varying according to the nesting structure described in §4.1, with no regression link. A 'c' indicates the parameter was spatially and temporally constant. An empty space indicates that the parameter was not involved in that particular model.

	Difference	
Dataset	Median	SD
Climate	0.00	0.00
ERC-FWI	-617.69	348.48
ERC	-738.98	342.80
FWI	-1030.65	395.42

TABLE 2 Comparison of log scores from the best count model, run on three different sets of covariates.

	Regression			twCRPS		Log score	
Model	κ	σ	ξ	Mean	SD	Median	SD
G_1	✓	s	s	0.00	0.00	0.00	0.00
	s	✓	s	5.39	14.03	-80.52	935.30
	✓	✓	s	12.45	17.96	-199.66	971.43
	✓	✓	✓	21.84	22.69	-432.41	1144.93

TABLE 3 Comparison of twCRPS and log scores for four parameter combinations of the G_1 EGPD model. Scores shown for twCRPS are multiplied by a factor of 10^2 to more easily see the magnitude of differences. Symbols are analogous to those in Table 1.

	twCRPS		Log score	
Dataset	Mean	SD	Median	SD
ERC-FWI	0.00	0.00	0.00	0.00
ERC	0.31	9.45	-8.40	90.86
FWI	0.74	9.28	-11.18	80.96
Climate	1.92	11.78	-34.33	934.27

TABLE 4 Comparison of twCRPS and log scores from the best G_1 parametrization ~~model~~, run on four different combinations of covariates.

	Regression			twCRPS		Log score		
Model	$\kappa(\mu)$	σ	ξ	Mean	SD	Median	SD	Time (hrs)
G_1	✓	s	s	0.00	0.00	0.00	0.00	6.90
G_3		✓	s	1.65	9.05	-64.56	80.44	89.21
G_4	✓	s	s	3.19	10.85	-16.70	100.07	129.09
G_2	✓	s	s	8.17	12.76	-39.83	70.60	11.55
Lognormal	(✓)	s		2.45×10^3	9.64×10^2	-2.74×10^3	8.17×10^3	5.23

TABLE 5 Comparison of twCRPS and log scores for one-parameter versions of five families of burn models run on the ERC-FWI dataset. Symbols and scale of scores are analogous to those in Table 4. The 'Time' column indicates how long the model took to complete 1000 MCMC iterations in the sampling phase (i.e. excluding warmup phase) using `Stan`.

	$\theta_t < 0$	$\theta_t > 0$
$\gamma_s < 0$	fewer counts, larger burned area	more counts, smaller burned area
$\gamma_s > 0$	fewer counts, smaller burned area	more counts, larger burned area

TABLE 6 Interpretation of the spatially varying values of the scaling parameter γ combined with the temporally varying values of the shared random effect θ.