QUBAHAN JOURNAL OF COACHING AND SPORTS SCIENCES
VOL. 4, NO. 4, November 2026
https://doi.org/10.48161/qjcs.v4n4a91

Tail-Aware Distribution Modeling for Extreme Weather Losses: Generalized Pareto Analysis of NOAA Data

Swapnil Anil Prabhudesai[1*]

[1] Basic Science and Humanities, NMIMS'S Mukesh Patel School of Technology Management and Engineering, NMIMS University, Mumbai, India;

* **Corresponding author:** swapnil71286@gmail.com.

ABSTRACT: Extreme events and rare outcomes are difficult to model because they occur infrequently, show strong skewness, and often produce consequences much larger than ordinary observations. This study develops and tests an empirical probability-distribution fitting framework for modeling extreme weather losses and rare outcomes using the National Oceanic and Atmospheric Administration (NOAA) Storm Events dataset. The analysis uses 1,962,332 storm-event records from 1950 to 2025 and focuses on 450,823 positive economic-loss observations after converting reported property and crop damage values from Thousands/Millions/Billions (K/M/B)-coded strings into numeric United States (U.S.) dollars. Candidate distributions include normal, lognormal, exponential, Weibull, Pareto, generalized extreme value, and generalized Pareto models. Parameters are estimated by maximum likelihood estimation (MLE), while performance is evaluated using log-likelihood, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Kolmogorov–Smirnov (KS) distance, graphical diagnostics, threshold sensitivity, and high-quantile tail-risk estimates. The empirical results show strong heavy-tailed behavior: positive damage has a median of $10.00 thousand, a 99th percentile of $10.59 million, a 99.9th percentile of $200.00 million, and a maximum reported value of $17.90 billion. The lognormal distribution provides the strongest full-body fit among conventional models, but it substantially underestimates the far upper tail. By contrast, the generalized Pareto peaks-over-threshold model at the 95th percentile produces tail estimates close to the empirical 99th and 99.9th percentiles, with a Kolmogorov–Smirnov (KS) distance of 0.076 for exceedances. The study contributes a transparent, reproducible framework for fitting probability distributions to rare event losses and shows that separate body and tail modeling is necessary for reliable risk estimation.

Keywords: Extreme value theory, heavy-tailed distributions, generalized Pareto distribution, peaks-over-threshold, economic loss modeling.

I. INTRODUCTION

Extreme events occupy a difficult position in empirical research. They are rare enough to be overlooked by conventional descriptive models, but consequential enough to determine the practical meaning of risk. In natural hazards, finance, engineering reliability, public health, and environmental management, the most important losses are often located in the far right tail rather than near the average. A storm event that produces no reported economic damage may be common, while a hurricane, flood, wildfire, or tornado that produces millions or billions of dollars of damage may be rare. The statistical challenge is therefore not only to describe the average record but also to identify which probability distribution can represent the behavior of the extreme tail.

Probability distribution fitting provides a practical method for this problem. It allows observed outcomes to be compared with theoretical distributions, parameters to be estimated from data, and risk quantities such as high quantiles and exceedance probabilities to be calculated. However, the choice of distribution is not a

neutral technical detail. A normal distribution imposes symmetry and thin tails. A lognormal distribution accommodates positive skewness. Weibull and exponential models describe positive lifetimes and severities but may still be too light-tailed for catastrophic losses. Pareto-type models and generalized Pareto distributions are designed for heavy tails and high-threshold exceedances, but they require careful threshold selection and diagnostic validation. The central methodological issue is therefore whether the same model can represent the main body of the data and the extreme tail, or whether body and tail behavior must be evaluated separately.

The empirical setting of this study is severe weather damage in the United States. The NOAA Storm Events Database provides long-term event-level records of storm and weather phenomena that involve loss of life, injuries, significant property damage, crop damage, disruption to commerce, or other meteorologically significant conditions. The uploaded Kaggle archive used in this study contains three CSV files derived from the NOAA Storm Events bulk data: an event-detail table, a fatality table, and a location table. The event-detail file is the main empirical source because it provides the event year, event type, state, injuries, deaths, property damage, crop damage, and magnitude information required for distribution fitting. The data are particularly suitable for rare-outcome research because the sample contains many ordinary observations but also a small number of very large reported losses.

The problem addressed in this study is that applied studies often select a probability distribution using only overall goodness-of-fit criteria, even though risk decisions are typically driven by the upper tail. A model can fit the center of the data well while seriously underestimating the 99th or 99.9th percentile. This is especially problematic in disaster-risk management, where preparedness, insurance reserves, infrastructure planning, and emergency-response capacity depend on high-severity outcomes. Therefore, the present study compares conventional distributions with extreme-value models and evaluates them using both full-sample fit and tail-risk accuracy.

The aim of the study is to develop and empirically demonstrate a distribution-fitting framework for modeling extreme weather losses and rare human-impact outcomes. The specific objectives are: (1) to clean and transform storm-event damage variables into a usable continuous loss measure; (2) to estimate candidate probability distributions for positive economic damage; (3) to compare model performance using information criteria and goodness-of-fit diagnostics; (4) to evaluate tail-risk estimates at the 95th, 99th, 99.5th, and 99.9th percentiles; and (5) to examine whether human-impact outcomes such as deaths and injuries display similar rare-outcome patterns.

The paper is structured as follows. Section 2 reviews literature on distribution fitting, extreme-value theory, disaster losses, and threshold selection. Section 3 identifies the research gap. Section 4 presents the study contribution. Section 5 describes the proposed method. Section 6 explains the dataset and analytical procedure. Section 7 reports empirical results. Sections 8 to 11 discuss theoretical contributions, practical implications, limitations, and conclusions.

II. LITERATURE REVIEW

Extreme-event research strengthened the statistical foundation for modeling observations located in the upper tail rather than near the center of a distribution. Climate-extreme methods emphasized detection, analysis, and uncertainty quantification for high-impact events [1]. Composite loss modeling further showed that frequent small losses and rare catastrophic losses may require separate body and tail components [2]. Automated threshold-selection research provided a reproducible alternative to subjective cutoff choice in extreme-value analysis [3]. Work on truncated Pareto-type distributions also warned that real loss tails may be affected by reporting limits, administrative thresholds, or physical constraints [4].

Risk-management and time-series applications clarify why distribution fitting must align with the purpose of the analysis. Pareto-type models are especially relevant when the objective is high-layer loss estimation, reinsurance, or capital planning rather than average prediction [5]. Time-series extreme-value research shows that dependence and exceedance behavior can affect rare-event interpretation [6]. Broader reviews of extreme statistics confirm that full-body distribution fitting and tail-specific inference answer different questions [7]. Spatial-extreme reviews add that environmental extremes may vary across locations and exposure contexts, which is important for storm-event damage data [8].

Spatial, disaster-loss, and software-oriented studies further support a tail-aware empirical design. Statistical modeling of spatial extremes shows that extreme weather processes cannot always be generalized across regions [9]. Natural-disaster loss research demonstrates that lognormal-Pareto and related structures can represent both ordinary losses and rare severe outcomes [10]. Software-oriented work on weather and climate extremes encourages transparent reporting of GEV and GPD specifications, parameters, and diagnostics [11]. GPD-based value-at-risk research confirms the usefulness of high-threshold exceedances for estimating tail-risk measures [12].

Recent spatial-extreme and climate-risk studies also highlight future methodological extensions. Reviews of spatial extremes identify dependence, nonstationarity, and regional heterogeneity as important directions for advanced rare-event modeling [13]. Climate-extreme statistics provide a post-2010 basis for high-quantile estimation and explicit tail uncertainty [14]. Economic studies of natural disasters show that loss estimates influence adaptation, insurance, public finance, and preparedness decisions [15]. Flood-risk research connects climate, exposure, vulnerability, and risk management, reinforcing the need to interpret severe losses in their hazard context [16].

Environmental and official-data studies reinforce the need to separate ordinary and extreme behavior. Rainfall research shows that low, moderate, and heavy intensities may require different distributional treatment rather than one universal model [17]. NOAA documentation confirms that storm-event records include fatalities, injuries, property damage, crop damage, event narratives, and storm classifications, making the data appropriate for the present empirical framework [18]. Improved threshold diagnostic plots support explicit GPD threshold evaluation instead of arbitrary cutoff selection [19]. Spatial-extreme modeling based on asymptotic independence further indicates that location and exposure may influence the structure of severe weather outcomes [20].

Historical loss databases and GPD extensions show that rare-event analysis depends on both data quality and model specification. Long-term flood-loss databases demonstrate the importance of event records for understanding hazard losses over time [21]. Extended generalized Pareto models show that classical GPD fitting may require diagnostic support and methodological refinement when tail behavior is unstable [22]. Design-life level research links return-level estimation with planning decisions under changing climate conditions [23]. Reviews of threshold uncertainty caution that peaks-over-threshold modeling is reliable only when threshold choice and uncertainty are reported transparently [24].

The most recent studies in the reviewed period strengthen the manuscript's emphasis on threshold sensitivity, climate extremes, and applied tail-risk estimation. Univariate threshold-selection procedures show that high-quantile estimates can change materially when the threshold is too low or too high [25]. Climate-assessment research emphasizes that weather and climate extremes require risk estimates that account for severity, exposure, and uncertainty [26]. Environmental rainfall applications demonstrate that GPD peaks-over-threshold models can produce practical return-level estimates for planning [27]. Disaster-loss database research highlights the importance of cleaning, reporting standards, and measurement consistency [28]. Seismic hazard applications show that GPD models are useful when low exceedance probabilities drive decisions [29]. Compound-event research finally shows that extreme losses may arise from interacting meteorological drivers, supporting cautious interpretation of event-level storm records [30].

Table 1 presents the literature-review comparison table.

No.	Study	Main focus	Relevance to this manuscript
[1]	AghaKouchak et al. (2013)	Synthesized methods for detecting, analyzing, and quantifying climate extremes under uncertainty.	Connects the present distribution-fitting framework with climate-risk and hydrological-extreme applications.
[2]	Aminzadeh and Deng (2017)	Developed Bayesian predictive modeling for exponential-Pareto composite distributions.	Supports body-tail thinking when frequent smaller losses coexist with rare catastrophic losses.

No.	Study	Main focus	Relevance to this manuscript
[3]	Bader et al. (2018)	Proposed automated threshold selection for extreme-value analysis using ordered goodness-of-fit tests.	Supports reproducible threshold choice and the threshold-sensitivity design used in this study.
[4]	Beirlant et al. (2016)	Examined extreme-value statistics for truncated Pareto-type distributions.	Warns that real loss tails may be shaped by reporting limits, truncation, or physical constraints.
[5]	Charpentier and Flachaire (2021)	Discussed Pareto-type models for risk management and high-layer loss analysis.	Supports the inclusion of Pareto-type reasoning for reinsurance, capital planning, and tail-risk estimation.
[6]	Chavez-Demoulin and Davison (2012)	Reviewed time-series modeling of extremes, dependence, and exceedance behavior.	Supports interpretation of storm-event losses as temporally structured rare outcomes.
[7]	Davison and Huser (2015)	Reviewed statistical theory and applications of extreme-value modeling.	Clarifies why full-body fit and tail-specific inference answer different research questions.
[8]	Davison et al. (2019)	Reviewed spatial-extreme models and environmental applications.	Positions storm-loss modeling within broader environmental statistics and motivates future spatial extensions.
[9]	Davison et al. (2012)	Presented statistical modeling approaches for spatial extremes.	Shows that extreme weather severity may vary across locations and exposure contexts.
[10]	Deng (2021)	Compared Bayesian predictive approaches for natural-disaster loss distributions.	Directly supports lognormal-Pareto and body-tail interpretations for disaster-loss severity.
[11]	Gilleland and Katz (2016)	Presented extRemes 2.0 for weather and climate extreme-value analysis.	Supports reproducible implementation, diagnostic reporting, and transparent EVT specification.
[12]	He et al. (2022)	Estimated value-at-risk using generalized Pareto distributions for high-threshold exceedances.	Supports using GPD models for tail-risk quantities rather than average-loss interpretation only.
[13]	Huser and Wadsworth (2022)	Reviewed advances in statistical modeling of spatial extremes.	Identifies dependence, nonstationarity, and spatial heterogeneity as important future extensions.
[14]	Katz (2010)	Examined statistical methods for climate extremes, high quantiles, and tail uncertainty.	Establishes the post-2010 statistical basis for tail-aware climate and weather risk analysis.
[15]	Kousky (2014)	Reviewed the economic cost of natural disasters and the relevance of adaptation.	Motivates using economic damage as the main severity outcome for rare-event risk estimation.
[16]	Merz et al. (2014)	Reviewed flood risk, climate, exposure, vulnerability, and risk-management implications.	Positions extreme-loss distribution fitting within broader flood and hazard-risk research.
[17]	Naveau et al. (2016)	Modeled low, moderate, and heavy rainfall intensities without relying on one threshold.	Supports the study argument that body behavior and tail behavior may require different treatments.
[18]	NOAA NCEI (2022)	Documents storm-event records including fatalities, injuries, property damage, crop damage, and event classifications.	Provides the empirical foundation and metadata context for the NOAA Storm Events dataset.
[19]	Northrop and Coleman (2014)	Improved threshold diagnostic plots for generalized Pareto modeling.	Supports explicit threshold evaluation rather than arbitrary cutoff selection.

No.	Study	Main focus	Relevance to this manuscript
[20]	Opitz (2016)	Modeled asymptotically independent spatial extremes using Laplace random fields.	Shows that location and spatial dependence can influence extreme-event severity.
[21]	Paprotny et al. (2018)	Developed a large historical flood-loss and exposure database for Europe.	Shows the value of long-term hazard records for understanding natural-disaster loss patterns.
[22]	Papastathopoulos and Tawn (2013)	Proposed extended generalized Pareto models for stable tail estimation.	Shows that GPD tail fitting requires diagnostic support and may need extensions.
[23]	Rootzen and Katz (2013)	Developed design-life level concepts for risk quantification under changing climate conditions.	Links return-level estimation with infrastructure, preparedness, and planning decisions.
[24]	Scarrott and MacDonald (2012)	Reviewed threshold estimation and uncertainty quantification in extreme-value analysis.	Guides the threshold-sensitivity strategy and cautions against untested threshold selection.
[25]	Schneider et al. (2021)	Developed threshold-selection procedures for univariate extreme-value analysis.	Supports reporting sensitivity across multiple thresholds rather than one fixed cutoff.
[26]	Seneviratne et al. (2021)	Assessed weather and climate extreme events in a changing climate.	Provides authoritative climate-risk context for rare-event severity modeling.
[27]	Singirankabo and Iyamuremye (2022)	Modeled extreme rainfall using a generalized Pareto peaks-over-threshold approach.	Provides an environmental application close to the storm-event risk setting of this paper.
[28]	Wirtz et al. (2014)	Discussed disaster-loss databases and challenges of database management.	Highlights the importance of cleaning, reporting standards, and data quality in disaster-loss modeling.
[29]	Zhang et al. (2021)	Applied generalized Pareto modeling to extreme ground motion and seismic hazard.	Shows GPD usefulness when low exceedance probabilities drive risk conclusions.
[30]	Zscheischler et al. (2020)	Developed a typology of compound weather and climate events.	Shows that extreme losses may arise from interacting meteorological drivers rather than isolated hazards only.

The reviewed literature shows that extreme-event modeling requires both distributional comparison and tail-specific diagnostics. Explanation: The present study builds directly on this literature by using full-body fitting, GPD threshold analysis, and rare-outcome interpretation in one empirical framework.

III. RESEARCH GAP

The literature provides strong statistical foundations for extreme-value theory, threshold exceedance modeling, and heavy-tailed loss distributions. However, three gaps remain important for applied rare-event research. First, many empirical studies focus either on theoretical distribution fitting or on domain-specific disaster description, but fewer combine large-scale official storm-event data with a systematic comparison of conventional and extreme-value distributions. Second, full-sample goodness-of-fit measures are often reported without checking whether the selected model accurately represents the far upper tail. This can create a misleading impression of fit when the model describes the center of the distribution but underestimates very rare losses. Third, NOAA storm-event data are often used descriptively, while fewer studies document the full transformation from coded damage strings into positive continuous damage outcomes suitable for likelihood-based probability modeling.

This study addresses these gaps by developing an empirical framework that separates three analytical tasks: cleaning and transforming storm-event losses, fitting candidate probability distributions to positive

damage values, and evaluating tail-risk estimates using both empirical and EVT-based diagnostics. The contribution is not the invention of a new distribution. Rather, the contribution is a reproducible workflow that shows how distribution fitting should be used when the research target is not only average fit but rare, high-impact outcomes.

IV. STUDY CONTRIBUTION

This study makes four contributions.

- First, it provides a dataset-based empirical demonstration of probability distribution fitting using a large storm-event database with 1,962,332 records and 450,823 positive economic-loss observations. This gives the paper an empirical foundation rather than leaving the topic at a theoretical level.
- Second, it compares conventional and heavy-tailed distributions using full-sample metrics and tail-specific diagnostics, showing why a model that performs well overall may still fail for the far tail.
- Third, it applies generalized Pareto peaks-over-threshold analysis at multiple thresholds and reports how high-quantile estimates change under threshold variation.
- Fourth, it links economic damage modeling with rare human-impact outcomes by summarizing deaths and injuries, thereby showing that rare-event modeling is relevant to both financial and social consequences.

V. RESEARCH QUESTIONS

The study is guided by four research questions.

- RQ1: Which probability distributions provide the strongest full-body fit for positive storm-event economic damage?
- RQ2: Do heavy-tailed and extreme-value distributions provide more realistic tail estimates than conventional models?
- RQ3: How sensitive are generalized Pareto estimates to threshold choice?
- RQ4: What do the distributions of deaths and injuries show about rare human-impact outcomes in storm-event data?

VI. RESEARCH HYPOTHESES

The study tests two working hypotheses:

- H1: conventional thin-tailed distributions will underperform for storm-event damage because the data are strongly right-skewed;
- H2: a generalized Pareto peaks-over-threshold model will provide more realistic estimates of upper-tail risk than full-sample conventional models.

VII. PROPOSED METHOD

The proposed method is a distribution-fitting and tail-risk framework for rare-event severity outcomes. The framework follows the logic that empirical losses contain two related but distinct components: a body of relatively common positive losses and a tail of rare high-severity losses. The body can be evaluated using conventional positive distributions, while the tail requires threshold-based extreme-value modeling.

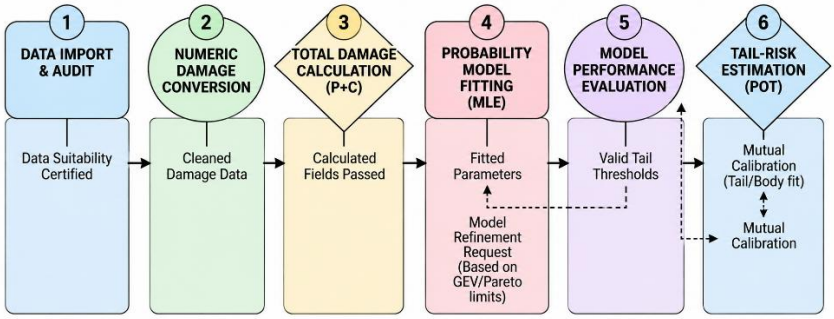


FIGURE 1. Empirical analysis workflow.

FIGURE 2 summarizes the empirical workflow used in the study. Explanation: The procedure is designed to be reproducible because each step moves from raw NOAA records to a measurable statistical output.

The first stage is data import and audit. The Data files are inspected for sample size, variable names, time coverage, and suitability for loss modeling.

The second stage is numeric damage conversion. NOAA damage values may appear as numeric strings or as values with K, M, or B multipliers. These must be standardized before analysis.

The third stage is construction of total damage as property damage plus crop damage.

The fourth stage is model fitting using maximum likelihood estimation.

The fifth stage is model evaluation using information criteria, KS distance, graphical diagnostics, and high-quantile errors.

The sixth stage is tail-risk estimation using generalized Pareto exceedances above selected thresholds.

VIII. METHODOLOGY

1. RESEARCH DESIGN

The study uses a quantitative empirical research design based on probability distribution fitting and extreme-value analysis. The unit of analysis is the storm-event record. The main outcome is positive total economic damage, defined as property damage plus crop damage, converted to numeric dollar values. The supplementary outcomes are total deaths and total injuries, each calculated as the sum of direct and indirect counts.

2. DATASET AND SAMPLE SELECTION

The archive contains NOAA Storm Events data. The event-detail table includes 1,962,332 records from 1950 to 2025. The main distribution-fitting sample retains only records with positive total damage because the objective is severity modeling rather than event occurrence modeling. This produces 450,823 positive damage observations. Zero values are not discarded as unimportant; rather, they are excluded from severity fitting because a zero-inflated occurrence model is a different statistical problem. The present manuscript focuses on the severity distribution conditional on a positive reported loss.

3. DAMAGE TRANSFORMATION

Damage fields are transformed using a multiplier function. Numeric values are retained as dollar amounts, while K, M, and B suffixes are converted to thousands, millions, and billions. Equation (1) shows the conversion rule:

$$D_i = v_i \times m_i, \quad m_i \in \{1, 10^3, 10^6, 10^9\}$$

Here, v_i is the numeric component of the reported value and m_i is the suffix multiplier. Total damage is then calculated as shown in Equation (2):

$$T_i = P_i + C_i$$

where T_i is total damage, P_i is converted property damage, and C_i is converted crop damage.

4. CANDIDATE DISTRIBUTIONS

The study fits normal, lognormal, exponential, Weibull, Pareto, GEV, and GPD models. The normal model is included as a baseline because it is widely known but theoretically weak for nonnegative heavy-tailed data. The lognormal model is included because loss data are often positive and right-skewed. The exponential and Weibull models are included as positive severity distributions. Pareto and generalized Pareto models are included because they are designed for heavy-tailed behavior.

Parameter estimation is performed using maximum likelihood. Equation (3) gives the likelihood objective:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \sum_{i=1}^{n} \log f(x_i \mid \theta)$$

where x_i is a positive damage observation and $f(x_i \mid \theta)$ is the candidate distribution density.

5. MODEL COMPARISON

Models are compared using AIC, BIC, KS distance, and tail-quantile accuracy. AIC and BIC are calculated as shown in Equation (4):

$$AIC = 2k - 2\ell(\hat{\theta}), \quad BIC = k\log(n) - 2\ell(\hat{\theta})$$

where k is the number of estimated parameters, n is the sample size, and $\ell(\hat{\theta})$ is the maximized log-likelihood. The KS distance is shown in Equation (5):

$$D_n = \sup_x \left|F_n(x) - F(x;\hat{\theta})\right|$$

where $F_n(x)$ is the empirical distribution function and $F(x;\hat{\theta})$ is the fitted cumulative distribution function.

6. GENERALIZED PARETO TAIL MODEL

For exceedances above a high threshold u, the generalized Pareto distribution is fitted to $Y = X - u \mid X > u$. Equation (6) gives the GPD distribution function:

$$G(y;\xi,\beta) = 1 - \left(1 + \frac{\xi y}{\beta}\right)^{-1/\xi}, \quad y > 0$$

where ξ is the shape parameter and β is the scale parameter. The unconditional tail quantile is estimated using Equation (7):

$$\mathrm{Q_p} = \mathrm{u} + \frac{\beta}{\xi}\left[\left(\frac{1-\mathrm{p}}{\hat{\mathrm{p}}_\mathrm{u}}\right)^{-\xi} - 1\right]$$

where $\hat{p}_u$ is the empirical exceedance probability above threshold u. Annual maxima are also examined as a supplementary GEV diagnostic using log-transformed annual maximum damage because raw-scale GEV fitting was unstable under extreme rounding and very large event losses.

7. VALIDITY, RELIABILITY, AND ETHICAL CONSIDERATIONS

The study uses publicly available event-level storm data and does not involve human participants or private individual information. The main validity concerns are measurement quality, nominal damage reporting, coded damage conversion, and threshold sensitivity. To improve transparency, the study reports the cleaning rule, positive-damage sample size, candidate distributions, model-comparison criteria, and multiple threshold estimates. Reliability is strengthened by using reproducible transformations and standard likelihood-based fitting procedures. However, NOAA damage amounts are broad reported estimates and are not adjusted for inflation; therefore, the results should be interpreted as distributional analysis of reported nominal damage rather than precise historical economic loss in constant dollars.

IX. RESULTS

1. DATASET AUDIT

The event-detail file contains 1,962,332 records covering 1950 to 2025. Positive property damage is reported in 435,749 records, positive crop damage in 38,439 records, and positive total damage in 450,823 records. Deaths are reported in 14,342 records, while injuries are reported in 24,723 records. This confirms that the dataset is suitable for both economic-loss severity modeling and supplementary rare-outcome analysis.

Table 2 summarizes the analytical stages and outputs.

Stage	Procedure	Output
Data audit	Inspect files, sample size, years, and candidate variables	Confirmed large event-level sample from 1950-2025

Stage	Procedure	Output
Damage cleaning	Convert K/M/B coded damage values to numeric dollars	Numeric property, crop, and total damage variables
Sample definition	Retain positive total damage for severity fitting	450,823 positive loss observations
Distribution fitting	Fit normal, lognormal, exponential, Weibull, Pareto, GPD, and GEV models	Estimated parameters and model diagnostics
Tail-risk analysis	Estimate Q95, Q99, Q99.5, and Q99.9	Quantified rare high-impact loss levels

The procedure translates the raw data into a reproducible severity-analysis sample. Explanation: The separation between event occurrence and positive-loss severity prevents zero values from distorting continuous distribution fitting.

2. DESCRIPTIVE DISTRIBUTION OF ECONOMIC DAMAGE

Positive total damage is highly skewed. The median positive loss is only $10.00 thousand, but the 99th percentile is $10.59 million, the 99.5th percentile is $25.00 million, and the 99.9th percentile is $200.00 million. The maximum reported positive damage is $17.90 billion. These values show that the distribution has a long right tail and that a small share of events dominates the extreme-loss profile.

Table 3. reports descriptive statistics for positive total damage.

Statistic	Total positive damage
Positive damage observations	450,823
Minimum	$0.10
Median	$10.00 thousand
75th percentile	$35.00 thousand
90th percentile	$250.00 thousand
95th percentile	$800.00 thousand
99th percentile	$10.59 million
99.5th percentile	$25.00 million
99.9th percentile	$200.00 million
Maximum	$17.90 billion

The descriptive statistics show a strong separation between ordinary positive losses and extreme losses. Explanation: The median event is small relative to the far tail, so mean-centered models are insufficient for estimating rare, high-impact damage.

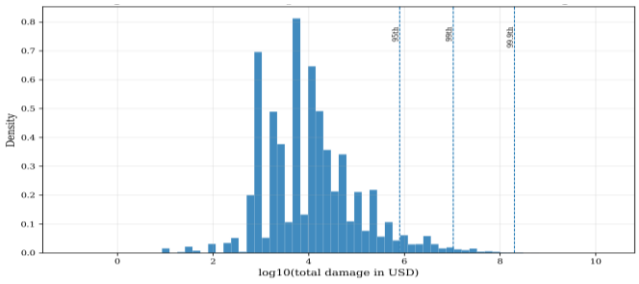


FIGURE 2. Distribution of positive storm-event economic damage.

The figure shows the distribution of positive damage on a log10 scale. Explanation: The log-scale histogram reveals a highly skewed loss distribution with visible concentration at smaller losses and a long upper tail.

3. EVENT TYPES ASSOCIATED WITH THE LARGEST LOSSES

Economic damage is not evenly distributed across storm-event types. Hurricane and typhoon records, flash floods, tornadoes, storm surge or tide, flood events, wildfires, hail, drought, coastal flood, and tropical storms are among the largest contributors to total reported damage.

Table 4 summarizes the top event types by total reported damage.

Event type	Records	Positive damage records	Total damage	Deaths	Injuries
Hurricane (Typhoon)	2,122	1,227	$167.51 billion	1,380	4,252
Flash Flood	105,758	38,411	$102.77 billion	2,005	6,577
Tornado	79,236	48,497	$75.80 billion	6,352	99,364
Storm Surge/Tide	1,651	564	$60.97 billion	44	43
Flood	68,784	20,320	$48.82 billion	759	2,456
Wildfire	9,079	2,192	$40.05 billion	572	2,858
Hail	411,905	34,604	$37.08 billion	29	1,682
Drought	78,977	4,949	$31.28 billion	0	358
Coastal Flood	4,352	752	$21.93 billion	59	17
Tropical Storm	7,110	2,709	$21.37 billion	293	430

The table confirms that extreme losses are concentrated in specific weather hazards. Explanation: Hurricanes, floods, tornadoes, storm surge, and wildfire events form the core empirical context for the upper tail of storm-event damage.

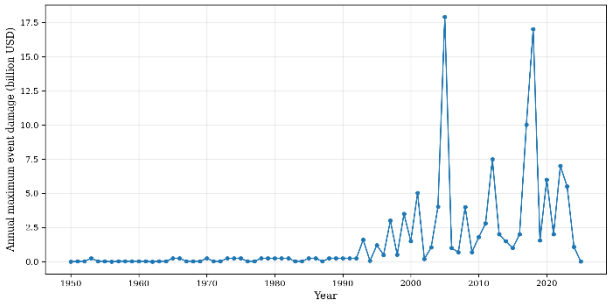


FIGURE 3. Annual maximum storm-event economic damage.

The figure presents the annual maximum positive damage value across the study period. Explanation: Annual maxima vary strongly across years, indicating that rare catastrophic events dominate the temporal profile of reported storm losses.

4. FULL-DISTRIBUTION MODEL FITTING

The model-comparison results show that the lognormal distribution provides the strongest conventional full-body fit, with AIC = 10,422,995 and KS D = 0.101. The Weibull model performs less well than the lognormal but better than the normal, exponential, and Pareto full-sample fits. The normal distribution performs poorly because the data are nonnegative, highly skewed, and dominated by a long right tail. The exponential model is also inadequate because it cannot represent the observed concentration of lower losses and the extreme upper tail simultaneously. A simple full-sample Pareto fit is unstable because the minimum positive damage is very small, making the fitted lower endpoint unrealistic for the full body of the data.

Table 5. reports the main goodness-of-fit results.

Model	Fitted sample	KS D	AIC	BIC
Normal	All positive damage observations	0.491	17,470,381	17,470,403
Lognormal	All positive damage observations	0.101	10,422,995	10,423,017
Exponential	All positive damage observations	0.799	13,712,676	13,712,687
Weibull	All positive damage observations	0.177	10,619,248	10,619,270
Pareto	All positive damage observations	0.492	11,508,290	11,508,313
GPD-POT(95%)	Threshold exceedances above 95th percentile	0.076	747,750	747,766

The table shows that full-body and tail-specific models should not be judged in the same way because the GPD is fitted only to exceedances. Explanation: The lognormal is strongest for the full positive-loss body, while the GPD is strongest for threshold exceedances.

5. TAIL-RISK COMPARISON

Tail-risk estimates reveal a different story from full-body information criteria. The empirical 99th percentile is $10.59 million, and the empirical 99.9th percentile is $200.00 million. The lognormal model estimates the 99th percentile at $2.29 million and the 99.9th percentile at $13.20 million, which means that it underestimates the far upper tail. The Weibull model also underestimates the far tail. The GPD-POT model fitted above the 95th percentile estimates the 99th percentile at $10.50 million and the 99.9th percentile at $203.47 million, closely matching the empirical tail levels.

Table 6 compares empirical and fitted tail quantiles.

Model	Q95	Q99	Q99.5	Q99.9
Empirical	$800.00 thousand	$10.59 million	$25.00 million	$200.00 million
Lognormal	$480.25 thousand	$2.29 million	$4.06 million	$13.20 million
Weibull	$813.07 thousand	$2.72 million	$4.02 million	$8.47 million
GPD-POT(95%)	$800.00 thousand	$10.50 million	$26.12 million	$203.47 million
Normal	$104.95 million	$147.82 million	$163.52 million	$195.88 million

The tail-quantile table shows why full-body fit is not enough for rare-outcome modeling. Explanation: The GPD-POT model is much closer to the empirical far tail than the lognormal and Weibull models, even though the lognormal performs best for the full body.

https://doi.org/10.48161/qjcs.v4n4a91

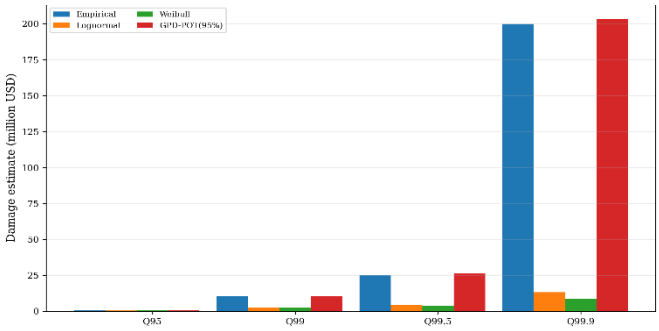


FIGURE 4. Empirical and fitted tail quantiles.

FIGURE 4. Compares empirical tail quantiles with selected model estimates. Explanation: The visual comparison shows that conventional full-sample models underestimate the far upper tail, while the GPD-POT model tracks extreme quantiles more closely.

6. GPD THRESHOLD SENSITIVITY

Threshold sensitivity was evaluated at the 90th, 95th, 97.5th, and 99th percentiles. The 90th percentile threshold uses more exceedances but may include observations that are not sufficiently tail-like. The 99th percentile threshold focuses on the most extreme values but leaves fewer observations for estimation. The 95th and 97.5th percentile thresholds provide a practical balance between bias and variance. The 95th percentile threshold retains 22,385 exceedances and produces a GPD shape estimate of 1.261, indicating a very heavy tail.

Table 7 reports threshold-sensitivity results.

Threshold	Cutoff	Exceedances	Shape xi	KS D	Estimated Q99	Estimated Q99.9
90.0%	$250.00 thousand	37,630	1.367	0.071	$10.23 million	$245.36 million
95.0%	$800.00 thousand	22,385	1.261	0.076	$10.50 million	$203.47 million
97.5%	$2.50 million	10,900	1.096	0.049	$11.35 million	$175.12 million
99.0%	$10.59 million	4,509	1.011	0.095	$10.60 million	$175.60 million

The threshold results show that the estimated 99.9th percentile remains within a high-risk range across thresholds. Explanation: Although threshold choice affects exact estimates, all GPD specifications indicate a very heavy upper tail.

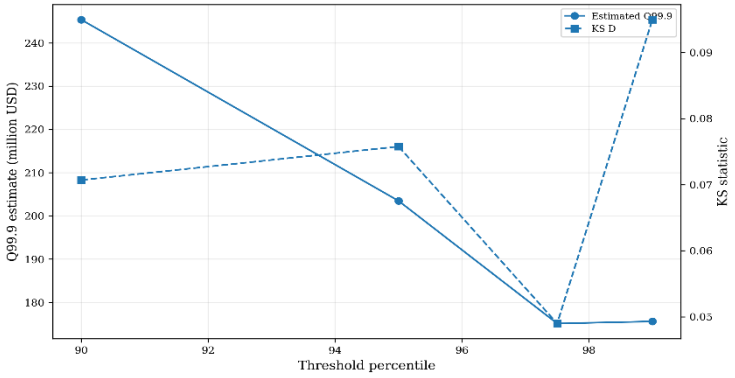


FIGURE 5. Sensitivity of GPD-POT estimates to threshold choice.

FIGURE 5. displays the relationship between threshold choice, the estimated 99.9th percentile, and KS distance. Explanation: The plot shows the practical trade-off between using more exceedances and obtaining a cleaner tail fit.

7. RARE HUMAN-IMPACT OUTCOMES

Deaths and injuries are rarer than economic damage, but they also display heavy concentration. Deaths occur in 14,342 records and injuries in 24,723 records. Among positive fatality records, the median is 1 death, the 99th percentile is 15 deaths, and the maximum is 638 deaths. Among positive injury records, the median is 2 injuries, the 99th percentile is 100 injuries, and the maximum is 2,400 injuries. These values indicate that human-impact outcomes are rare but highly consequential when they occur.

Table 8 summarizes rare human-impact outcomes.

Outcome	Nonzero records	Median positive	95th percentile positive	99th percentile positive	Maximum
Total deaths	14,342	1	5	15	638
Total injuries	24,723	2	24	100	2,400

The table shows that deaths and injuries are sparse but severe in the upper tail. Explanation: Human-impact outcomes support the same general conclusion as damage outcomes: rare event severity cannot be represented well by averages alone.

X. DISCUSSION

The empirical results support the study's central argument: probability distribution fitting for extreme events must distinguish between overall fit and tail fit. The lognormal model performs best among conventional full-sample distributions because it captures positive skewness better than the normal, exponential, Weibull, and simple Pareto alternatives. This is consistent with the general observation that economic losses are nonnegative and multiplicative. However, the lognormal model underestimates the 99th and 99.9th percentiles in this dataset. Therefore, selecting the lognormal model solely on AIC or BIC would reasonably describe the main body of positive damage while providing weak guidance for rare catastrophic losses.

The GPD-POT model provides a more realistic representation of the upper tail. At the 95th percentile threshold, it estimates Q99 and Q99.9 values very close to the empirical quantiles. This does not mean that the GPD model is universally correct or that the 95th percentile is the only defensible threshold. It means that the tail-specific model is better aligned with the risk quantity of interest. For emergency management, insurance planning, and infrastructure design, this distinction is essential. A decision-maker concerned with a 1-in-100 or 1-in-1000 type loss should not rely on a distribution selected only because it fits the center of the data.

The results also show that a simple full-sample Pareto model is not automatically appropriate. Although Pareto-type behavior is theoretically relevant for extremes, fitting a Pareto distribution to all positive observations is not the same as fitting a Pareto-type model to the tail. In the present data, the minimum positive loss is extremely small, making a full-sample Pareto fit unrealistic. This reinforces the importance of threshold-based modeling: the Pareto logic belongs in the upper tail, not necessarily across the entire positive loss distribution.

The rare human-impact results strengthen the interpretation. Deaths and injuries occur in a small share of storm-event records, but their positive distributions are also concentrated in the upper tail. This suggests that rare-outcome modeling should not be limited to financial damage. A complete disaster-risk framework should model economic severity, deaths, injuries, and possibly location-specific exposure. However, the statistical treatment of counts may require zero-inflated, hurdle, negative-binomial, or discrete extreme-value models, which are beyond the main scope of this paper.

The findings have two broader implications. First, distribution fitting should be treated as a diagnostic process rather than a mechanical ranking exercise. The best model depends on whether the research question concerns the full body, the center, or the upper tail. Second, rare-event modeling benefits from reporting multiple diagnostics: AIC and BIC, KS distance, graphical plots, quantile comparison, and threshold sensitivity. No single statistic fully captures fit quality for heavy-tailed data.

XI. PRACTICAL AND POLICY IMPLICATIONS

For disaster-risk management, the findings suggest that planning should be based on tail-aware estimates rather than average damages. The empirical 99.9th percentile of positive damage is $200.00 million, while the maximum reported event damage is $17.90 billion. These values are far above the median positive loss of $10.00 thousand. Emergency planners, insurers, infrastructure agencies, and public-finance analysts should therefore use models that explicitly evaluate high quantiles and exceedance probabilities.

For researchers using NOAA storm-event data, the study provides a practical cleaning rule and a model-comparison template. Damage values should be converted from coded strings, positive-loss severity should be modeled separately from zero-loss occurrence, and results should be reported with full-body and tail-specific diagnostics. For journals, the study demonstrates how a general probability-distribution topic can be turned into a dataset-based empirical manuscript with transparent methods and numerical findings.

XII. LIMITATIONS AND FUTURE RESEARCH

The study has limitations. First, NOAA damage values are reported estimates and are not adjusted for inflation. Therefore, the analysis models reported nominal damage rather than constant-dollar economic loss. Future research should convert historical damage to real values using an appropriate price index or hazard-specific normalization procedure. Second, the study focuses on univariate severity modeling. Storm losses are influenced by event type, location, exposure, population, infrastructure, and reporting practices. Future work should estimate stratified or covariate-adjusted tail models. Third, the main distribution-fitting sample includes only positive damage observations. A full occurrence-severity framework would model the probability of nonzero loss and the severity of positive loss jointly. Fourth, GEV modeling of annual maxima is treated as supplementary because annual maximum damage values are highly variable and affected by large historical events. Future studies may use regional or event-type-specific block maxima to improve stability.

XIII. CONCLUSION

This study advanced a tail-aware probability-distribution fitting framework for modeling extreme weather losses and rare outcomes using NOAA Storm Events data. Based on 1,962,332 storm-event records from 1950 to 2025 and 450,823 positive economic-loss observations, the findings confirm that storm-loss severity is driven not by average behavior but by a limited number of high-impact events in the far upper tail. Positive damage had a median of $10.00 thousand, a 99th percentile of $10.59 million, a 99.9th percentile of $200.00 million, and a maximum reported loss of $17.90 billion, revealing a scale gap that mean-centered models cannot represent adequately. The model-comparison results provide the main methodological conclusion. The lognormal distribution produced the strongest conventional full-body fit because it captured positive skewness better than normal, exponential, Weibull, and full-sample Pareto alternatives; however, it substantially underestimated catastrophic upper-tail losses. By contrast, the generalized Pareto peaks-over-threshold model, particularly at the 95th percentile threshold, generated tail estimates much closer to the empirical 99th and 99.9th percentiles. Therefore, distribution selection for rare outcomes should not depend only on AIC, BIC, or whole-sample goodness-of-fit statistics. A defensible rare-event model must evaluate whether the fitted distribution represents the risk quantity that matters most: the probability and severity of extreme exceedances. The study contributes both a statistical message and an applied risk-management message. Statistically, body and tail modeling should be treated as complementary tasks. Practically, emergency planning, insurance reserving, infrastructure protection, and climate-risk assessment require tail-specific diagnostics, threshold sensitivity analysis, and high-quantile reporting. Supplementary evidence

from deaths and injuries further confirms that rare human-impact outcomes are sparse but highly concentrated when they occur. Overall, the study demonstrates that transparent data cleaning, distributional comparison, EVT-based tail modeling, graphical diagnostics, and practical interpretation can convert probability fitting into a rigorous empirical framework for low-probability, high-impact events.

XIV. DATA AND CODE AVAILABILITY

The empirical analysis is based on the uploaded Kaggle archive containing NOAA Storm Events bulk CSV files. The analysis used the event-detail file as the main dataset and retained the fatalities and locations files as supplementary sources. The computational workflow included damage conversion, positive-loss filtering, maximum-likelihood distribution fitting, goodness-of-fit comparison, GPD threshold sensitivity, and figure generation. Data are publicly available through the NOAA Storm Events Database and Kaggle dataset mirror, subject to the terms and documentation of those sources.

Author Contributions

The author conducted the conceptualization, methodology, data analysis, investigation, writing, review, editing, and final approval of the manuscript.

Funding

This research received no external funding.

Data Availability

The dataset will be available from the author upon reasonable request.

Conflicts of Interest

The author declares no conflict of interest.

REFERENCES

[1]. AghaKouchak, A., Easterling, D., Hsu, K., Schubert, S., & Sorooshian, S. (Eds.). (2013). Extremes in a changing climate: Detection, analysis and uncertainty. Springer.
[2]. Aminzadeh, M. S., & Deng, M. (2017). Bayesian predictive modeling for exponential-Pareto composite distribution. Variance, 12(1), 59-78.
[3]. Bader, B., Yan, J., & Zhang, X. (2018). Automated threshold selection for extreme value analysis via ordered goodness-of-fit tests with adjustment for false discovery rate. The Annals of Applied Statistics, 12(1), 310-329.
[4]. Beirlant, J., Fraga Alves, I., Gomes, I., & Meerschaert, M. M. (2016). Extreme value statistics for truncated Pareto-type distributions. Extremes, 19, 1-29.
[5]. Charpentier, A., & Flachaire, E. (2021). Pareto models for risk management. In G. Dufrenot and T. Matsuki (Eds.), Recent econometric techniques for macroeconomic and financial data (pp. 125-145). Springer.
[6]. Chavez-Demoulin, V., & Davison, A. C. (2012). Modeling time series extremes. REVSTAT - Statistical Journal, 10(1), 109-133.
[7]. Davison, A. C., & Huser, R. (2015). Statistics of extremes. Annual Review of Statistics and Its Application, 2, 203-235.
[8]. Davison, A. C., Huser, R., & Thibaud, E. (2019). Spatial extremes. In A. E. Gelfand, M. Fuentes, J. A. Hoeting, & R. L. Smith (Eds.), Handbook of environmental and ecological statistics (pp. 711-744). CRC Press.
[9]. Davison, A. C., Padoan, S. A., & Ribatet, M. (2012). Statistical modeling of spatial extremes. Statistical Science, 27(2), 161-186.
[10]. Deng, M. (2021). Bayesian predictive analysis of natural disaster losses. Risks, 9(1), 12.
[11]. Gilleland, E., & Katz, R. W. (2016). extRemes 2.0: An extreme value analysis package in R. Journal of Statistical Software, 72(8), 1-39.
[12]. He, Y. I., Chen, R., Lee, T. T., & Lou, K. (2022). Risk analysis via generalized Pareto distributions. Risks, 10(6), 117.
[13]. Huser, R., & Wadsworth, J. L. (2022). Advances in statistical modeling of spatial extremes. WIREs Computational Statistics, 14(1), e1537.
[14]. Katz, R. W. (2010). Statistics of extremes in climate change. Climatic Change, 100(1), 71-76.

[15]. Kousky, C. (2014). Informing climate adaptation: A review of the economic costs of natural disasters. Energy Economics, 46, 576-592.
[16]. Merz, B., Aerts, J. C. J. H., Arnbjerg-Nielsen, K., Baldi, M., Becker, A., Bichet, A., Bloschl, G., Bouwer, L. M., Brauer, A., Cioffi, F., Delgado, J. M., Gocht, M., Guzzetti, F., Harrigan, S., Hirschboeck, K., Kilsby, C., Kron, W., Kwon, H. H., Lall, U., Merz, R., Nissen, K., Salvatti, P., Swierczynski, T., Ulbrich, U., Viglione, A., Ward, P. J., Weiler, M., Wilhelm, B., & Nied, M. (2014). Floods and climate: Emerging perspectives for flood risk assessment and management. Natural Hazards and Earth System Sciences, 14, 1921-1942.
[17]. Naveau, P., Huser, R., Ribereau, P., & Hannart, A. (2016). Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. Water Resources Research, 52(4), 2753-2769.
[18]. NOAA National Centers for Environmental Information. (2022). Storm Events Database. National Oceanic and Atmospheric Administration.
[19]. Northrop, P. J., & Coleman, C. L. (2014). Improved threshold diagnostic plots for extreme value analyses. Extremes, 17, 289-303.
[20]. Opitz, T. (2016). Modeling asymptotically independent spatial extremes based on Laplace random fields. Spatial Statistics, 16, 1-18.
[21]. Paprotny, D., Sebastian, A., Morales-Napoles, O., & Jonkman, S. N. (2018). HANZE: A pan-European database of exposure to natural hazards and damaging historical floods since 1870. Earth System Science Data, 10, 565-581.
[22]. Papastathopoulos, I., & Tawn, J. A. (2013). Extended generalized Pareto models for tail estimation. Journal of Statistical Planning and Inference, 143(1), 131-143.
[23]. Rootzen, H., & Katz, R. W. (2013). Design life level: Quantifying risk in a changing climate. Water Resources Research, 49(9), 5964-5972.
[24]. Scarrott, C., & MacDonald, A. (2012). A review of extreme value threshold estimation and uncertainty quantification. REVSTAT - Statistical Journal, 10(1), 33-60.
[25]. Schneider, L. F., Krajina, A., & Krivobokova, T. (2021). Threshold selection in univariate extreme value analysis. Extremes, 24, 881-913.
[26]. Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S. M., Wehner, M., & Zhou, B. (2021). Weather and climate extreme events in a changing climate. In V. Masson-Delmotte et al. (Eds.), Climate change 2021: The physical science basis. Cambridge University Press.
[27]. Singirankabo, E., & Iyamuremye, A. (2022). Modeling extreme rainfall events in Kigali city using generalized Pareto distribution. Meteorological Applications, 29(3), e2076.
[28]. Wirtz, A., Kron, W., Loewe, P., & Steuer, M. (2014). The need for data: Natural disasters and the challenges of database management. Natural Hazards, 70, 135-157.
[29]. Zhang, M., Wen, R., Ren, Y., & Ji, K. (2021). Application of generalized Pareto distribution for modeling extreme ground motion in probabilistic seismic hazard analysis. Natural Hazards, 107, 1527-1547.
[30]. Zscheischler, J., Martius, O., Westra, S., Bevacqua, E., Raymond, C., Horton, R. M., van den Hurk, B., AghaKouchak, A., Jezequel, A., Mahecha, M. D., Maraun, D., Ramos, A. M., Ridder, N. N., Thiery, W., & Vignotto, E. (2020). A typology of compound weather and climate events. Nature Reviews Earth & Environment, 1, 333-347.