Earth's Future

RESEARCH ARTICLE
10.1029/2025EF007485

Accounting for Extremes in Modeling the Size and Likelihood of Large Fires in the United States

Amirali Asadian[1], Masoud Zaerpour[2], Simon Papalexiou[3], Amir AghaKouchak[4,5,6], and Farshid Vahedifard[1,6]

[1]Department of Civil and Environmental Engineering, Tufts University, Medford, MA, USA, [2]Department of Civil Engineering, University of Calgary, Calgary, AB, Canada, [3]Institute for Global Water Security, Hamburg University of Technology, Hamburg, Germany, [4]Department of Civil and Environmental Engineering, University of California Irvine, Irvine, CA, USA, [5]Department of Earth System Science, University of California Irvine, Irvine, CA, USA, [6]United Nations University Institute for Water, Environment and Health (UNU-INWEH), Richmond Hill, ON, Canada

Key Points:

- Relying solely on body-only statistics can significantly misrepresent the size and exceedance probability of extremely large fires
- Ignoring the distribution's heavy tail leads to substantially misrepresenting the probability of extreme fire events
- The best-fitting tail model varies by region and spatial scale, with no single distribution dominating across all ecoregions and Geographic Area Coordination Centers

Correspondence to:
F. Vahedifard,
Farshid.vahedifard@tufts.edu

Citation:
Asadian, A., Zaerpour, M., Papalexiou, S., AghaKouchak, A., & Vahedifard, F. (2026). Accounting for extremes in modeling the size and likelihood of large fires in the United States. *Earth's Future*, *14*, e2025EF007485. https://doi.org/10.1029/2025EF007485

Received 3 OCT 2025
Accepted 4 APR 2026

Author Contributions:
Conceptualization: Amir AghaKouchak, Farshid Vahedifard
Data curation: Amirali Asadian
Formal analysis: Amirali Asadian, Farshid Vahedifard
Funding acquisition: Farshid Vahedifard
Investigation: Amirali Asadian, Masoud Zaerpour, Simon Papalexiou, Farshid Vahedifard
Methodology: Amirali Asadian, Masoud Zaerpour, Simon Papalexiou, Amir AghaKouchak, Farshid Vahedifard
Project administration: Farshid Vahedifard
Resources: Simon Papalexiou, Farshid Vahedifard
Software: Amir AghaKouchak
Supervision: Amir AghaKouchak, Farshid Vahedifard
Validation: Amirali Asadian
Visualization: Amirali Asadian
Writing – original draft: Amirali Asadian

Abstract Wildfires pose growing threats to ecosystems, infrastructure, and communities, yet fire size distributions are often modeled without sufficient attention to rare, high-impact events. Most statistical approaches emphasize the central body of the distribution, which can obscure the behavior of extreme wildfires (i.e., tail events) that have the most significant impact. Here, we introduce a framework that explicitly distinguishes between the statistical characteristics of body and tail fire size distributions. We analyzed 30,331 large fire perimeters from the Monitoring Trends in Burn Severity (MTBS) data set (1984–2024), disaggregated across 105 Level III ecoregions and 10 Geographic Area Coordination Centers (GACCs) across the United States. For each region, we fit three candidate distributions (Pareto Type II, lognormal, Weibull) to both the full data set and an exceedance-based tail sample. Results show that models calibrated only to the body can substantially misestimate the exceedance probability (by up to three orders of magnitude) and burned area (by hundreds of thousands of acres) for large fires. For the body of the distribution, the lognormal and Pareto Type II distributions generally outperform the Weibull. For the tail, the picture is more nuanced: the Pareto Type II and lognormal perform comparably at the ecoregion level, while the Weibull, despite being the weakest body model, provides the best tail fit at the GACC scale. These findings highlight the need to model body and tail behavior separately and show that optimal distributions vary across scales, underscoring the importance of region-specific tail models for suppression budgeting, fuel treatment, and resilience planning.

Plain Language Summary Wildfires are becoming more destructive across the United States, making it increasingly important to understand how large they can grow. This information helps guide infrastructure design, emergency response, and long-term wildfire planning. Scientists often rely on statistical models to describe past fire sizes and estimate the likelihood of very large fires. However, many commonly used methods focus on smaller, more frequent fires and do not accurately capture rare, extreme events that cause the most damage. In this study, we analyzed data from more than 30,000 large fires across the United States and tested different ways of modeling fire size. We found that focusing only on the "average" range of fires can significantly misestimate the size (by hundreds of thousands of acres) and likelihood (by up to three orders of magnitude) of the largest fires. We also found that the best model for typical fires is not always the best for extreme fires, and the choice depends on region and scale. These findings highlight the need to model moderate and extremely large fires differently, improving decisions on prevention, resource allocation, and community preparedness.

1. Introduction

In recent decades, the global wildfire landscape has become increasingly complex. While burned areas have declined in some regions, the exposure, frequency, and intensity of large wildfires have increased in others. Consequently, the overall exposure of people and assets to large wildfires has grown globally (e.g., Abatzoglou et al., 2021; Cunningham et al., 2024; Cunningham et al., 2025; Davis et al., 2025; Dennison et al., 2014; Mueller et al., 2020; Teymoor Seydi et al., 2025), posing increasing threats to ecosystems, communities, infrastructure, and regional economies in the United States and beyond (Abdollahi et al., 2023, 2024; AghaKouchak et al., 2025; Ermagun et al., 2025; Tao et al., 2025; Vahedifard et al., 2024). Several studies have documented this intensifying trend in the western United States (Salguero et al., 2020; Westerling, 2016; Westerling et al., 2006; Zhuang

Writing – review & editing: Masoud Zaerpour, Simon Papalexiou, Amir AghaKouchak, Farshid Vahedifard

et al., 2021), while emerging research confirms similar increases in fire size and frequency in the eastern United States (Brown et al., 2021; Donovan et al., 2023). Projections further suggest continued growth in fire activity across many regions under future climate scenarios (Brown et al., 2021; Liu et al., 2013; Prestemon et al., 2016). These impacts are spatially heterogeneous, driven by variations in vegetation, land use, climate, and management practices (Chelsea Nagy et al., 2018; Yang & Vahedifard, 2026). From 1984 to 2015, anthropogenic climate change contributed to an additional 10.3 million acres (4.2 million hectares) burned in western United States forests (Abatzoglou & Williams, 2016), while between 1985 and 2017, high-severity burned areas in the western United States increased eightfold (Parks & Abatzoglou, 2020).

These alarming trends highlight the need to better understand the statistical behavior of wildfire size distributions, especially their extremes, to support risk assessment, emergency response, and long-term resilience planning (Holmes et al., 2008). Despite their importance, the statistical characteristics of extreme wildfires remain insufficiently understood. Many existing analyses have examined burned area and frequency trends (e.g., Salguero et al., 2020; Taylor et al., 2013), rather than the statistical behavior of the rare but highly destructive extremes (Holmes et al., 2008). This emphasis can lead to systematic underestimation of the risks posed by exceptionally large wildfires, potentially leaving ecosystems, communities, and critical infrastructure inadequately prepared for these catastrophic events. The challenge of accurately modeling extremes extends well beyond wildfire science. It is a central problem in different fields such as natural hazards, actuarial science and finance (Embrechts et al., 1997; Klugman et al., 2012), geophysics, critical phenomena (Sornette, 2006), and hydrology (Papalexiou et al., 2013). Across these disciplines, failing to account for heavy-tailed behavior or misidentifying the underlying distribution can result in systematic bias and underestimation of risk, particularly for low-probability, high-impact events.

The mean or central tendency (the "body") of wildfire burned area distributions provides insights into typical fire behavior, such as identifying common fire sizes and frequencies that are critical for allocating firefighting resources, guiding prescribed burning practices, informing land-use planning decisions, and developing fire prevention programs. However, it is the tail of these distributions, representing the rare but catastrophic events, that holds the key to understanding extreme risk (Holmes et al., 2008). In other words, wildfires with unusually large burned areas, though infrequent, account for a disproportionately high percentage of the total area burned and can overwhelm suppression capabilities and mitigation infrastructure (Sadegh et al., 2025). These extreme (or "tail") events can also reshape ecosystems and cause multi-billion-dollar damage. In 2018 alone, large fires were responsible for more than $148 billion in damages in California (Wang et al., 2020). Extremely large wildfires (the largest of the large) are also responsible for increasing out-migration in regions affected by extreme wildfires (McConnell et al., 2024). These statistical tails have profound physical consequences. Analyzing the data of large fires in the continental United States from 1984 to 2024 (MTBS, 2025) shows that the top 1% of fire events in the data set account for 28.1% of the total burned area. Furthermore, the 41 annual exceedance events alone, representing just 0.14% of the total record, contribute 8.3% of the total area burned, highlighting the disproportionate impact of rare, extreme events on the national fire regime.

While extreme events are important, focusing only on extreme values or record-breaking events, without a realistic understanding of their likelihood and tail behavior, can also lead to a mischaracterization of fire risk. It is important to view these events within the broader context of their statistical and environmental patterns. Furthermore, existing research often overlooks the variability of wildfire extremes across different spatial scales. Comprehensive comparisons of extreme wildfire distributions at both ecological and administrative scales on a national level remain limited. Without such analyses, we might risk missing important regional differences in extreme fire characteristics across various geographic and ecological regions.

This study addresses the limitations of wildfire statistical modeling approaches that rely solely on central distribution metrics, which can misrepresent the likelihood and impact of large fire events. Here, we introduce a framework that explicitly distinguishes between the statistical characteristics of body and tail fire size distributions. We analyzed 30,331 large fire perimeters that occurred in the United States between 1984 and 2024. The data are disaggregated across various ecoregions and Geographic Area Coordination Centers (GACCs) to capture regional variability. For each spatial unit, we fit three widely used heavy-tailed distributions, Pareto Type II, lognormal, and Weibull, to two distinct samples: the full distribution of fire sizes and a subset representing annual exceedance events in the upper tail. Model performance is evaluated using a modified mean square error norm, complemented by tail-focused diagnostic tools, including quantile-quantile (QQ) and probability-probability (PP)

plots. Using recent major fires as benchmarks, we demonstrate that body-only models can substantially under- or over-estimate both the Annual Exceedance Probability (AEP) of extreme fire events and their associated burned areas. These findings have direct implications for wildfire risk analysis, emphasizing the importance of tail-focused models for predicting extreme fire events, resource allocation, and long-term resilience planning.

2. Background

Wildfire sizes (measured by area burned) are widely recognized as heavy-tailed (Cumming, 2001; Ricotta et al., 1999; Song et al., 2001; Malamud et al., 2005), and several theories have been proposed to explain their form. One framework is the concept of self-organized criticality (SOC), which posits that dynamic systems evolve toward a critical state characterized by scale-invariant behavior and power-law distributions (Bak et al., 1987; Holmes et al., 2008; Malamud et al., 1998). The classic forest fire model developed by Drossel and Schwabl (1992) illustrates this concept: as tree density increases or fire-weather becomes more frequent, the system approaches a threshold beyond which large-scale fires suddenly emerge, consuming large sections of the modeled landscape grid. Near this threshold, the system exhibits power-law behavior in fire sizes (Newman, 2005).

While SOC-inspired models have been extensively employed in the literature to explain why wildfire sizes follow power-law patterns (e.g., Bak et al., 1987; Malamud et al., 1998; Pueyo, 2007; Ricotta et al., 2001), they have important limitations in representing real-world fire dynamics. The classic Drossel-Schwabl model is not strictly self-organized; it reaches a power-law regime only if ignition and growth parameters are finely tuned within a narrow range (Palmieri & Jensen, 2020). Outside this range, the model produces either mostly small fires or one large burn, failing to reproduce realistic heavy tails. In contrast, actual fire systems are shaped by external and non-stationary drivers. Fuel loads vary seasonally and climatically. Ignition patterns are non-random, influenced by power lines, human activity, suppression efforts, and episodic extreme weather. In this sense, wildfire systems are not autonomously critical but are instead "tuned" by external forces. As a result, the heaviness of wildfire size distributions can vary significantly across regions and over time.

Previous research has examined wildfire distributions using various statistical approaches. Malamud et al. (1998) analyzed data from Alaskan boreal forests, the western United States, and the Australian Capital Territory, identifying power-law behavior with exponents around 1.3 to 1.5. Similarly, Ward et al. (2001) found a power-law pattern in 25 years of wildfire data from Ontario's boreal forest, highlighting that fire management intensity can influence tail heaviness. Moritz (1997) employed Generalized Extreme Value (GEV) distributions, particularly the heavy-tailed type, to model annual maximum fire sizes in Los Padres National Forest in California, emphasizing that extreme events significantly influence total burned area. More recent comparative studies have also tested alternative distributions, including lognormal and Weibull, finding these distributions often provide a superior fit across a wider range of fire sizes, with power-law behavior holding at most over limited size ranges (Hantson et al., 2016; Reed & McKelvey, 2002).

Given the limitations of SOC-based theories, there is a clear need for empirically grounded parametric models that can flexibly describe wildfire size distributions and quantify the likelihood of extreme events. Several heavy-tailed distributions have been proposed for this purpose, most notably the Pareto (power-law), lognormal, and Weibull families. Holmes et al. (2008), for example, compared Pareto and lognormal fits for large wildfire events and found both captured the heavy-tailed nature of the data. Other studies show that in certain regimes fire sizes may deviate from pure power-law behavior. For example, Cumming (2001) found a truncated Pareto fit wildfire data in Canadian forests, while Schoenberg et al. (2003) introduced a tapered Pareto for Southern California, which incorporates a faster decay in the far tail, recognizing that even heavy-tailed phenomena are constrained by physical limits. Collectively, these studies suggest that empirical models can better account for the complexity of real-world fire regimes, including finite-size effects, changing climate conditions, and management interventions, than idealized SOC-based frameworks. However, comparative evaluations across space and scale remain limited.

Finally, the applied relevance of tail-focused modeling is clear. Statistical models that focus on extremes can directly inform how agencies anticipate and prepare for the rare events that drive the majority of suppression costs and damages. Generating exceedance probability curves at ecoregion and GACC levels provides managers with benchmarks for resource allocation, contingency budgeting, and insurance modeling. Previous work has demonstrated the utility of partitioning wildfire behavior by administrative units: Salguero et al. (2020) document burn-area trends within GACC regions, and integrated risk frameworks in the United States use spatial units to translate fire likelihood into management priorities (Scott et al., 2013; Thompson et al., 2011). For fire modeling

10.1029/2025EF007485

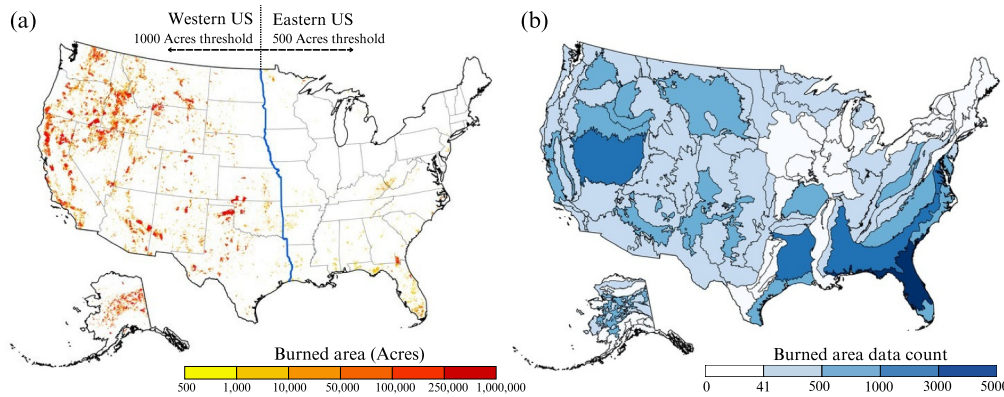

Figure 1. Overview of the spatial distribution and frequency of large fire events in the United States from 1984 to 2024. (a) Location and size of large burned areas across the continental United States, based on Monitoring Trends in Burn Severity data. (b) Number of large fire events recorded in each Level III ecoregion, highlighting regional variation in fire occurrence.

in the continental United States, Joseph et al. (2019) employed Bayesian approaches to characterize size extremes across spatial units, showing the importance of modeling heterogeneity in tail behavior across regions. Landscape simulation models (Finney et al., 2011) further tie extreme fire risk to spatial fuel and weather distributions. More recent spatial modeling frameworks (e.g., Cisneros et al., 2023; Zhou & Bradley, 2026) highlight the value of multiscale and spatial dependence methods for capturing regional variation in fire behavior. In the Southeastern US, spatially resolved Bayesian risk models have also been used to link fire risk to socio-ecological vulnerability (Nepal et al., 2023). By grounding policy and practice in empirically calibrated tail behavior at relevant spatial units, decision-makers can avoid the systematic under- or over-estimation of risk that arises from body-focused models.

Together, this background highlights both the theoretical and empirical progress that has been made in characterizing wildfire extremes, as well as the remaining gaps. Building on these insights, the present study applies heavy-tailed models to a comprehensive national data set, explicitly focusing on the tails of wildfire size distributions across ecoregions and GACCs to evaluate their performance and implications for risk management.

3. Data

The primary data set used in this study is sourced from the Monitoring Trends in Burn Severity (MTBS) project (MTBS, 2025; https://www.mtbs.gov/). MTBS is a joint interagency initiative designed to consistently map the extent and severity of large fires across the United States from 1984 to 2024. It is important to note that the MTBS record for 2024 was incomplete at the time of analysis. While some 2024 fire perimeters were available and included, the MTBS database is updated regularly as new fires are mapped and existing perimeters are refined. All analyses in this study were conducted using the MTBS data set accessed on 25 February 2025, and results reflect the database status as of that time. Given that the analysis emphasizes long-term distributional structure and tail behavior over four decades, the inclusion of partial 2024 observations does not materially affect the fitted distributions or the study conclusions.

The MTBS program provides comprehensive coverage of large fire events for the conterminous United States, Alaska, Hawaii, and Puerto Rico, defined as those exceeding 1,000 acres (≈404 ha) in the western United States and 500 acres (≈202 ha) in the eastern United States (Figure 1a). As a result, the analyses presented in this study describe the statistical behavior and tail properties of the large-fire population captured by MTBS, rather than the full population of all wildfire ignitions, including small fires. To assign fire events to spatial units, we clipped each fire perimeter to ecoregion boundaries. As a result, any single fire event spanning multiple ecoregion boundaries was split into separate polygons. Consequently, the statistical distributions analyzed in this study represent the within-region burned area for each spatial unit, rather than the total size of the original whole-fire event. Aggregated burned-area totals across regions therefore reflect the sum of clipped portions and should not be interpreted as unique national burned area.

To investigate spatial variability in fire behavior, the analysis is conducted at two spatial scales. The first is the Level III Ecoregion framework, developed by the United States Environmental Protection Agency (EPA) in collaboration with state and international partners (EPA, 2025; Omernik & Griffith, 2014). These regions delineate areas with broadly similar ecological characteristics, including patterns in climate, geology, vegetation, soils, hydrology, and land use. The study area is divided into 105 such ecoregions, which provide an ecologically grounded basis for assessing wildfire characteristics. Figure 1b displays these ecoregions alongside the number of large fire events observed in each. Of these 105 ecoregions, 100 contained at least one large fire in the MTBS record.

The second scale follows the GACC boundaries used by the National Interagency Coordination Center (2025). The conterminous United States and Alaska are segmented into 10 GACC regions, each coordinated by an interagency group representing federal and state land management agencies. These administrative zones serve as operational units for wildfire management, enabling strategic coordination of firefighting resources and response planning. GACC-level analysis provides a complementary perspective for evaluating wildfire risk and response capacity across broader administrative regions.

4. Methodology

The methodology consists of three core steps: distribution fitting, parameter estimation, and goodness-of-fit evaluation. These steps are performed independently for two subsets of the data: the full set of fire sizes to characterize the central body of the distribution, and a subset of the largest fire events, to examine tail behavior. We evaluate three widely used heavy-tailed parametric distributions: the Pareto Type II distribution, the lognormal distribution, and the Weibull distribution (which exhibits heavy-tailed behavior for certain shape parameter values). For the full-sample analysis, each distribution is fitted in its three-parameter form; for the tail analysis, a two-parameter form is used. The estimation procedures differ between the two analyses, as described below.

4.1. Distribution Definitions

Each distribution is first defined on the domain $x \geq x_0$, where x_0 is the location parameter. This added flexibility allows the models to shift the support of the distribution to reflect regional variability in minimum fire sizes, improving alignment with real-world data thresholds and fire behavior.

The Pareto Type II distribution is a power-law distribution with the following Probability Density Function (PDF) and Exceedance Probability Function (EPF):

$$f_{\mathrm{PII}}(x) = \frac{1}{\beta}\left(1 + \gamma\frac{x - x_0}{\beta}\right)^{-\frac{1}{\gamma}-1}, x \geq x_0 \tag{1}$$

$$E_{\mathrm{PII}}(x) = \left(1 + \gamma\frac{x - x_0}{\beta}\right)^{-\frac{1}{\gamma}} \tag{2}$$

where $x_0 > 0$ is the location parameter, representing the minimum fire size for the distribution support, $\beta > 0$ is the scale parameter, controlling the spread of the distribution, $\gamma \geq 0$ is the shape (tail) parameter, determining the heaviness of the tail. Larger values of γ result in heavier tails, indicating more frequent large fires. The mean and variance are undefined for $\gamma \geq 1$ and $\gamma \geq 0.5$, respectively.

The PDF and EPF of the lognormal distribution, often used to model multiplicative processes, are defined by the following expressions:

$$f_{\mathrm{LN}}(x) = \frac{1}{\sqrt{\pi}\,\gamma\,(x - x_0)}\exp\left(-\ln^2\left(\frac{x - x_0}{\beta}\right)^{\frac{1}{\gamma}}\right), x \geq x_0 \tag{3}$$

$$E_{\mathrm{LN}}(x) = \frac{1}{2}\mathrm{erfc}\left(\ln\left(\frac{x - x_0}{\beta}\right)^{\frac{1}{\gamma}}\right) \tag{4}$$

where erfc() function is the complementary error function, $x_0 > 0$ is the location parameter, shifting the distribution's support to the right and $\beta > 0$ the scale parameter. The shape, or tail-thickness, parameter $\gamma \geq 0$ governs how widely the shifted lognormal tail spreads: a larger γ stretches the right-hand tail, so extremely large fires become comparatively more common and the tail decays more slowly. Unlike the Pareto Type II distribution, all ordinary moments of the shifted lognormal exist for every positive γ. Specifically, the mean and variance grow exponentially with γ.

The Weibull distribution, widely applied in reliability and survival analysis, is expressed as:

$$f_W(x) = \frac{\gamma}{\beta}\left(\frac{x - x_0}{\beta}\right)^{\gamma-1} exp\left(-\left(\frac{x - x_0}{\beta}\right)^{\gamma}\right), x \geq x_0 \tag{5}$$

$$E_W(x) = \exp\left(-\left(\frac{x - x_0}{\beta}\right)^{\gamma}\right) \tag{6}$$

where $x_0 > 0$ is the location parameter representing the minimum support, $\beta > 0$ is the scale parameter, controlling the distribution's spread, $\gamma > 0$ is the shape parameter, which governs tail behavior. When $\gamma < 1$ the distribution has a heavy tail; when $\gamma = 1$, it becomes exponential, and when $\gamma > 1$, the tail is lighter and decays more quickly.

In the next step, we define the same set of three distributions: Pareto Type II, lognormal, and Weibull, each in their two-parameter form by setting the location parameter $x_0 = 0$, leaving each model characterized solely by a scale parameter β and a shape parameter γ. Two-parameter forms that exclude the location term are expected to yield less flexible fits and result in higher modeling errors (Bølviken & Hobæk Haff, 2024). Using distributions with the same number of parameters avoids any misinterpretation due to differences in model complexity (i.e., varying numbers of parameters) (Burnham & Anderson, 2002).

4.2. Parameter Estimation

Parameters for the three-parameter distributions are estimated by maximum likelihood (ML). Because the log-likelihood surface may contain multiple local optima, we employ a multi-stage random search procedure to increase the probability of locating the global maximum. In the first stage, 200,000 candidate parameter vectors are drawn uniformly at random from broad initial bounds and at the end, each candidate is evaluated by computing the negative log-likelihood (NLL) over the full data set, and the candidate with the lowest value is retained for refinement.

For the tail analysis, to isolate tail behavior, we adopt an annual-exceedance (partial-duration) approach (Gupta, 2011; Papalexiou et al., 2013). For each spatial unit, a threshold is selected that, on average, one fire per year exceeds it. Given the 41-year MTBS record (1984–2024), this corresponds to selecting the 41 largest fire events from each spatial unit. The resulting sample of $N = 41$ exceedances form the tail data set. Of the 105 Level III ecoregions present in the data set, 70 contained at least 41 large fire events after filtering and were retained for distributional analysis. All 10 GACC regions met this threshold. This approach retains all substantial fire events, regardless of whether multiple occur in a single year, and thus provides a more accurate representation of the parent distribution's upper tail than traditional methods that retain only a single annual maximum.

For the tail sample, the empirical exceedance probability is computed using Weibull plotting positions referenced to the full data set size. If $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(N)}$ are the N largest observations in ascending order, drawn from a full sample of N_{total} fires, the empirical exceedance probability is defined as:

$$E_{\text{emp}}\left(x_{(i)}\right) = 1 - \frac{r_i}{N_{\text{total}} + 1} \tag{7}$$

where r_i is the rank of observation $x_{(i)}$ within the full data set of N_{total} observations sorted in ascending order (i.e., $r_i = N_{\text{total}} - N + i$ for $i = 1,\ldots,N$). Referencing probabilities to the full data set ensures that each exceedance reflects its rarity relative to the entire fire record.

Unlike the full-sample analysis, which employs maximum likelihood, parameters for the two-parameter tail distributions are estimated by directly minimizing the N_1 norm, a scale-free, probability-based error metric:

$$N_1 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{E_{\text{the}}\left(x_{(i)}\right)}{E_{\text{emp}}\left(x_{(i)}\right)} - 1\right)^2 \tag{8}$$

where $E_{\text{the}}\left(x_{(i)}\right)$ is the theoretical exceedance probability at the ith order statistic and $E_{\text{emp}}\left(x_{(i)}\right)$ is the empirical exceedance probability from Equation 7. This approach directly minimizes discrepancies in probability space, prioritizing accurate representation of tail behavior. The optimization uses a multi-stage random search procedure, with β drawn from a log-uniform distribution to ensure adequate coverage across orders of magnitude. Search bounds are progressively narrowed around the current best solution at each stage.

4.3. Error Metric Selection

This N_1 criterion is particularly well suited to modeling wildfire extremes for three key reasons. First, it is dimensionless, representing a relative error between theoretical and empirical exceedance probabilities, and is therefore invariant to the units used for burned area (e.g., acres, hectares, or square kilometers). Second, the criterion operates in probability space rather than area space, preventing the largest fires from dominating the fit. Burned area sizes can span three to four orders of magnitude, from as small as 500 acres (≈202 ha) to more than 1 million acres (≈404,686 ha); minimizing a term such as $(x_i - \hat{x}_i)^2$ would disproportionately emphasize the largest events and distort the fit for the remainder of the tail. The ratio-based formulation of N_1 is particularly sensitive to relative discrepancies at low exceedance probabilities, where accurate modeling of extreme event frequencies is most critical for risk assessment. Third, the N_1 metric aligns directly with decision-relevant quantities such as exceedance probabilities and return periods, which are commonly used in both wildfire risk analysis and practical planning frameworks. Because return period is simply the reciprocal of the exceedance probability, minimizing N_1 directly optimizes the quantity that practitioners rely on for risk estimation and resource allocation.

The choice of different estimation procedures for the full-sample and tail analyses reflects their distinct objectives. Maximum likelihood is a statistically efficient estimator that leverages the full information content of the data and is well suited for characterizing the overall distributional form. For the tail analysis, however, the goal is to match the observed exceedance structure as closely as possible. Minimizing N_1 directly targets this objective, producing parameter estimates that are optimized for exceedance probability accuracy rather than overall likelihood.

5. Results

5.1. Body Distribution Analysis

Table 1 reports L-moment statistics: L-location (λ_1), L-scale (λ_2), and L-skewness (τ_3), for fire burned areas computed across 100 Level III ecoregions (that have at least 1 large fire) and 10 GACC regions. L-moments are used in this study because they provide a more robust and informative summary of distributional characteristics, particularly when dealing with heavy-tailed data like wildfire burned areas. Unlike traditional moments (e.g., mean, variance, skewness), which can be heavily influenced by extreme values, L-moments are based on linear combinations of order statistics (Vogel & Fennessey, 1993), making them more resistant to the influence of outliers. This results in more stable and interpretable estimates of location, scale, and shape. L-moments are therefore particularly well-suited for analyzing data characterized by large variability and scale differences (Vogel et al., 2025), such as wildfire sizes spanning multiple orders of magnitude across diverse geographic regions.

The results reveal higher central tendency at the GACC scale compared to individual ecoregions. The mean L-location increases from approximately 7,876 acres (≈3,187 ha) at the ecoregion scale to nearly 9,990 acres (≈4,043 ha) at the GACC scale. A similar pattern is observed in dispersion: the mean L-scale rises from 5,212 acres (≈2,109 ha) to 6,942 acres (≈2,809 ha), indicating that larger spatial units encompass a broader range of fire sizes. L-skewness values confirm the dominance of extreme events at both scales, though the effect becomes more uniform at the GACC level. The median L-skewness increases from 0.64 to 0.70, and inter-regional variability decreases (standard deviation drops from 0.16 to 0.03), though this reduction is partly attributable to the smaller

Table 1
Summary of L-Moment Statistics for Burned Areas Clipped to Level III Ecoregions and Geographic Area Coordination Center Regions of the United States

	Ecoregion level III			GACC regions		
	L-location (λ_1, acres)	L-scale (λ_2, acres)	L-skewness (τ_3)	L-location (λ_1, acres)	L-scale (λ_2, acres)	L-skewness (τ_3)
Min	895.26	215.76	0.12	2368.35	1357.92	0.62
Q5	1065.71	291.35	0.33	2565.28	1481.19	0.65
Q25	2129.43	981.28	0.51	6647.16	4257.37	0.69
Q50 (Median)	6113.55	4024.51	0.64	9313.90	6489.11	0.70
Q75	11052.68	7145.43	0.71	11598.97	8124.04	0.71
Q95	22555.28	16464.83	0.79	20249.28	14567.22	0.74
Max	38572.79	26276.47	0.98	22721.60	15710.63	0.74
Mean	7875.92	5212.47	0.60	9988.70	6942.20	0.70
SD	6954.41	5123.70	0.16	5923.66	4373.35	0.03

number of GACC units. This suggests that while GACCs consistently exhibit heavy right tails, individual ecoregions vary widely, with some exhibiting strongly skewed distributions and others showing more moderate skewness due to fewer large fires.

Table A1 presents the fitted parameters for the three candidate distributions: lognormal, Pareto Type II, and Weibull, across both spatial aggregation levels. To identify the best-fitting model in each region based on the complete fire burned-area record, we selected the distribution that minimized the NLL. This evaluation was conducted separately for (a) Level III ecoregions and (b) GACCs, with results visualized in Figure 2. In the figure, colors indicate the distribution with the best fit: green for lognormal, blue for Pareto Type II, and red for Weibull. Regions shaded in gray had fewer than 41 qualifying fires (as defined in Section 4.2) and were therefore excluded from the model comparison.

At the Level III ecoregion scale (Figure 2a), a distinct spatial pattern emerges. The lognormal distribution (green) dominates much of the western United States, including part of California, the Great Basin, the northern and central Rockies, and the Pacific Northwest, as well as large portions of Alaska. In these regions, the lognormal distribution effectively captures the curvature of the burned-area distribution. The Weibull distribution (red) is preferred across portions of the upper Midwest, and parts of the Interior Plateau and above, where fires are generally smaller but exhibit considerable variability under moderately moist fuel conditions; its shape flexibility provides a better fit in these settings. The Pareto Type II distribution (blue) emerges as the best fit in scattered regions, particularly across the central and southern Great Plains and parts of the Desert Southwest, where

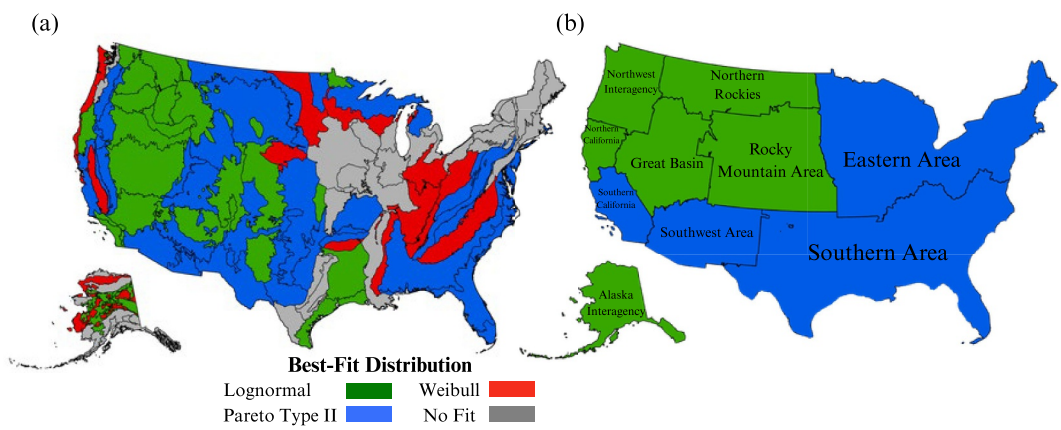

Figure 2. Spatial distribution of the best distribution fits for the entire observations for (a) ecoregion level III and (b) Geographic Area Coordination Center regions.

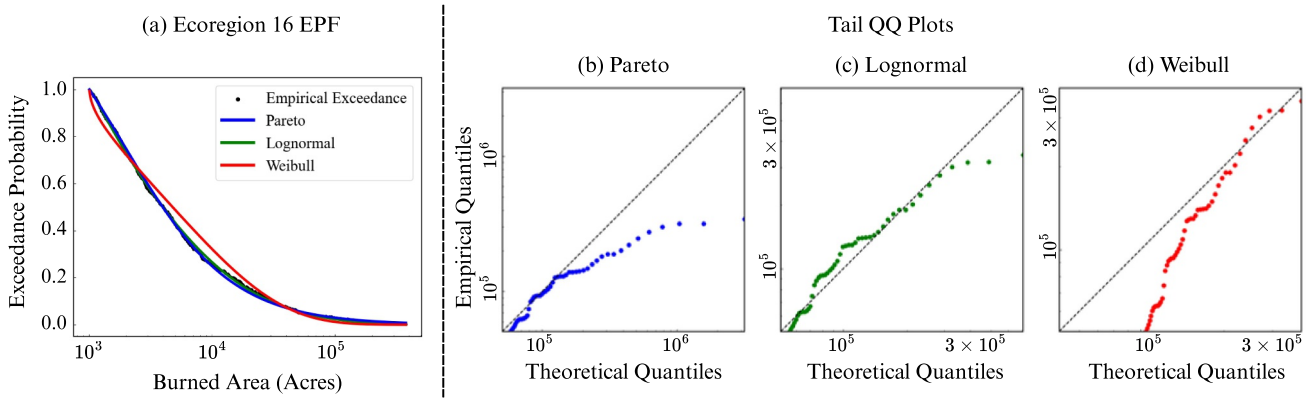

Figure 3. Illustration of the mismatch between full-distribution fits and tail performance using the Idaho Batholith ecoregion (Level III, code 16) as a representative example. The figure demonstrates how candidate distributions that fit the overall burned area distribution well may still fail to capture the behavior of extreme wildfire events in the upper tail. (a) Empirical exceedance probability function (EPF; black dots) overlaid with body-fitted EPFs from the Pareto Type II (blue), lognormal (green), and Weibull (red) distributions on a log-log scale. (b) Tail-specific QQ plot for the Pareto Type II distribution, showing overestimation of the most extreme events (points below the 1:1 line). (c) Tail-specific QQ plot for the lognormal distribution, revealing an S-shaped deviation, with underestimation of intermediate extremes and overestimation of the largest fires. (d) Tail-specific QQ plot for the Weibull distribution, indicating overestimation of exceedance probabilities throughout the upper tail.

heavier-tailed distributions are more representative of local fire size behavior. A substantial number of ecoregions in the central and eastern United States (gray) were excluded due to insufficient data.

When the analysis is repeated at the coarser GACC scale (Figure 2b), the regional variability observed at the ecoregion level gives way to a more uniform pattern. The lognormal distribution provides the best fit in 6 of the 10 GACCs, Northwest Interagency, Northern Rockies, Great Basin, Northern California, Rocky Mountain Area, and Alaska Interagency, all located in the western and interior United States. The Pareto Type II distribution is selected in the remaining 4 GACCs: Eastern Area, Southern Area, Southern California, and Southwest. These regions span a range of fire regimes, from the fire-prone chaparral landscapes of Southern California to the broad fuel-connected environments of the eastern and southern United States, where large fires, though less frequent, can reach substantial sizes and produce heavier-tailed distributions. Notably, no GACC is best characterized by the Weibull distribution once local variability is averaged out at this coarser spatial scale.

These findings yield two key insights. First, spatial scale significantly influences model selection: at finer ecological resolutions, the best-fitting distribution varies meaningfully with regional differences in climate, fuel characteristics, and topography, and all three candidate distributions appear as best fits across different ecoregions. At broader administrative scales, however, the competition narrows to two models, lognormal in the western interior and Pareto Type II in the southern, eastern, and Southern California GACCs. Second, the geographic pattern of heavier-tailed body distributions at the GACC scale suggests that Pareto Type II behavior is not confined to a single fire regime but emerges in regions with diverse fuel structures and fire drivers, from the shrub-dominated Southwest to the expansive fuel-connected landscapes of the East and South. This observation is explored further in the following tail analysis.

Figure 3 highlights a single representative Level III ecoregion, the Idaho Batholith (code 16), to demonstrate why a model that fits the full distribution well may still fail to capture extreme tail behavior. Figure 3a overlays the empirical exceedance probability function (EPF, black dots) with the theoretical EPFs of the three body-fitted distributions. On a log-log scale, both the Pareto Type II (blue) and lognormal (green) distributions follow the empirical curve closely across nearly four orders of magnitude. Based on conventional goodness-of-fit metrics applied to the entire data set, both would be considered acceptable representations of fire size behavior.

However, Figures 3b through 3d zoom in on the upper 5% of fire sizes, those between the 95th and 99.9th percentiles, using tail-focused QQ plots. These plots reveal important differences in how each distribution performs in the extreme right tail. The Pareto Type II distribution (Figure 3b) performs reasonably well for moderately large fires (up to approximately 10^5 acres (≈40,469 ha) but overestimates the likelihood of the largest events, as evidenced by points falling below the 1:1 reference line. In contrast, the lognormal distribution

Table 2
Summary Statistics From the Fitting of the Three Distribution Into the Tail Samples of Burned Area (in Acres) for Ecoregion Level III

	Pareto type II			Lognormal			Weibull		
	N_1	β	γ	N_1	β	γ	N_1	β	γ
Min	0.003	115.6	0.001	0.004	6.1	0.796	0.005	0.6	0.160
Mode	0.018	1942.0	0.653	0.016	1468.4	2.269	0.017	1424.4	0.387
Mean	0.025	3649.2	0.553	0.024	2663.2	2.124	0.028	2999.2	0.528
Median	0.021	2275.4	0.598	0.021	1561.5	2.178	0.022	1658.0	0.442
Max	0.078	22831.1	1.139	0.092	19406.1	4.461	0.107	21795.9	1.319
SD	0.017	4191.5	0.281	0.017	3516.5	0.736	0.019	4200.6	0.243
Skew	1.366	2.9	−0.343	1.707	3.3	0.641	1.672	2.9	1.271

(Figure 3c) underestimates intermediate tail quantiles (points above the 1:1 line), then overestimates the most extreme events. The Weibull distribution (Figure 3d) exhibits the poorest tail fidelity, consistently overestimating exceedance probabilities across most of the tail.

This mismatch between good full-sample fit (Figure 3a) and poor tail performance (Figures 3b–3d) underscores the importance of evaluating fire models using tail-specific diagnostics. As a result, the remainder of this study shifts its focus to modeling extreme-size behavior directly by fitting the three candidate distributions to annual-exceedance (top-N) tail samples for each Level III ecoregion and GACC region. This approach isolates the dynamics of the most consequential fire events and enables spatially explicit comparisons of which distributions most reliably capture regional wildfire extremes.

5.2. Tail Distribution Analysis

Table A2 summarizes the 41-year tail samples constructed by selecting the 41 largest fire events from each spatial unit, corresponding to an average of one exceedance per year over the 1984–2024 MTBS record. Each row reports the distribution of a given descriptive statistic across all ecoregions or GACCs; for example, the "Mean" row gives the average value of each statistic across spatial units. At the Level III ecoregion scale, the average median burned area is approximately 21,868 acres (≈8,850 ha), the average mean fire size is roughly 38,852 acres (≈15,723 ha), and the average standard deviation is 40,213 acres (≈16,274 ha). These values underscore the high variability and heavy-tailed nature of extreme wildfire sizes even at localized ecological scales.

When aggregated to the 10 GACC regions, all descriptive metrics increase substantially. The average median tail fire size rises to nearly 153,668 acres (≈62,183 ha), the average mean fire size increases to about 147,318 acres (≈59,617 ha), and the average standard deviation exceeds 62,925 acres (≈25,462 ha). Notably, skewness declines at the GACC scale, with a mean value of −0.39 compared to 1.31 at the ecoregion level. This indicates that, within aggregated regions, extreme fires are less pronounced as outliers and instead reflect a more uniformly distributed upper tail.

These results highlight a key insight: spatial scale strongly influences the distributional properties of wildfire extremes. The average median tail fire size at the GACC level is roughly seven times larger than at the ecoregion level. This scale dependence reinforces the limitations of fitting distributions to the full data set without explicitly modeling the upper tail and motivates the targeted tail analysis presented in the remainder of this section.

Table 2 summarizes the performance of two-parameter tail fits applied across 70 Level III ecoregions. Among the three candidate distributions, all three achieve broadly similar mean N_1 values: lognormal (0.024), Pareto Type II (0.025), and Weibull (0.028). While the lognormal distribution achieves the lowest mean N_1, the margins are narrow, indicating competitive performance across all three models. The Pareto Type II distribution achieves the lowest single-region N_1 (0.003), demonstrating strong performance in specific ecoregions, though its N_1 values are right-skewed (≈1.37), indicating less consistent performance across all regions. The Weibull distribution produces the highest mean error and the most right-skewed N_1 values (≈1.67), suggesting it is the least reliable of the three candidates for tail modeling.

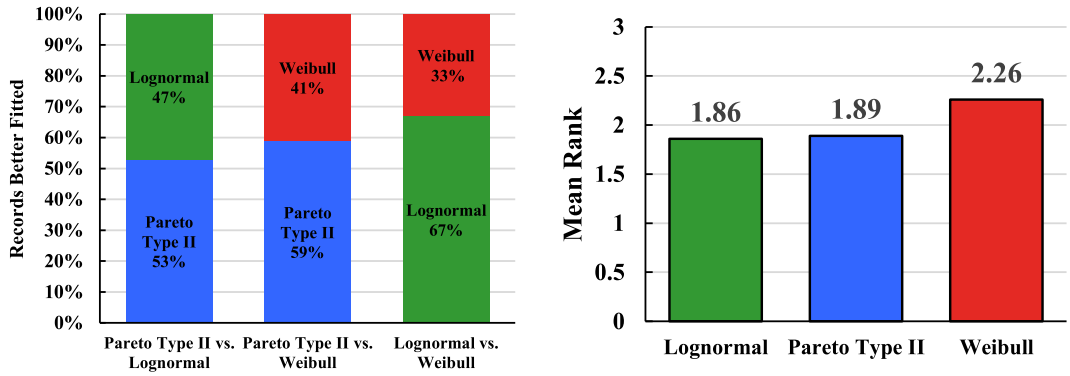

Figure 4. Tail model performance across 70 Level III ecoregions. (left) Pairwise comparisons of tail model performance in terms of N_1. For each pair, the percentage of ecoregions where each distribution achieves the lower N_1 is shown. (right) Mean ranks of the tails for all records. The best-fitted tail is ranked as 1, while the worst-fitted is 3. A lower average rank indicates better performance.

Parameter estimates further illustrate differences in tail behavior across the three distributions. The Pareto Type II shape parameter (γ) has a median of 0.60 and a mode of 0.65, indicating heavy tails across most ecoregions. In a number of regions, the shape parameter falls below 0.5, implying infinite variance, though the majority of ecoregions have shape values above this threshold. The lognormal shape parameter (γ) centers around a median of 2.18, with relatively moderate variability (SD = 0.74), indicating stable and consistently heavy-tailed fits across diverse regions. The Weibull shape parameter (γ) has a median of 0.44 and a mode of 0.39, both well below 1, confirming that the Weibull fits are capturing heavy-tailed behavior rather than exponential or light-tailed decay. Despite this, the Weibull's slightly higher N_1 values suggest that its stretched-exponential tail form is less flexible than the power-law (Pareto) or log-multiplicative (lognormal) alternatives for representing wildfire extremes.

Figure 4 provides a model-comparative summary of tail performance across all 70 Level III ecoregions that are included in the analysis. Figure 4a displays pairwise comparisons of N_1 between candidate tail models. For each ecoregion, we identify which of the two distributions achieves the lower N_1. In the Pareto Type II versus lognormal comparison, the Pareto Type II outperforms the lognormal in 53% of ecoregions, with the lognormal prevailing in the remaining 47%, a near-even split indicating that neither model consistently dominates the other. The Pareto Type II versus Weibull comparison shows a wider margin, with the Pareto Type II preferred in 59% of cases versus 41% for the Weibull. The lognormal versus Weibull comparison yields a clearer result: the lognormal is preferred in 67% of ecoregions, indicating stronger overall performance relative to Weibull.

Figure 4b summarizes these outcomes using a ranking-based approach. Each ecoregion assigns rank 1 to the best-fitting model (lowest N_1), rank 2 to the second-best, and rank 3 to the worst. Averaged across all regions, the lognormal distribution achieves a slightly lower mean rank (1.86) than the Pareto Type II (1.89), despite the Pareto Type II winning more head-to-head comparisons. This apparent paradox reflects the consistency of the lognormal's performance: when it does not achieve the best fit, it nearly always ranks second, rarely falling to third place. The Pareto Type II, by contrast, exhibits more variable performance; it wins outright in more ecoregions but also ranks third more frequently, which raises its average rank. The Weibull distribution has the highest mean rank (2.26), indicating that while it occasionally provides competitive fits, it is most often the weakest of the three candidates for tail modeling. Together, these results suggest that the Pareto Type II and lognormal distributions offer broadly comparable tail performance, with the Pareto Type II providing the single best fit in slightly more regions and the lognormal delivering more consistently strong fits across the full set of ecoregions. The Weibull distribution's stretched-exponential form is generally less suited to capturing the heaviness of large fires tail distributions.

Figure 5 maps the best-fitting tail distribution for each region: green indicates the lognormal, blue represents the Pareto Type II, red shows the Weibull, and gray denotes regions with fewer than 41 tail observations. At the Level III ecoregion scale (Figure 5a), no single distribution dominates as clearly as in the full-sample analysis. The Pareto Type II distribution (blue) emerges as the best-fitting tail model across much of the Great Plains, parts of the northern and middle Rockies, and scattered regions in the Desert Southwest and eastern seaboard, where the most extreme fires produce particularly heavy upper tails that favor a power-law form. The Weibull distribution

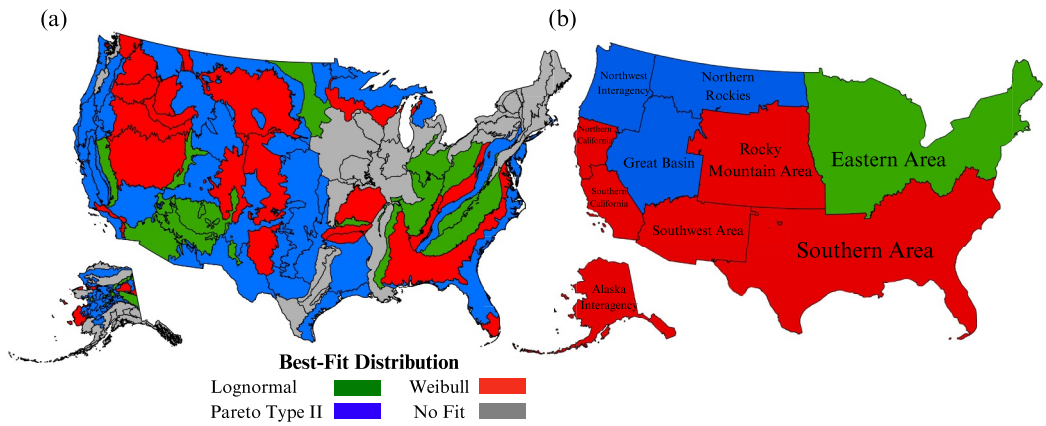

Figure 5. Spatial distribution of the best-fitting tail distribution for (a) Level III ecoregions and (b) Geographic Area Coordination Center regions.

(red) is preferred across a substantial number of ecoregions, including portions of the southern Rockies, the Great Basin interior, parts of the Southeast, and several Alaskan ecoregions. In these areas, the tail, while still heavy, decays at a rate better captured by the Weibull's stretched-exponential form. The lognormal distribution (green) provides the best tail fit in a smaller but geographically diverse set of ecoregions, primarily in the eastern United States, including portions of the Appalachians, the Interior Low Plateaus, and the Piedmont, as well as a few western ecoregions such as the Sierra Nevada and Wasatch-Uinta Mountains.

At the broader GACC scale (Figure 5b), the heterogeneity observed at the ecoregion level consolidates into a clearer geographic pattern. The Weibull distribution emerges as the best-fitting tail model in 6 of the 10 GACCs, Rocky Mountain Area, Southern Area, Southwest, Southern California, Northern California, and Alaska Interagency, spanning a wide range of fire regimes. The Pareto Type II distribution is selected in three GACCs, Northwest Interagency, Northern Rockies, and Great Basin, all in the northwestern interior, where the most extreme fires produce tails heavy enough to require a power-law representation. The lognormal distribution provides the best tail fit in only the Eastern Area GACC. This contrasts sharply with the full-sample analysis (Figure 2b), where the lognormal dominated 6 of 10 GACCs. The shift reflects a key finding: distributions that best characterize the overall body of fire sizes do not necessarily provide the best representation of the extreme upper tail.

Together, Figures 4 and 5 reinforce this conclusion. While the lognormal and Pareto Type II distributions perform comparably at the ecoregion level in terms of mean rank (Figure 4b), the spatial mapping reveals that their strengths are geographically distinct. The Pareto Type II excels in the northwestern interior where tail heaviness is most pronounced, the Weibull captures tail behavior across the broadest range of GACCs, and the lognormal, despite its strong full-sample performance, is less frequently the best tail model at both spatial scales. These results highlight the importance of evaluating tail behavior separately from the distributional body, as the optimal model can differ substantially between the two.

Table A3 summarizes the fitted parameters and model performance for the three tail distributions across all 10 GACCs. The Weibull distribution achieves the lowest mean N_1 (0.015) with the smallest variability (SD = 0.006), indicating the most consistent tail performance at this spatial scale. The lognormal distribution ranks second (mean $N_1 = 0.016$, SD = 0.010), followed by the Pareto Type II (mean $N_1 = 0.019$, SD = 0.013). These results are consistent with Figure 5b, where the Weibull provides the best fit in 6 of 10 GACCs, the Pareto Type II in 3, and the lognormal in 1.

Parameter estimates shed light on why the Weibull performs well at this scale. The Weibull shape parameter (γ) has a median of 0.36, well below 1, confirming that the fitted Weibull tails are heavy rather than exponential. This contrasts with the full-sample analysis, where Weibull shape values were higher, and suggests that when fitting is restricted to extreme events, the Weibull's stretched-exponential form is flexible enough to capture heavy tail behavior. The Pareto Type II shape parameter (γ) has a median of 0.51, with all values falling between 0.29 and 0.76, indicating consistently heavy power-law tails. The lognormal shape parameter (γ) centers around a median of 2.21, with moderate variability (SD = 0.65).

10.1029/2025EF007485

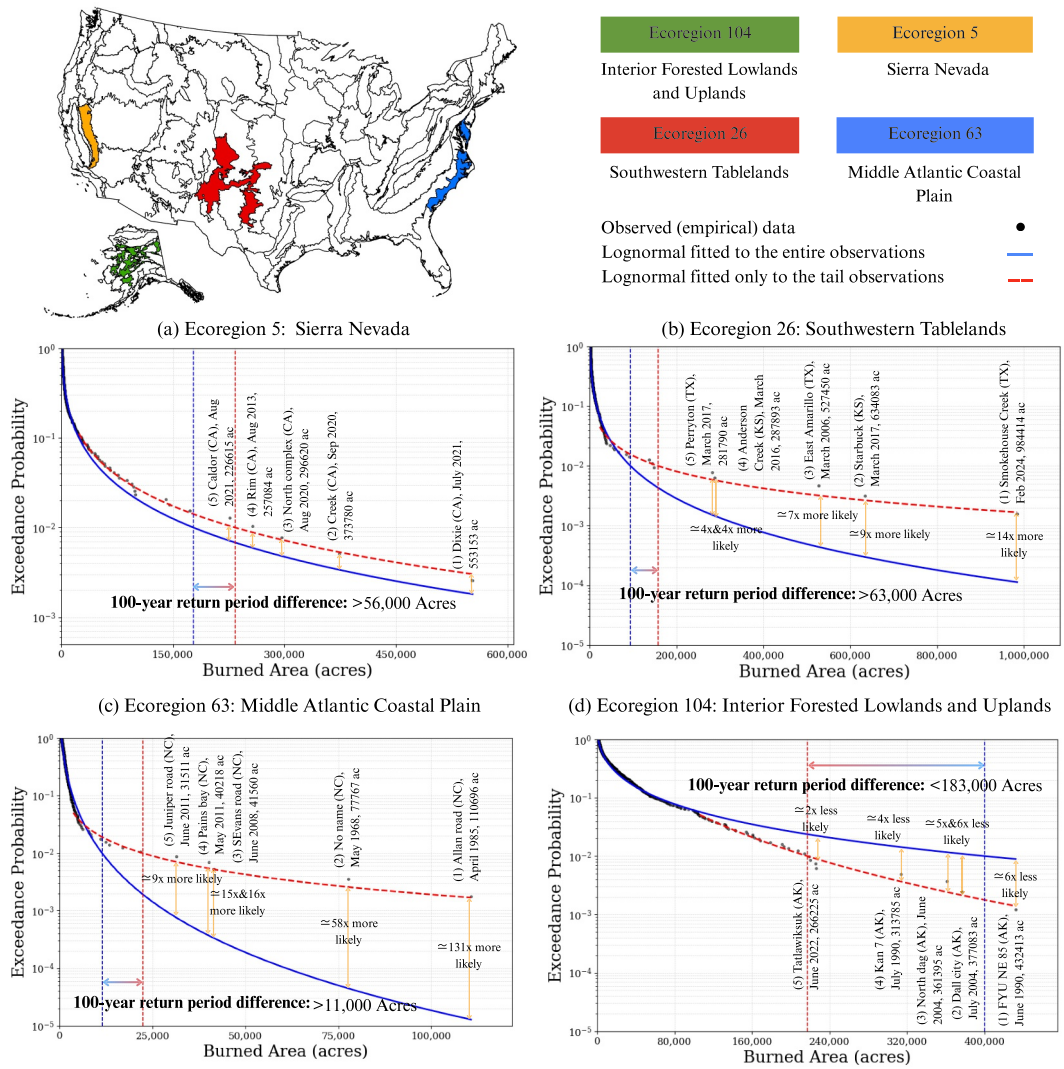

Figure 6. Exceedance probability plot for four different ecoregions: (a) Sierra Nevada, (b) Southwestern Tablelands, (c) Middle Atlantic Coastal Plain, and (d) Interior Forested Lowlands and Uplands. Black dots represent the empirical exceedance probabilities derived from annual-exceedance tail samples. The blue solid curve shows a three-parameter lognormal distribution fitted to the full sample of observed fire sizes, while the red dashed curve depicts a two-parameter lognormal model calibrated exclusively to the top-N annual exceedances.

These GACC-scale results present a notable contrast with the ecoregion-level findings. While the lognormal and Pareto Type II distributions performed comparably at the ecoregion scale (Figure 4b), the Weibull emerges as the strongest tail model once data are aggregated into larger administrative units. This suggests that the pooling of fire records across diverse ecoregions within each GACC smooths out the most extreme local variability, producing tail behavior that the Weibull's functional form can represent effectively. The Pareto Type II retains its advantage in the northwestern interior GACCs, where the heaviest tails persist even after aggregation.

5.3. Body-Tail Discrepancies

Figure 6 illustrates a central insight of this study: full-sample models, even those that appear to fit the overall distribution well, can significantly mischaracterize upper-tail behavior, leading to systematic under- or over-estimation of extreme fire risk. To isolate the effect of the fitting approach from the choice of distributional form, we use the lognormal distribution for both body and tail comparisons, as it provides consistently strong performance across both analyses. We examine four representative ecoregions: Sierra Nevada (Figure 6a), Southwestern Tablelands (Figure 6b), Middle Atlantic Coastal Plain (Figure 6c), and Interior Forested Lowlands and

Uplands (Figure 6d). In each subfigure, the blue solid line represents a three-parameter lognormal distribution fitted to the full data set, while the red dashed line represents a lognormal model calibrated to the top-N annual exceedances using the N_1 norm.

Across all 70 ecoregions, the body-fitted model overestimates the 100-year return fire size in 41 ecoregions (59%) and underestimates it in 29 (41%). The magnitude of these discrepancies is substantial: overestimation averages approximately 34,800 acres (≈14,085 ha) and reaches a maximum of 222,303 acres (≈89,960 ha) in the Yukon Flats of Alaska, while underestimation averages approximately 19,800 acres (≈8,013 ha) and reaches 74,905 acres (≈30,315 ha) in the Edwards Plateau. For example, the 100-year fire size is underestimated by 56,000 acres (≈22,662 ha) in the Sierra Nevada (Figure 6a) and 63,000 acres (≈25,500 ha) in the Southwestern Tablelands (Figure 6b). In contrast, all seven Alaskan ecoregions exhibit overestimation, with discrepancies ranging from approximately 1,700 to 222,300 acres; the Interior Forested Lowlands and Uplands (Figure 6d) overpredicts the 100-year fire size by approximately 183,000 acres (≈74,060 ha).

In contrast, the tail-calibrated model (red dashed line) more accurately captures upper-tail behavior. The red curve closely follows the empirical observations (black dots) across the extreme range, producing return-period estimates that align with observed extremes and preserving the frequency of extreme events. This discrepancy has direct implications for wildfire risk analysis and decision-making. At the size of the largest observed event in each region, likelihood estimates differ substantially between the two approaches. For example, in the Southwestern Tablelands (Figure 6b), the 2024 Smokehouse Creek Fire is assigned an Annual Exceedance Probability (AEP) of $\sim 1.1 \times 10^{-4}$ by the full-sample fit, compared to $\sim 1.6 \times 10^{-3}$ under the tail model, a more than fourteen-fold increase. In the most extreme case, the Southern Coastal Plain (ecoregion 75), the tail model assigns a probability more than 1,000 times higher than the body model, indicating near-complete failure of the full-sample fit to represent the extreme tail.

Wildfires in the eastern United States are generally smaller than those in the west, resulting in smaller discrepancies in return-period estimates; however, differences in exceedance probabilities remain substantial. For instance, in the Middle Atlantic Coastal Plain (Figure 6c), the largest and second-largest fires are estimated to be 131 and 58 times more likely, respectively, under the tail model than under the full-sample model.

Beyond improved accuracy, the tail-based N_1 calibration offers several advantages. It provides a unified, interpretable error metric across candidate distributions, avoids instability associated with maximum-likelihood estimation in heavy-tailed samples, and aligns directly with decision-relevant quantities such as exceedance probabilities and return periods.

These results highlight a key implication: relying solely on full-sample models can substantially misrepresent the likelihood and magnitude of rare but high-impact wildfires. Calibrating models to the upper tail provides a more accurate representation of extreme fire behavior, with direct relevance for resource allocation, infrastructure resilience, insurance modeling, and long-term planning.

5.4. Heaviness of the Tail

Figure 7 examines how the heaviness of large-fire size distributions varies across space and between modeling approaches by mapping the shape parameters of two heavy-tailed models, Pareto Type II and lognormal, fitted to both the full data set (Figures 7a–7d) and the top-N exceedances (Figures 7e–7h). The shape parameter controls tail thickness: larger values indicate heavier tails, meaning that extremely large fires occur more frequently and exert greater influence on the distribution.

Full-sample ("body") fits reveal strong regional contrasts. In the Pareto Type II distribution fitted to all observations (Figure 7a), the heaviest tails ($\gamma \approx 0.83$–1.26) occur in the Intermountain West, northern Rockies, and Alaska, regions characterized by infrequent but very large fires. In several of these regions, γ exceeds 1, implying that the theoretical mean burned area diverges and highlighting the disproportionate influence of extreme events. In contrast, the humid Southeast exhibits much lighter tails ($\gamma < 0.25$), indicating finite mean and variance and relatively infrequent large fires.

Aggregation to the GACC scale (Figure 7b) reduces local variability while preserving the broad east–west contrast. The California, Northern Rockies, and Northwest GACCs continue to exhibit the heaviest tails, whereas the Southern Area GACC centers near $\gamma \approx 0.46$. Lognormal fits to the full data set (Figures 7c and 7d)

10.1029/2025EF007485

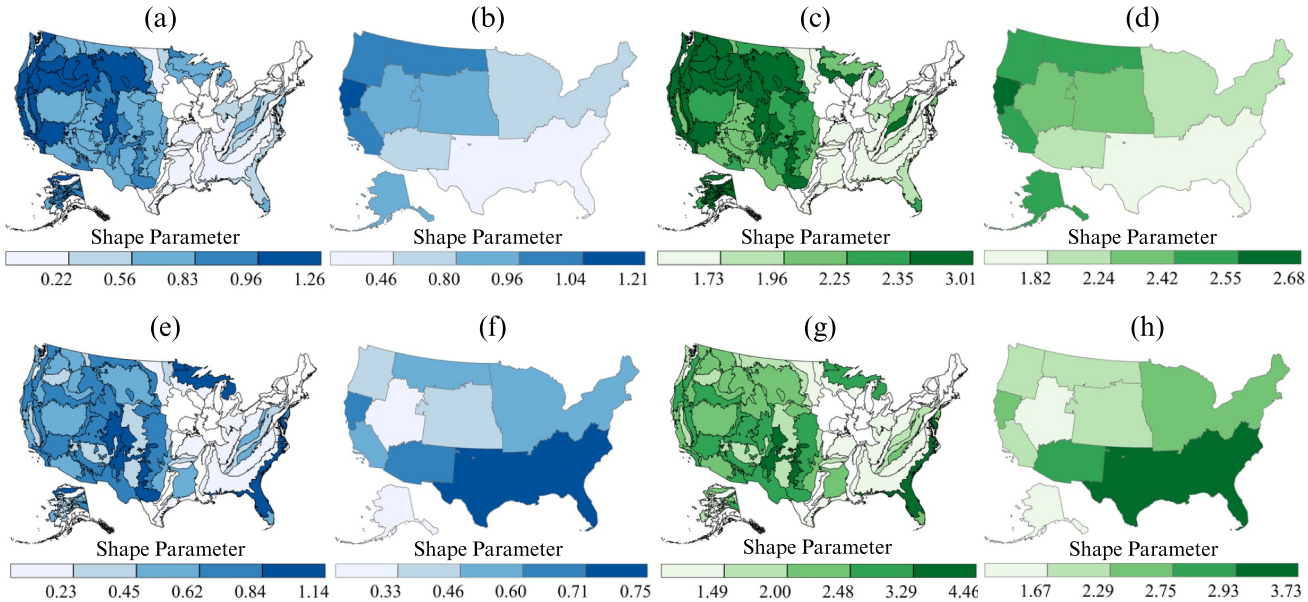

Figure 7. Spatial distribution of estimated shape parameters for the Pareto Type II and lognormal distributions, illustrating variation in tail heaviness across the United States. Panels (a–d) show shape parameters derived from full-sample ("body") fits: (a) Pareto Type II in Level III ecoregions; (b) Pareto Type II in Geographic Area Coordination Center (GACC) regions; (c) lognormal in Level III ecoregions; (d) lognormal in GACC regions. Panels (e–h) show shape parameters derived from tail-only fits using top-N exceedances: (e) Pareto Type II in Level III ecoregions; (f) Pareto Type II in GACC regions; (g) lognormal in Level III ecoregions; (h) lognormal in GACC regions.

show similar spatial gradients. At the ecoregion scale, γ ranges from approximately 1.18 in the Southeast to over 3.01 in the northern Rockies and coastal ranges. Because the lognormal mean and variance increase exponentially with γ, this implies substantial regional differences in expected fire size and variability. At the GACC scale (Figure 7d), western regions maintain $\gamma > 2.42$, confirming persistent heavy-tailed behavior despite aggregation, while the Southern Area GACC exhibits the lowest value.

Tail-only fits ("top-N exceedances") alter this pattern. When restricted to the top 41 observations, estimated tail heaviness decreases across regions. In the Pareto Type II model (Figure 7e), no region exceeds $\gamma = 1.14$, implying that the mean of the tail sample is finite in nearly all cases. Many eastern regions fall below $\gamma = 0.5$, indicating finite variance as well. Aggregation to GACCs (Figure 7f) further reduces tail heaviness, with most regions exhibiting $\gamma < 0.75$, making extreme fire sizes more statistically stable. However, the Southern Area ($\gamma \approx 0.75$) and Northern California ($\gamma \approx 0.71$) GACCs remain among the most heavy-tailed.

For the lognormal model (Figures 7g and 7h), the shape parameter ranges from approximately 0.80 to 4.46 at the ecoregion scale and from 1.52 to 3.73 at the GACC scale. The spatial pattern broadly mirrors that of the Pareto Type II, with heavier tails concentrated in western and southern regions. Although mean and variance remain finite under the lognormal model, spatial differences in tail heaviness remain substantial, with implications for planning and resource allocation, particularly for surge-capacity firefighting under extreme conditions.

6. Discussion and Conclusions

This study analyzed 30,331 large fire records (1984–2024) across the continental United States, totaling 213.6 million acres (≈86.4 million ha), using a dual modeling strategy: a three-parameter fit for the body of the fire size distribution and a two-parameter fit for the extreme right tail. We evaluated three common heavy-tailed distributions, Pareto Type II, lognormal, and Weibull, across 70 Level III ecoregions and 10 GACCs, providing a spatially explicit and comprehensive assessment of fire size distributions.

Our findings reveal a clear scale- and analysis-dependent model preference. For the central body of the distribution, the lognormal and Pareto Type II distributions frequently alternate as best fits across individual

ecoregions. When data are aggregated to the GACC scale, the lognormal provides the best body fit in 6 of 10 GACCs, with the Pareto Type II selected in the remaining 4. The tail analysis reveals a markedly different pattern. At the ecoregion level, the Pareto Type II and lognormal distributions perform comparably, with nearly identical mean ranks (1.89 and 1.86, respectively), though the Pareto Type II wins slightly more head-to-head comparisons (53% vs. 47%). The lognormal achieves its edge in mean rank through greater consistency, rarely ranking last. The Weibull distribution, despite being the weakest body model, emerges as a strong tail model, particularly at the GACC scale, where it achieves the lowest mean N_1 and provides the best tail fit in 6 of 10 GACCs. The Pareto Type II is selected in 3 GACCs concentrated in the northwestern interior, and the lognormal in 1. These results demonstrate that the distribution best suited for the body does not necessarily represent the extreme tail, with direct implications for wildfire risk assessment and planning.

6.1. Interpreting Body-Tail Discrepancies

One of the most consequential findings of this study is the clear divergence in statistical behavior between the central body of the large-fire size distribution and its extreme right tail. This divergence has direct implications for both modeling accuracy and practical decision-making. When fitted to the full data set, models such as the three-parameter lognormal or Pareto Type II can produce good visual agreement and low average error over the bulk of the distribution. However, these body-oriented fits can still fail to capture the true frequency and scale of the most extreme events, those of greatest concern to emergency planners, engineers, and policymakers.

Figure 6 illustrates this contrast using four representative ecoregions. The lognormal distribution fitted to the full data set (blue curve) tracks most observations well but diverges sharply in the tail, either underestimating or overestimating rare, high-return-period fires depending on the region. In contrast, fitting the same distribution solely to the top-N exceedances (red dashed curve) yields a substantially improved representation of tail behavior. For example, in the Southwestern Tablelands ecoregion, the exceedance probability of the largest wildfire increases from $\sim 1.1 \times 10^{-4}$ under the full-sample fit to $\sim 1.6 \times 10^{-3}$ under the tail-calibrated model, more than fourteen-fold increase. Importantly, this discrepancy is evaluated at an observed event rather than a theoretical extrapolation.

These differences translate directly into biased risk estimates for 50-, 100-, or 500-year fire events. Across the 70 ecoregions analyzed, the 100-year return size from the body-fitted model can differ by tens of thousands of acres from the tail-calibrated estimate, up to approximately 75,000 acres (≈30,350 ha) in underestimation and over 220,000 acres (≈89,030 ha) in overestimation. Similar discrepancies between body- and tail-based estimates have been reported in wildfire and natural-hazard applications, where full-sample fits misrepresent extremes despite acceptable global goodness-of-fit scores (e.g., Joseph et al., 2019; Papalexiou et al., 2013). These differences can influence decisions related to suppression capacity, contingency budgeting, and infrastructure design thresholds. Wildfire suppression expenditures in the United States are particularly sensitive to large events, with a small number of fires accounting for a disproportionate share of total costs (e.g., Calkin et al., 2015).

This body-tail misalignment is especially problematic when using goodness-of-fit metrics such as AIC, BIC, or Kolmogorov-Smirnov statistics, which are dominated by the bulk of the data. A model may perform well globally while still misrepresenting the extreme tail. This issue is well documented in extreme value and heavy-tail literature, where models that perform adequately on global fit criteria can yield biased tail inferences, particularly in finite samples (e.g., Clauset et al., 2009; Coles et al., 2001; Davison & Smith, 1990). In effect, decisions focused on rare events are being guided by metrics dominated by common events, introducing structural bias into planning.

Consistent behavior has been observed across natural hazards. Papalexiou et al. (2013) showed that fitting rainfall extremes with light-tailed distributions leads to systematic underestimation of extreme values. Similar patterns have been documented in flood frequency analysis, storm surge modeling, and compound hazards (e.g., Koutsoyiannis, 2004; Sornette, 2006), where underestimation of tail heaviness leads to biased design thresholds and safety margins. In wildfire systems, where climate change and anthropogenic influences are expanding the envelope of extremes, accurate tail characterization is increasingly important (Abatzoglou et al., 2021; Cunningham et al., 2024).

Moreover, the body–tail discrepancy is spatially variable. In some regions, such as the Southeastern Plains, body and tail behavior are more closely aligned, whereas in others, such as the Southwestern Tablelands or Southern

Coastal Plain, the divergence is more pronounced. This variability reinforces the need for region-specific modeling approaches that account for ecological and statistical differences across fire-prone landscapes. Such spatial variation is consistent with prior work documenting regional contrasts in wildfire regimes and fire–climate relationships across the United States (e.g., Abatzoglou & Williams, 2016; Joseph et al., 2019; Westerling et al., 2006).

These findings demonstrate that full-sample models are insufficient for characterizing extreme wildfire behavior. Risk-informed analysis should distinguish between modeling objectives: one model to represent common fire activity (e.g., budgeting, suppression strategy, fuel treatment) and a separate, tail-calibrated model to estimate rare, high-impact events relevant to infrastructure design, insurance, and emergency planning. This body–tail separation is analogous to approaches used in other risk domains, where different components of the distribution are modeled using methods tailored to their decision context (Embrechts et al., 1997; Klugman et al., 2012).

6.2. Model Robustness: Three- Versus Two-Parameter Forms

This study also highlights the importance of using appropriately parameterized models depending on which part of the distribution is being examined. When fitting to the full fire-size distribution, models that incorporate three parameters, including a location parameter to account for the minimum observable fire size, provide greater flexibility and typically reduce overall error. The inclusion of a minimum threshold is especially useful given that many fire records are truncated due to reporting standards or minimum mapping units. In such cases, allowing the model to estimate this lower bound improves fidelity of the body fit and provides a more accurate depiction of the distribution's interior. This approach is consistent with modern applications in environmental and actuarial statistics, where distributions with a location (threshold) parameter are widely used to model truncated samples arising from reporting thresholds or minimum mapping units (e.g., Coles et al., 2001; Embrechts et al., 1997).

However, this advantage diminishes when attention shifts to the distribution's extreme right tail. In tail modeling, where the number of observations is inherently limited (in this study, 41 annual exceedances), the inclusion of a location parameter can destabilize the model. With small samples, simultaneous estimation of location, scale, and shape parameters may produce weakly identified parameter combinations and highly variable tail behavior, particularly when extrapolating to long return periods. These challenges are well documented in the extreme value literature (e.g., Coles et al., 2001; Davison & Smith, 1990), where estimating multiple parameters in small samples can lead to identifiability issues and unstable shape-parameter estimates, especially under high thresholds and limited data. Similar stability concerns have been reported in applications of generalized Pareto and GEV models (e.g., Davison et al., 2012; Reiss & Thomas, 2001; Zaerpour et al., 2024), motivating the use of reduced-parameter or threshold-fixed formulations.

To address these issues, this study adopts a two-parameter fitting strategy for the tail, estimating only the scale and shape parameters. This simplification enhances numerical stability, improves convergence, and reduces parameter uncertainty, particularly for the shape parameter, which governs extrapolation into rare-event regimes. Fixing the location parameter also enables more consistent comparison of parameters across regions, helping to reveal large-scale spatial patterns in wildfire behavior. Importantly, this two-parameter approach does not materially reduce model accuracy in the tail. Because the focus is on extreme events, small biases near the lower bound of the tail domain are of limited consequence. Instead, accurate estimation of exceedance probabilities and return levels for very large fires is paramount, as these quantities directly inform infrastructure design thresholds, contingency planning, and wildfire resilience investments. Similar considerations have motivated parsimonious tail parameterizations in hydrological and climatological extreme-value studies (Papalexiou et al., 2013). In this context, model parsimony enhances robustness rather than limiting performance.

The broader implication is that model selection should be guided not only by goodness-of-fit metrics but also by the intended application. Full-sample modeling with three parameters is appropriate for characterizing overall fire behavior and informing operational decisions where the central distribution is most relevant. In contrast, tail-focused modeling with two parameters is better suited for estimating the likelihood and magnitude of rare, high-impact events that dominate long-term risk. Recognizing when each approach is appropriate enables more robust, credible, and actionable wildfire risk assessments across spatial scales. Accordingly, body and tail models should be viewed as complementary components of a unified risk framework, each calibrated and evaluated using diagnostics aligned with its specific decision context.

6.3. Implications for Risk Management and Practice

The results have direct implications for risk management and planning. Exceedance-based metrics translate statistical tail behavior into decision-relevant quantities such as return levels and probabilities. GACCs function as the operational framework for resource allocation in the United States, coordinating personnel, aviation assets, and equipment during periods of elevated fire activity. Estimating exceedance probabilities at this scale provides a regionally aggregated statistical signal of the likelihood of very large fires. While statistical distributions do not dictate operational decisions directly, they provide quantitative context for anticipating the magnitude and frequency of extreme fire demand. In this sense, exceedance-based modeling strengthens the empirical foundation upon which preparedness and coordination strategies are built.

It is important to note that these administrative distributions represent a spatial aggregation of diverse ecological processes. From a geostatistical perspective, a GACC often encompasses multiple distinct fire regimes, ranging from high-frequency, low-severity grass fires to infrequent, high-severity events. Aggregating these distinct processes into a single regional distribution provides a necessary portfolio view for national-level strategic budgeting and inter-regional resource allocation. However, this aggregation inherently masks intra-regional heterogeneity. Applying a GACC-wide tail metric to local land-use planning or community-level risk assessment would be inappropriate, as it could obscure the concentration of extreme behavior within specific fuel types or subregions. Accordingly, these metrics are intended to inform macro-scale strategic decisions rather than localized applications.

The exceedance-based framework also has important implications for financial risk management. Wildfire management expenditures in the United States frequently reach billions of dollars annually (Riddle, 2020), and a disproportionate share of these costs is associated with a relatively small number of large and complex fires (North et al., 2015). Suppression expenditures are strongly influenced by high-impact events, with large fires accounting for a substantial portion of annual costs (Calkin et al., 2015; Thompson et al., 2013). When extreme events are underestimated in probabilistic models, long-term budgeting and contingency planning may fail to capture the true financial exposure associated with rare but consequential fire seasons. Improved characterization of the statistical tail does not eliminate structural funding challenges, but it provides a more realistic quantitative basis for stress-testing budgets against low-probability, high-consequence scenarios.

The results also have implications for resource allocation strategies beyond immediate suppression. Land management agencies must decide where to invest in treatments such as prescribed burning and thinning. Traditional models focused on central tendencies may suggest that moderate fires dominate risk, leading to strategies that prioritize reducing medium-sized events. However, the exceedance analysis demonstrates that in many regions, a small number of very large fires account for a disproportionate share of total burned area and damage. In these settings, effective risk reduction requires strategies explicitly designed to reduce the likelihood or severity of extreme events. This may include strengthening fuel break networks, prioritizing treatment in high-risk ignition corridors, and strategically pre-positioning suppression resources in regions prone to extreme fire behavior.

6.4. Implications for Extreme-Event Prediction and Wildfire Policy

The findings of this study carry important implications for predictive modeling and wildfire policy across local, regional, and national scales. As fire regimes intensify under a changing climate, decision-makers increasingly rely on models that must credibly estimate not only average fire behavior but also the likelihood and magnitude of rare, high-impact events. The results show that models calibrated to the full distribution frequently misrepresent the extreme right tail, leading to systematic under- or over-estimation of the largest fires. Because extreme events disproportionately influence seasonal outcomes, such mischaracterization propagates into planning, budgeting, and preparedness decisions.

At the national scale, where coordination and forecasting are centralized, no single distribution dominates tail performance across all GACCs. The Weibull distribution provides the best tail fit in 6 of 10 GACCs, the Pareto Type II in 3, and the lognormal in 1. This diversity of best-fitting models indicates that a uniform national tail model is inappropriate. Instead, the results support a regionally differentiated modeling strategy, where each GACC adopts the distribution that best represents its aggregated extreme fire behavior. Despite this variation, the

framework, based on fitting two-parameter distributions to annual-exceedance samples and evaluating performance with the N1 norm, remains consistent and transferable across regions.

At regional and sub-regional scales, a similar pattern emerges. At the ecoregion level, the Pareto Type II and lognormal distributions perform comparably, with the Pareto Type II winning slightly more comparisons and the lognormal exhibiting greater consistency. Where the Pareto Type II is preferred, the tail follows a power-law decay, implying slower decline in exceedance probabilities and greater weight in extreme quantiles. Where the lognormal is preferred, the tail decays faster but remains heavy, still requiring explicit consideration of extreme events, though with relatively less emphasis on the most extreme outcomes.

A notable result is the reversal of the Weibull distribution's role between body and tail analyses. While rarely selected for the body and never preferred as a body model at the GACC scale, the Weibull emerges as the dominant tail model at the GACC level, with γ values well below 1 (median $\gamma \approx 0.36$), confirming heavy-tailed behavior. This indicates that, when data are aggregated across heterogeneous fire regimes, the Weibull's stretched-exponential form can effectively represent the smoothed tail structure. More broadly, this finding underscores that model suitability depends on the portion of the distribution being represented and cannot be inferred from body performance alone.

These results support a dual-regime modeling strategy for wildfire risk assessment. Body-level decisions, such as seasonal planning, suppression staffing, and fuel treatment prioritization, can be informed by models that best capture the central distribution (typically lognormal or Pareto Type II). In contrast, contingency planning and extreme-event preparedness require tail-calibrated models, which vary by region and scale. This separation ensures that both common events and rare extremes are represented using appropriate functional forms.

Beyond immediate applications, these findings inform efforts to model non-stationarity under climate change. Recent assessments (e.g., Westerling, 2018) have incorporated climate drivers but have still underestimated the size of wildfires over the past five years. This suggests that if the underlying distribution does not correctly represent tail decay, even dynamic parameter adjustments will fail to capture emerging extremes. While this study does not explicitly model temporal trends, it identifies distributional forms that better represent tail behavior, providing a structural basis for future non-stationary modeling efforts.

For the distribution body, the consistent performance of the lognormal model supports the widespread use of log-transformed burned area metrics in climate–fire studies (e.g., Westerling et al., 2006; Abatzoglou & Williams, 2016). However, this justification does not extend to the extreme tail, where the lognormal is not uniformly optimal and where Weibull and Pareto Type II models often provide better representations.

Overall, these results demonstrate that incorporating body-tail decomposition into fire modeling frameworks improves the reliability of extreme-event prediction. Models that focus only on central tendencies overlook the disproportionate influence of rare events. A modeling strategy that explicitly separates body and tail behavior, and selects regionally appropriate distributions, provides a more robust basis for wildfire risk assessment and policy. In a warming climate, where extreme fires are expected to play an increasing role, such approaches are essential for maintaining credible and adaptive decision-making.

Appendix A

Tables A1, A2, and A3 provide a comprehensive statistical characterization of burned area distributions across the continental United States, with a particular focus on both full records and extreme events. Table A1 summarizes the fitted parameters for the three candidate distributions applied to the complete data set. Table A2 shifts the focus to extremes by presenting descriptive statistics of tail samples, defined as the largest 41 events over a 41-year record, disaggregated by Level III ecoregions and Geographic Area Coordination Center (GACC) regions, thereby highlighting regional variability in large-fire behavior. Building on this, Table A3 evaluates the fit of the same candidate distributions specifically to these tail samples across GACC regions, providing insight into the ability of each distribution to capture heavy-tailed behavior and extreme wildfire events.

Table A1
Summary Statistics From Fitting the Three Distributions to the Full Observations of Burned Areas (in Acres)

	Ecoregions level III			GACC regions		
	Location	Scale	Shape	Location	Scale	Shape
	Pareto type II			Pareto type II		
Min	499.06	379.35	0.06	499.08	694.53	0.46
Mode	1009.99	1482.29	0.86	1000.11	1780.41	0.97
Mean	829.80	2113.45	0.71	900.01	2199.94	0.90
Median	1000.90	1646.43	0.80	999.50	1954.10	0.95
Max	1100.69	13290.19	1.26	1002.64	6235.29	1.21
SD	245.79	1987.67	0.34	211.05	1546.97	0.21
Skew	−0.60	3.33	−0.64	−1.78	2.24	−0.94
	Lognormal			Lognormal		
Min	272.51	278.89	1.18	475.53	587.27	1.82
Mode	991.65	1316.96	2.40	991.19	1786.54	2.47
Mean	796.91	1944.76	2.17	888.69	2152.28	2.37
Median	984.07	1534.37	2.29	989.94	1889.70	2.45
Max	1037.80	11151.30	3.01	996.43	5775.42	2.68
SD	267.52	1782.41	0.46	213.71	1458.82	0.25
Skew	−0.68	2.76	−0.78	−1.78	1.84	−1.25
	Weibull			Weibull		
Min	499.09	496.12	0.50	499.06	1206.01	0.53
Mode	1010.04	2600.30	0.60	1000.02	4162.03	0.59
Mean	829.80	4310.01	0.67	900.02	5074.30	0.60
Median	1000.93	3158.41	0.62	999.51	4388.08	0.59
Max	1100.74	25127.40	1.08	1002.55	13721.90	0.69
SD	245.81	4194.81	0.15	211.06	3584.22	0.05
Skew	−0.60	2.51	1.39	−1.78	1.64	0.64

Table A2
Descriptive Statistics of the Tail Samples Defined for a 41-Year Record as the 41 Largest Values in Level III Ecoregions and Geographic Area Coordination Centers of the Continental United States

	Ecoregion level III					GACC regions				
	Median	Mean	SD	Max	Skew	Median	Mean	SD	Max	Skew
Min	7408.2	13947.7	17475.8	105539.4	2.69	66035.5	72671.4	41843.7	154790.0	0.85
Q5	7663.8	14638.5	18170.8	108971.8	2.62	67290.2	75026.7	42999.0	160298.2	0.85
Q25	8653.1	17724.8	21157.3	122776	2.34	84751.0	88263.7	45256.3	170471.7	0.40
Q50	11010.6	24982.7	29000.9	154135.8	1.97	101742.2	110007.4	53157.4	214189.0	0.44
Q75	22099.7	41536.7	44473.2	197946.7	1.39	165082.0	167092.9	76901.5	310223.0	0.07
Q95	69040.5	106691.2	113878.1	527450.3	1.58	349784.3	339965.1	145375.7	527450.3	−0.58
Max	138163.7	202223.5	211718.1	984414.1	1.63	560573.6	541597.9	271987.3	984414.1	0.28
Mean	21867.6	38851.6	40212.8	174035.8	1.31	153667.8	147318.3	62924.7	249604.7	−0.39
SD	23996.8	38703.1	40075.6	191437.3	1.55	98756.7	97932.1	48868.1	177006.6	0.35
Skew	2.66	2.75	1.10	5.29	0.52	2.52	2.34	0.57	3.06	−0.89

Table A3
Summary Statistics From the Fitting of the Three Distribution Into the Tail Samples of Burned Area (in Acres) for Geographic Area Coordination Center Regions

	Pareto type II			Lognormal			Weibull		
	N_1	β	γ	N_1	β	γ	N_1	β	γ
Min	0.007	599.2	0.293	0.007	43.4	1.520	0.007	3.5	0.181
Mode	0.012	4934.3	0.504	0.010	2339.9	2.163	0.012	1336.2	0.353
Mean	0.019	7200.0	0.530	0.016	4468.1	2.361	0.015	2473.0	0.348
Median	0.014	6186.1	0.512	0.011	2744.0	2.206	0.014	1459.3	0.360
Max	0.045	22719.0	0.755	0.034	18323.8	3.726	0.025	11053.8	0.506
SD	0.013	6599.6	0.156	0.010	5513.7	0.653	0.006	3284.0	0.093
Skew	1.219	1.6	−0.061	0.993	2.1	0.886	0.148	2.3	−0.109

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

The wildfire burned area data used for analyzing extreme fire events in this study are available from the MTBS program (MTBS Project, 2025) at https://www.mtbs.gov/direct-download. The Ecoregion Level III shapefile used for regional classification and spatial analysis is available from the US Environmental Protection Agency (EPA, 2025) at https://www.epa.gov/eco-research/level-iii-and-iv-ecoregions-continental-united-states. The Geographic Area Coordination Center (GACC) boundaries used for fire management zone analysis are available from the National Interagency Fire Center at https://gacc.nifc.gov (National Interagency Coordination Center, 2025). All data sets are publicly accessible and provided under open access conditions.

Acknowledgments
This material is based upon work supported by the United States Department of Agriculture (USDA), the National Institute of Food and Agriculture (NIFA) under Grant 2021-67022-35908.

References

Abatzoglou, J. T., Battisti, D. S., Williams, A. P., Hansen, W. D., Harvey, B. J., & Kolden, C. A. (2021). Projected increases in Western US forest fire despite growing fuel constraints. *Communications Earth & Environment*, *2*(1), 227. https://doi.org/10.1038/s43247-021-00299-0

Abatzoglou, J. T., & Williams, A. P. (2016). Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences*, *113*(42), 11770–11775. https://doi.org/10.1073/pnas.1607171113

Abdollahi, M., Vahedifard, F., & Leshchinsky, B. A. (2024). Hydromechanical modeling of evolving post-wildfire regional-scale landslide susceptibility. *Engineering Geology*, *335*, 107538. https://doi.org/10.1016/j.enggeo.2024.107538

Abdollahi, M., Vahedifard, F., & Tracy, F. T. (2023). Post-wildfire stability of unsaturated hillslopes against rainfall-triggered landslides. *Earth's Future*, *11*(3), e2022EF003213. https://doi.org/10.1029/2022ef003213

AghaKouchak, A., Hjelmstad, A., Lucy, J., Duku, J., Sadhasivam, N., Werth, S., et al. (2025). Building urban fire resilience to enhance national security. *Nature Cities*, *2*(9), 778–780. https://doi.org/10.1038/s44284-025-00296-w

Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/ f noise. *Physical Review Letters*, *59*(4), 381–384. https://doi.org/10.1103/PhysRevLett.59.381

Bølviken, E., & Hobæk Haff, I. (2024). Loss modeling with many-parameter distributions. *Scandinavian Actuarial Journal*, *2024*(8), 763–780. https://doi.org/10.1080/03461238.2024.2309987

Brown, E. K., Wang, J., & Feng, Y. (2021). US wildfire potential: A historical view and future projection using high-resolution climate data. *Environmental Research Letters*, *16*(3), 034060. https://doi.org/10.1088/1748-9326/aba868

Burnham, K. P. & Anderson, D. R. (Eds.) (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. Springer New York.

Calkin, D. E., Thompson, M. P., & Finney, M. A. (2015). Negative consequences of positive feedbacks in US wildfire management. *Forest Ecosystems*, *2*(1), 9. https://doi.org/10.1186/s40663-015-0033-8

Chelsea Nagy, R., Fusco, E., Bradley, B., Abatzoglou, J. T., & Balch, J. (2018). Human-related ignitions increase the number of large wildfires across U.S. ecoregions. *Fire*, *1*(1), 4. https://doi.org/10.3390/FIRE1010004

Cisneros, D., Hazra, A., & Huser, R. (2023). Spatial wildfire risk modeling using mixtures of tree-based multivariate Pareto distributions. arXiv preprint arXiv:2308.03870.

Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*(4), 661–703. https://doi.org/10.1137/070710111

Coles, S., Bawa, J., Trenner, L., & Dorazio, P. (2001). *An introduction to statistical modeling of extreme values* (Vol. 208, p. 208). Springer.

Cumming, S. G. (2001). A parametric model of the fire-size distribution. *Canadian Journal of Forest Research*, *31*(8), 1297–1303. https://doi.org/10.1139/cjfr-31-8-1297

Cunningham, C. X., Abatzoglou, J. T., Kolden, C. A., Williamson, G. J., Steuer, M., & Bowman, D. M. J. S. (2025). Climate-linked escalation of societally disastrous wildfires. *Science*, *390*(6768), 53–58. https://doi.org/10.1126/science.adr5127

Cunningham, C. X., Williamson, G. J., & Bowman, D. M. J. S. (2024). Increasing frequency and intensity of the most extreme wildfires on Earth. *Nature Ecology & Evolution*, *8*(8), 1420–1425. https://doi.org/10.1038/s41559-024-02452-2

Davis, F. W., Parkinson, A.-M., Moritz, M. A., Park, I. W., & D'Antonio, C. M. (2025). Increasing vulnerability of an endemic Mediterranean-climate conifer to changing climate and fire regime. *Frontiers in Forests and Global Change*, *8*, 1516623. https://doi.org/10.3389/ffgc.2025.1516623

Davison, A. C., Padoan, S. A., & Ribatet, M. (2012). Statistical modeling of spatial extremes. *Statistical Science*, *27*(2), 161–186. https://doi.org/10.1214/11-sts376

Davison, A. C., & Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *52*(3), 393–425. https://doi.org/10.1111/j.2517-6161.1990.tb01796.x

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

Donovan, V. M., Crandall, R., Fill, J., & Wonkka, C. L. (2023). Increasing large wildfire in the Eastern United States. *Geophysical Research Letters*, *50*(24), e2023GL107051. https://doi.org/10.1029/2023GL107051

Drossel, B., & Schwabl, F. (1992). Self-organized critical forest-fire model. *Physical Review Letters*, *69*(11), 1629–1632. https://doi.org/10.1103/PhysRevLett.69.1629

Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Springer.

EPA. (2025). Ecoregions of North America [Dataset]. *United States Environmental Protection Agency*. Retrieved from https://www.epa.gov/eco-research/level-iii-and-iv-ecoregions-continental-united-states

Ermagun, A., Thompson, D., Vahedifard, F., & Silver, R. C. (2025). Emergency managers' challenges with wildfires and related cascading hazards in California. *Journal of Environmental Management*, *374*, 124008. https://doi.org/10.1016/j.jenvman.2024.124008

Finney, M. A., McHugh, C. W., Grenfell, I. C., Riley, K. L., & Short, K. C. (2011). A simulation of probabilistic wildfire risk components for the continental United States. *Stochastic Environmental Research and Risk Assessment*, *25*(7), 973–1000. https://doi.org/10.1007/s00477-011-0462-z

Gupta, S. K. (2011). *Modern hydrology and sustainable water development*. John Wiley & Sons.

Hantson, S., Pueyo, S., & Chuvieco, E. (2016). Global fire size distribution: From power law to log-normal. *International Journal of Wildland Fire*, *25*(4), 403–412. https://doi.org/10.1071/wf15108

Holmes, T. P., Huggett, Jr., R. J., & Westerling, A. L. (2008). Statistical analysis of large wildfires. In T. P. Holmes, J. P. Prestemon, & K. L. Abt (Eds.), *The economics of forest disturbances: Wildfires, storms, and invasive species* (Vol. 79, pp. 59–77). Springer. https://doi.org/10.1007/978-1-4020-4370-3

Joseph, M. B., Rossi, M. W., Mietkiewicz, N. P., Mahood, A. L., Cattau, M. E., St. Denis, L. A., et al. (2019). Spatiotemporal prediction of wildfire size extremes with Bayesian finite sample maxima. *Ecological Applications*, *29*(6), e01898. https://doi.org/10.1002/eap.1898

Klugman, S. A., Panjer, H. H., & Willmot, G. E. (2012). *Loss models: From data to decisions* (4th ed.). John Wiley & Sons.

Koutsoyiannis, D. (2004). Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records / Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: II. Recherche empirique sur de longues séries de précipitations. *Hydrological Sciences Journal*, *49*(4), 610. https://doi.org/10.1623/hysj.49.4.591.54424

Liu, Y., Goodrick, S. L., & Stanturf, J. A. (2013). Future US wildfire potential trends projected using a dynamically downscaled climate change scenario. *Forest Ecology and Management*, *294*, 120–135. https://doi.org/10.1016/j.foreco.2012.06.049

Malamud, B. D., Millington, J. D. A., & Perry, G. L. W. (2005). Characterizing wild-fire regimes in the United States. *Proceedings of the National Academy of Science*, *102*(13), 4694–4699. https://doi.org/10.1073/pnas.0500880102

Malamud, B. D., Morein, G., & Turcotte, D. L. (1998). Forest fires: An example of self-organized critical behavior. *Science*, *281*(5384), 1840–1842. https://doi.org/10.1126/science.281.5384.1840

McConnell, K., Fussell, E., DeWaard, J., Whitaker, S., Curtis, K. J., St. Denis, L., et al. (2024). Rare and highly destructive wildfires drive human migration in the U.S. *Nature Communications*, *15*(1), 6631. https://doi.org/10.1038/s41467-024-50630-4

Moritz, M. A. (1997). Analyzing extreme disturbance events: Fire in Los Padres National Forest. *Ecological Applications*, *7*(4), 1252. https://doi.org/10.2307/2641212

MTBS Project. (2025). MTBS burned areas boundaries dataset [Dataset]. Retrieved from https://www.mtbs.gov/direct-download

Mueller, S. E., Thode, A. E., Margolis, E. Q., Yocom, L. L., Young, J. D., & Iniguez, J. M. (2020). Climate relationships with increasing wildfire in the southwestern US from 1984 to 2015. *Forest Ecology and Management*, *460*, 117861. https://doi.org/10.1016/J.FORECO.2019.117861

National Interagency Coordination Center. (2025). Geographic Area Coordination Center (GACC) boundaries [Dataset]. *National Interagency Fire Center*. Retrieved from https://gacc.nifc.gov/

Nepal, S., Pomara, L. Y., Gould, N. P., & Lee, D. C. (2023). Wildfire risk assessment for strategic forest management in the southern United States: A Bayesian network modeling approach. *Land*, *12*(12), 2172. https://doi.org/10.3390/land12122172

Newman, M. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, *46*(5), 323–351. https://doi.org/10.1080/00107510500052444

North, M. P., Stephens, S. L., Collins, B. M., Agee, J. K., ApleT, G., Franklin, J. F., & Fule, P. Z. (2015). Reform forest fire management. *Science*, *349*(6254), 1280–1281. https://doi.org/10.1126/science.aab2356

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

Palmieri, L., & Jensen, H. J. (2020). The forest fire model: The subtleties of criticality and scale invariance. *Frontiers in Physics*, *8*, 257. https://doi.org/10.3389/fphy.2020.00257

Papalexiou, S. M., Koutsoyiannis, D., & Makropoulos, C. (2013). How extreme is extreme? An assessment of daily rainfall distribution tails. *Hydrology and Earth System Sciences*, *17*(2), 851–862. https://doi.org/10.5194/hess-17-851-2013

Parks, S. A., & Abatzoglou, J. T. (2020). Warmer and drier fire seasons contribute to increases in area burned at high severity in western U.S. forests from 1985 to 2017. *Geophysical Research Letters*, *47*(22), e2020GL089858. https://doi.org/10.1029/2020GL089858

Prestemon, J. P., Shankar, U., Xiu, A., Talgo, K., Yang, D., Dixon, E., et al. (2016). Projecting wildfire area burned in the south-eastern United States, 2011–60. *International Journal of Wildland Fire*, *25*(7), 715–729. https://doi.org/10.1071/WF15124

Pueyo, S. (2007). Self-organised criticality and the response of wildland fires to climate change. *Climatic Change*, *82*(1–2), 131–161. https://doi.org/10.1007/S10584-006-9134-2

Reed, W. J., & McKelvey, K. S. (2002). Power-law behaviour and parametric models for the size-distribution of forest fires. *Ecological Modelling*, *150*(3), 239–254. https://doi.org/10.1016/s0304-3800(01)00483-5

Reiss, R.-D., & Thomas, M. (2001). *Statistical analysis of extreme values* (2nd ed.). Birkhäuser.

Ricotta, C., Arianoutsou, M., Dıaz-Delgado, R., Duguy, B., Lloret, F., Maroudi, E., et al. (2001). Self-organized criticality of wildfires ecologically revisited. *Ecological Modelling*, *141*(1–3), 307–311. https://doi.org/10.1016/s0304-3800(01)00272-1

Ricotta, C., Avena, G., & Marchetti, M. (1999). The flaming sandpile: Self-organized criticality and wildfires. *Ecological Modelling*, *119*(1), 73–77. https://doi.org/10.1016/s0304-3800(99)00057-5

Riddle, A. A. (2020). Federal wildfire management: Ten-year funding trends and issues. (FY2011-FY2020).

Sadegh, M., Abatzoglou, J. T., AghaKouchak, A., & Seydi, S. T. (2025). Ignition matters. *Nature Sustainability*, *8*(4), 327–328. https://doi.org/10.1038/s41893-025-01527-7

Salguero, J., Li, J., Farahmand, A., & Reager, J. T. (2020). Wildfire trend analysis over the contiguous United States using remote sensing observations. *Remote Sensing*, *12*(16), 2565. https://doi.org/10.3390/rs12162565

Schoenberg, F. P., Peng, R., & Woods, J. (2003). On the distribution of wildfire sizes. *Environmetrics*, *14*(6), 583–592. https://doi.org/10.1002/env.605

Scott, J. H., Thompson, M. P., & Calkin, D. E. (2013). A wildfire risk assessment framework for land and resource management.

Song, W., Weicheng, F., Binghong, W., & Jianjun, Z. (2001). Self-organized criticality of forest fire in China. *Ecological Modelling*, *145*(1), 61–68. https://doi.org/10.1016/s0304-3800(01)00383-0

Sornette, D. (2006). *Critical phenomena in natural sciences: Chaos, fractals, selforganization and disorder: Concepts and tools*. Springer Science & Business Media.

Tao, T. J., Estes, K. D., Holman, E. A., Vahedifard, F., & Silver, R. C. (2025). Understanding climate change anxiety and anticipatory climate disaster stress: A survey of residents in a high-risk California county during wildfire season. *BMJ Mental Health*, *28*(1), e301331. https://doi.org/10.1136/bmjment-2024-301331

Taylor, S. W., Woolford, D. G., Dean, C. B., & Martell, D. L. (2013). Wildfire prediction to inform fire management: Statistical science challenges. *Statistical Science*, *28*(4), 586–615. https://doi.org/10.1214/13-STS451

Teymoor Seydi, S., Abatzoglou, J. T., Jones, M. W., Kolden, C. A., Filippelli, G., Hurteau, M. D., et al. (2025). Increasing global human exposure to wildland fires despite declining burned area. *Science*, *389*(6762), 826–829. https://doi.org/10.1126/science.adu6408

Thompson, M. P., Calkin, D. E., Finney, M. A., Ager, A. A., & Gilbertson-Day, J. W. (2011). Integrated national-scale assessment of wildfire risk to human and ecological values. *Stochastic Environmental Research and Risk Assessment*, *25*(6), 761–780. https://doi.org/10.1007/s00477-011-0461-0

Thompson, M. P., Calkin, D. E., Finney, M. A., Gebert, K. M., & Hand, M. S. (2013). A risk-based approach to wildland fire budgetary planning. *Forest Science*, *59*(1), 63–77. https://doi.org/10.5849/forsci.09-124

Vahedifard, F., Abdollahi, M., Leshchinsky, B. A., Stark, T. D., Sadegh, M., & AghaKouchak, A. (2024). Interdependencies between wildfire-induced alterations in soil properties, near-surface processes, and geohazards. *Earth and Space Science*, *11*(2), e2023EA003498. https://doi.org/10.1029/2023ea003498

Vogel, R. M., & Fennessey, N. M. (1993). L moment diagrams should replace product moment diagrams. *Water Resources Research*, *29*(6), 1745–1752. https://doi.org/10.1029/93wr00341

Vogel, R. M., Papalexiou, S. M., Lamontagne, J. R., & Dolan, F. C. (2025). When heavy tails disrupt statistical inference. *The American Statistician*, *79*(2), 221–235. https://doi.org/10.1080/00031305.2024.2402898

Wang, D., Guan, D., Zhu, S., Kinnon, M. M., Geng, G., Zhang, Q., et al. (2020). Economic footprint of California wildfires in 2018. *Nature Sustainability*, *4*(3), 252–260. https://doi.org/10.1038/s41893-020-00646-7

Ward, P. C., Tithecott, A. G., & Wotton, B. M. (2001). Reply: A re-examination of the effects of fire suppression in the boreal forest. *Canadian Journal of Forest Research*, *31*(8), 1467–1480. https://doi.org/10.1139/cjfr-31-8-1467

Westerling, A. L. (2016). Increasing western US forest wildfire activity: Sensitivity to changes in the timing of spring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1696), 20150178. https://doi.org/10.1098/rstb.2015.0178

Westerling, A. L. (2018). *Wildfire simulations for California's fourth climate change assessment: Projecting changes in extreme wildfire events with a warming climate: A report for California's fourth climate change assessment* (p. 57). California Energy Commission.

Westerling, A. L., Hidalgo, H. G., Cayan, D. R., & Swetnam, T. W. (2006). Warming and earlier spring increase western U.S. forest wildfire activity. *Science*, *313*(5789), 940–943. https://doi.org/10.1126/science.1128834

Yang, H. Q., & Vahedifard, F. (2026). AI-augmented modeling of post-fire soil hydraulic conductivity across environmental gradients for geohazard assessment. *Reliability Engineering & System Safety*, *272*, 112486. https://doi.org/10.1016/j.ress.2026.112486

Zaerpour, M., Papalexiou, S. M., Pietroniro, A., & Nazemi, A. (2024). How extreme are flood peak distributions? A quasi-global analysis of daily discharge records. *Journal of Hydrology*, *631*, 130849. https://doi.org/10.1016/j.jhydrol.2024.130849

Zhou, S., & Bradley, J. R. (2026). Multiscale multi-type spatial Bayesian analysis for high-dimensional data with application to wildfires and migration. *Environmetrics*, *37*(1), e70066. https://doi.org/10.1002/env.70066

Zhuang, Y., Fu, R., Santer, B. D., Dickinson, R. E., & Hall, A. (2021). Quantifying contributions of natural variability and anthropogenic forcings on increased fire weather risk over the western United States. *Proceedings of the National Academy of Sciences*, *118*(45), e2111875118. https://doi.org/10.1073/pnas.2111875118