



Extreme Weather Loss and Damage Estimation Using a Hybrid Simulation Technique

Charles Doktycz¹ · Mark Abkowitz¹ · Hiba Baroud¹

Accepted: 18 July 2022 / Published online: 1 August 2022
© The Author(s) 2022

Abstract History has shown that occurrences of extreme weather are becoming more frequent and with greater impact, regardless of one's geographical location. In a risk analysis setting, what will happen, how likely it is to happen, and what are the consequences, are motivating questions searching for answers. To help address these considerations, this study introduced and applied a hybrid simulation model developed for the purpose of improving understanding of the costs of extreme weather events in the form of loss and damage, based on empirical data in the contiguous United States. Model results are encouraging, showing on average a mean cost estimate within 5% of the historical cost. This creates opportunities to improve the accuracy in estimating the expected costs of such events for a specific event type and geographic location. In turn, by having a more credible price point in determining the cost-effectiveness of various infrastructure adaptation strategies, it can help in making the business case for resilience investment.

Keywords Climate change · Cost estimation · Extreme weather · Loss and damage · Loss normalization · Monte Carlo simulation

1 Introduction

Citing dire consequences, the Intergovernmental Panel on Climate Change has recently reported a compelling need to accelerate global climate change adaptation

(Masson-Delmotte et al. 2021). This finding is based on substantial evidence that the effect of global warming is causing increased frequency and severity of a variety of weather extremes, which are likely to persist for at least several decades. The consequential impacts of these developments can be expected to cause additional economic, social, and environmental suffering, creating an imperative for society to invest, post haste, in adaptation strategies that reduce community vulnerability and strengthen resilience.

The impact of an extreme weather event may be felt at several levels—local, regional, national, and global. A single catastrophic event can lead to human casualties, property damage, loss of assets, community disruption, severed supply chains, mental health issues, and other negative economic, social, and environmental impacts (Botzen and Van Den Bergh 2009). The magnitude of these impacts brings into focus the need for policy responses to prepare for what is considered a new era of natural catastrophes and ease the burden of these occurrences on society (Kunruether 2008).

Yet, widespread implementation of adaptation strategies only becomes possible if one can make a business case for investing in such initiatives. This requires an ability to quantify the full range of loss and damage (L&D) associated with an extreme weather event, such that the benefits accrued (that is, L&D avoidance) from deploying an adaptation strategy can be assessed relative to the cost of implementation. Publicly available L&D data are typically focused on quantifiable tangible direct costs, with little or no consideration of other direct costs, commonly known as tangible indirect costs and intangible costs. This limits our complete understanding of the impacts of an extreme weather event. Empirical L&D data can also suffer from inconsistent definitions and potential biases (Doktycz and Abkowitz 2019). Furthermore, utilization of L&D data for decision making follows the same difficulties as with data gathering.

✉ Charles Doktycz
charles.j.doktycz@vanderbilt.edu

¹ School of Civil and Environmental Engineering, Vanderbilt University, Nashville, TN 37235, USA

Overcoming these deficiencies is a fundamental aspect of our effort, and the reason for developing a hybrid simulation approach. As development of such an approach is ultimately intended to help practitioners make more risk-informed resource investment decisions, we are mindful that local authorities require a simple, practical framework for decision making given limited time and resources. Therefore, it must balance between not requiring an overabundance of technical requirements while also effectively using available data. We develop and apply this methodology using L&D data obtained from the National Oceanic and Atmospheric Administration (NOAA) Storm Events Database, subsequently normalizing the data to allow for spatial/temporal comparisons, fitting the data to statistical distributions, and finally using Monte Carlo simulation (MCS) to return L&D costs of specific extreme event types in a region over a specified time horizon.

The objective of the study described in this article was to generate a representative model of L&D costs for specific regions and segmented by extreme weather type, and subsequently to employ the results in a Monte Carlo simulation. This study took a unique approach in the utilization of L&D databases, to develop a cost model for hazards within the United States. Development of an accurate base historical cost model will then allow for additional cost projections and analysis in future work. As mentioned, future climate projections along with other costs such as indirect or intangible damages will help to create a more accurate picture of the impacts society will face due to climate change. This work highlights an overlooked data source, loss and damage data can provide insight towards climate impacts through use of another tool to improve cost benefit considerations for adaptation decisions in the face of future uncertainty. By building this base model, it provides opportunities to consider future climate change scenarios for cost analysis where input of different parameters can influence how the costs may change in the future.

2 Literature Review

When investigating trends in extreme weather events over time, it is important to account for development and wealth that may have occurred during that same period, as this affects the level of L&D exposure. To date, there have been several approaches used to normalize L&D data for spatial and temporal aggregation, with varying results. Although normalization is yet a perfect science, it is a crucial step in data analysis when comparing results by ensuring that the data can be compared consistently across all records in the database.

Weinkle et al. (2018) performed a study of normalized losses due to hurricane landfall for the period from 1900 to

2017 in the continental United States. Multiple normalization methodologies were utilized in this effort, including adjusting the historical loss data for inflation, per capita wealth, and the population of affected counties. These methodologies tried to account for wealth added in the areas in which the storms took place by using a conventional approach to L&D data normalization, one which focused on the economic value of a region. This was based on the concept that as an area develops over time, it generates more valuable assets, which increases the possible L&D outcomes that did not exist in years prior. However, this methodology does not account for the evolution of technology, building codes, improved knowledge, and investment in risk adaptation that might reduce the L&D realized in more recent extreme weather events. Furthermore, the rate of inflation or the rate of value increase over time used in these normalization methods could mask any concurrent rate changes in extreme weather frequency and severity.

Nordhaus (2010) found that annual hurricane damage in the United States increased by USD 10 billion (at 2005 incomes), 0.08% of gross domestic product (GDP), and is possibly an underestimate. The findings were based on three primary factors: (1) number of storms; (2) maximum wind speed at landfall; and (3) GDP. The assumption in development of the hurricane damage function was that damage per storm is conditional on wind speed and proportional to nominal GDP. This assumption is based on the grounds of economic growth, assuming no changes in technology and the location and structure of economic activity. His argument for why this is likely an underestimate of damage is that it omits consideration of the social impact of the destruction of communities, as well as the culture and history in the impacted areas.

In a study of various disaster trend analyses in South Korea, Choi et al. (2019) grouped normalization methods into those that used inflation, wealth, and societal factors. It was determined that each hazard type had its own characteristics and disaster-specific variables to better fit the normalization. Generally speaking, it was found that a larger spatial scope required a more simple and general normalization methodology, while a longer time period improved the quality of the statistical analysis.

Social vulnerability is a difficult metric to develop due to the array of variables that may be used to establish a representative factor. To address this consideration, the Social Vulnerability Index (SoVI) was developed by Cutter et al. (2003); it contains around 30 U.S. Census variables, including socioeconomic, household composition and disability, minority status and language, housing type, and transportation. This index provides a powerful normalization tool because the index utilizes consistent rating scales that can be employed in comparisons between regions in time or space, thus solving a major hurdle in disaster cost normalization.

Additionally, social vulnerability can be measured using the U.S. Center for Disease Control (CDC) Social Vulnerability Index (SVI), which measures a community's vulnerability to respond to and recover from disasters by ranking each census tract based on 15 social factors grouped into four themes (socioeconomic status, household composition, race/ethnicity/language, and housing/transportation). The result is a vulnerability score between 0 (least vulnerable) and 1 (most vulnerable). The CDC SVI and SoVI are two of the most commonly used vulnerability indices. Various studies have compared vulnerability indices and evaluated the contribution of different factors to the assessment of community vulnerability (Flanagan et al. 2011). The hybrid simulation technique developed in this study uses the CDC SVI, which constructs the data in a percentile rank (Tarling 2017).

Loss and damage data normalization must explain the change in exposure and vulnerability within and between areas of comparison, while controlling for the change in the value of a dollar over time. There is no standard method for cost normalization and much of it depends on the spatial region and temporal setting of analysis. The normalization input factors are dependent on the type of analysis performed and is limited to the available data. It is clear normalization for L&D must account for the socioeconomic vulnerability of an area combined with the potential exposure. Furthermore, the dollar values must be adjusted to reflect inflation that has occurred over the time period in question.

Effective disaster management policies require a full understanding of the cost of disasters. Direct damage estimations can help to provide insight from which key vulnerable sectors and mitigating factors can be determined. Furthermore, these estimates can be built upon with other costs, such as indirect, intangible damages, and future long-term climate predictions to understand the full scope of risk a community face (Botzen et al. 2020). Common natural hazard-related disaster modeling techniques build disaster loss and damage models based on physical characteristics, as in the case of catastrophe modeling. Although these models often derive costs in the form of annual expected damages, L&D databases provide actual data that can be utilized for analysis.

As an example of the application of L&D databases, Kahn (2005) examined the direct impacts of natural hazards measured by fatalities, using raw data on deaths from the Centre for Research on the Epidemiology of Disasters (CRED). Through analyzing deaths from different disasters in 57 countries, conclusions towards the role of income, geography, and institutions in minimizing death counts were made. Other studies utilize the CRED database to compare hazards on a global scale to understand the drivers of risk from extreme weather events (Shen et al. 2018). It is worth noting that events are only included in the CRED database if they meet the database classification of an extreme event,

which includes ten or more fatalities, declaration of a state of emergency, hundreds or more reported affected, or a call for international assistance. The following presented methodology uses L&D resources to develop a model for natural hazards of all cost ranges to aid decision making for future climate adaptation to natural hazard-related disasters.

3 Model Development

The quality of extreme weather data and uncertainty in how they are collected and reported make loss and damage data difficult to utilize at a local level, particularly due to sample size limitations when the data are segmented by extreme weather type. For this reason, some level of data aggregation is recommended; one means for doing so is to adopt a geographical region approach, such as has been defined by NOAA climatological divisions (Fig. 1). These nine regions are considered to be internally climatically consistent and therefore useful for putting climate anomalies into a historical perspective (Karl and Koss 1984). Using a regional approach while maintaining internal climate consistency is an important consideration in creating a more robust sample of extreme weather events from which L&D estimation can be performed. This does not suggest, however, that smaller geographical units, such as state or county, should be ignored if a sufficient sample size exists to provide a more resolute view of L&D. Since each region is considered similar climatologically, the grouping of data by this categorization was the initial starting point to expand the modeling sample size. However, in cases where it was determined that an adequate sample size was available for a smaller geographical unit, individual state data analyses were also performed.

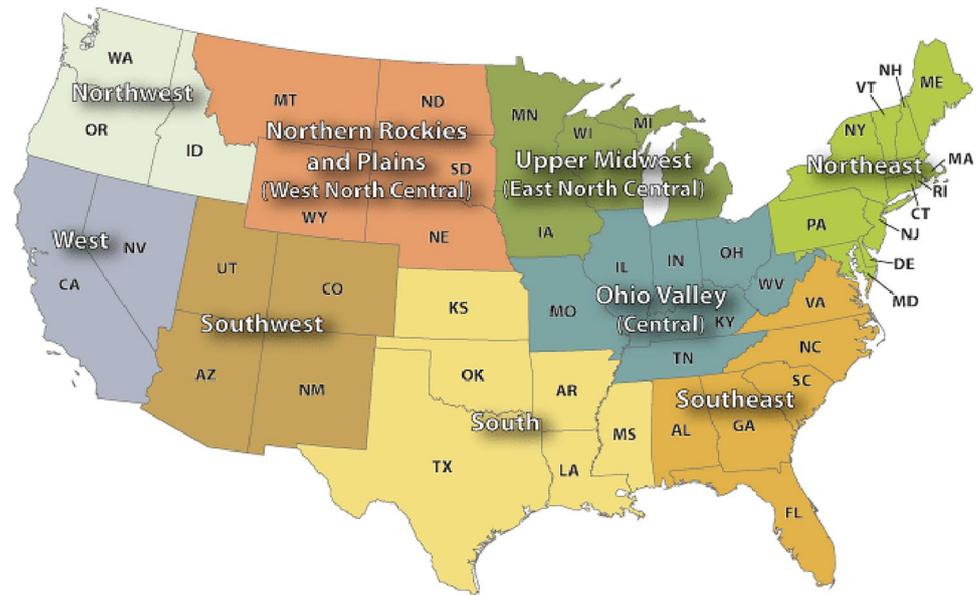
Socioeconomic status can be a meaningful indicator of risk exposure to extreme weather impacts, and such information is available at the county level as provided by the U.S. Census Bureau. Combining this information with the extreme weather profile for the area of interest can provide a more downscaled perspective of the risk with as suitable a level of disaggregation as the data allow.

We limited our use of extreme weather event data from the NOAA Storm Events Database to the period of 2000 to 2019. This time period is associated with an event history that reflects more recent trends, and conforms with a stretch of time where NOAA maintained consistency in how different extreme weather events were recorded.

The NOAA database includes 49 different event types, with each record containing date, location (state and county), property damage, crop damage, injuries, and fatalities. In our methodology, we aggregated the 49 event types into "main event" categories, following the Integrated Research on Disaster Risk (IRDR) peril classification and hazard glossary, as shown in Integrated Research on Disaster Risk (2014).

Fig. 1 National Oceanic and Atmospheric Administration (NOAA) climatological divisions. *Source* Karl and Koss (1984)

U.S. Climate Regions



The main event categories consisted of the following: Earthquake, Volcanic Activity, Flooding, Landslide, Wave Action, Convective Storm, Tornado, Winter Convective Storm, Heat, Cold, Fog, Tropical Cyclone, Drought, and Wildfire. These events were further sorted by the state and county in which they occurred along with the year in which the event took place, with the goal of normalizing the data based on these factors.

A simple, scalable L&D normalization equation that is both spatially and temporally consistent is applied to the cost values. As outlined in the literature review, there are a variety of normalization methods implemented for trend analysis of natural hazard costs. For example, Weinkle et al. (2018) developed the following normalization equation (Eq. 1). First, the cost total (D_y) is inflation adjusted with the adjustment variable (I_y), then it is further normalized with the real wealth per capita of the impacted area ($RWPC_y$) and a county population adjuster ($P_{2018/y}$). The indices functioned as multiplicative indices to generate the normalized cost data (D_{2018}) as shown in Eq. 1.

$$D_{2018} = D_y \times I_y \times RWPC_y \times P_{2018/y} \tag{1}$$

A similar approach is used in this study with a focus on larger spatial scale and multiple natural hazards. The normalized damage, D_n , is calculated in Eq. 2. To adjust for inflation, the monetary value of property damage was adjusted to the CPI for calendar year 2018 from the year the event actually occurred. The CPI data were retrieved from the Bureau of Labor Statistics at the U.S. Department of Labor. The SVI is included to account for the capability of a household to withstand event impact. Finally, population

density is included to explain the magnitude of the impact, based on the assumption that more people exposed increases L&D potential.

$$D_n = D_i \times SVI_{(y,c)} \times PD_{y,c} \tag{2}$$

In Eq. 2, D_i is the inflation adjusted to 2018 event cost, $SVI_{(y,c)}$ is the CDC SVI percentile rank in year y for county c , and PD is the population density percentile rank based on event year and county. The CDC SVI is one of the most widely used indicator models for social vulnerability (Wood et al. 2021). The CDC SVI database has indices from the year 2000, 2010, 2014, 2016, and 2018 and the CDC will continue to release new data every other year. At the time of this study, the 2020 data had not been released. In this application, metrics for the years between CDC SVI release years are linearly extrapolated to fill in during normalization. The CDC SVI uses a percentile ranking for each census tract, which allows for easily interpretable data and as a result shows the spread of vulnerability more evenly without explicitly displaying vulnerability outliers (Tarling 2017). In this normalization approach the data are grouped by county rather than census tract since the loss and damage cost data are provided at the county level. This county grouping would account for the potential loss of census track outliers the percentile ranking may miss since the values would be averaged out across a county grouping. The percentile rank functions as a consistent comparative metric facilitating direct comparison across counties in the United States and across the years in which the data are sampled to account for changes in socioeconomic vulnerabilities over time. In order to keep the normalization equation variables consistent, the population

density values were also converted to a percentile rank. This ranking applies a homogenous vulnerability index for comparison between counties. Reducing vulnerability outliers keeps general vulnerability patterns for the impacted area but prevents unique locations from entirely changing the shape of the damage data. As loss and damage data become more resolute and accurate to impacted locations, comparison at the census tract level may be necessary, in which case vulnerability at a smaller scale will be crucial to understand.

If the desired outcome consists of cost estimates over a period of time, understanding the general distribution of costs and the nature of extreme weather occurrence in the location of interest requires statistical representation to develop a model for cost estimation. In order to account for the uncertainty in L&D estimation, we fit the natural logarithm of the data to probability distributions from which potential L&D can be generated using Monte Carlo simulation. The normalized data for a specific extreme event type and location were fitted against 85 different candidate functions using the Kolmogorov-Smirnov test to determine the best fit. Once a best-fit distribution was obtained for a specific extreme weather type and region, that function was subsequently used to represent the extreme weather type and region profile in the succeeding model simulation.

During this process, it was discovered that the fitted distributions typically did not capture rare and highly damaging events, leading to distributions with infinite moments and failing to accurately replicate the historical data consistently. This issue was addressed by Blackwell (2014), who observed a USD 200 billion difference in damages based on the distribution used to represent the hurricane impact. This is important because it strongly impacts the cost-benefit analysis results in justifying whether to invest in risk mitigation measures.

Based on the extreme value theory (EVT), using a separate distribution to replicate heavy tail events is preferred for modeling extreme losses (Katz 2020). The EVT has a growing application in measuring high-cost losses from natural hazards, for example, modeling windstorm, rainfall, wildfire, earthquake, and snowfall losses. In many cases, the goal of the application is to price the events for future expectations and probabilities to inform decision making. For example, Zimbidis et al. (2007) measured earthquake risk via the EVT to assign respective probabilities to the extreme high-cost events. Recent application has seen a similar approach using the EVT to compute probabilities of unobserved rare heatwaves not seen in historical records (French et al. 2018).

When looking at cost extremes, studies often use a specific cost threshold to separate high-cost events from less costly events to build the end tail distributions, specifically in the case of the EVT because of the focus on modeling extremes. In one particular example, McNeil who used the EVT to estimate end tail loss of Danish wildfires fitted

multiple distributions to find a single best performing distribution of extreme events (extreme events were classified as events over one million Danish Krone) from 1980 to 1990 (McNeil 1997). In our study, the application of the EVT led to similar results when compared to fitting the individual distributions to end tail. We do see, however, the need for more robust testing of EVT distributions for application within our hybrid simulation technique as a topic of future research.

Rather than establishing a specific monetary threshold as done in other EVT studies, we define the top 10% of the data for each region/hazard combination as high-cost events to be separated for data fitting in order to more accurately represent the dynamic of the data at the end tail in a region. This cutoff was implemented to both represent the unique features of the specific region/hazard combination and ensure that a large enough sample size was obtained to accurately fit the data. Furthermore, although many end tail fitting methodologies focus on fitting a single distribution to a single event data space, our approach fit many different hazard and regional data combinations to different distributions depending on the best fit found during testing. The value of the threshold can be revised and evaluated using a sensitivity analysis to determine the most appropriate threshold for different regions and hazards.

Taking the most extreme instances and fitting them separately allows for a more accurate characterization of these events. The reason for the selection of the top 10% part of the loss and damage cost distribution is that the “end tail” behavior is a unique problem due to the limited number of low probability, high consequence events in the sample. Furthermore, the top 10% of the distribution was selected using the Kolmogorov-Smirnov test in order to include enough event samples to create a confident distribution fit that can be consistently applied to each event and location. An example of the two-step data fitting process is presented in Fig. 2. The best-fit distributions for the convective storms in the Midwest were determined to be the generalized logistic and the Johnson SB distribution. This process was performed for each hazard/region combination, resulting in different distributions for each combination, depending on the distribution that fit best when tested.

For example, along with convective storms, the following list consists of the other best-fit distributions for hazards in the Midwest:

- Winter convective storm: A T-distribution for the nonextreme events and an exponential power distribution for the end tail;
- Flooding: Log-normal distribution for the nonextreme best-fit and a Gompertz (truncated Gumbel) distribution for the end tail distribution.

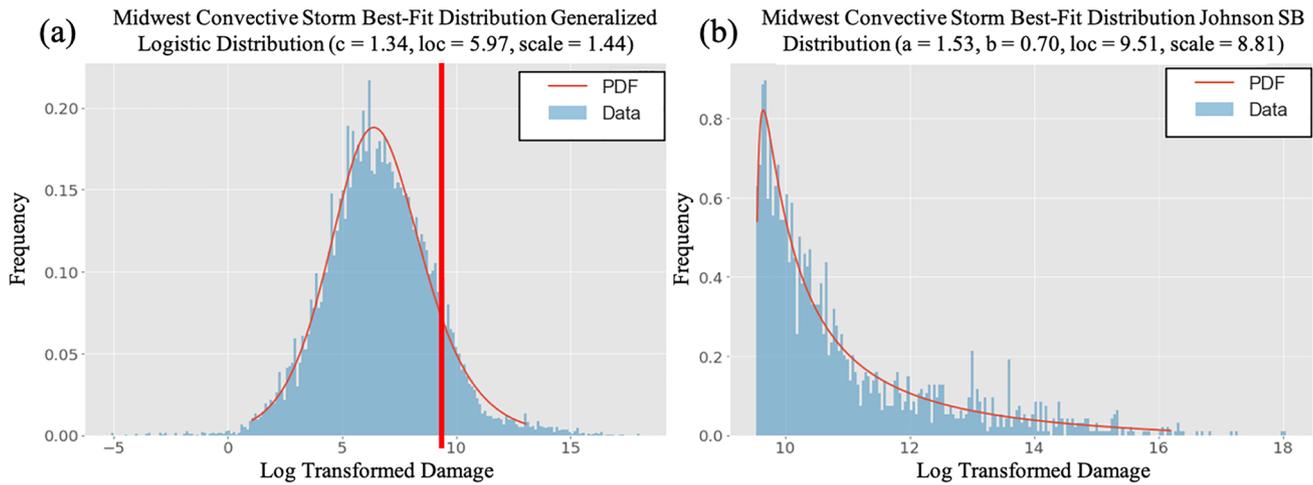


Fig. 2 Two-step data fitting process. The log-normalized data for Midwest convective storms was fitted across the entire data (a) and then for the events in the top 10th percentile (b). The vertical line

before 10 on a indicates the top 10% of the data that was fitted again on the right (b). The red curves in a and b represent the probability density function (PDF) of the best-fit distribution.

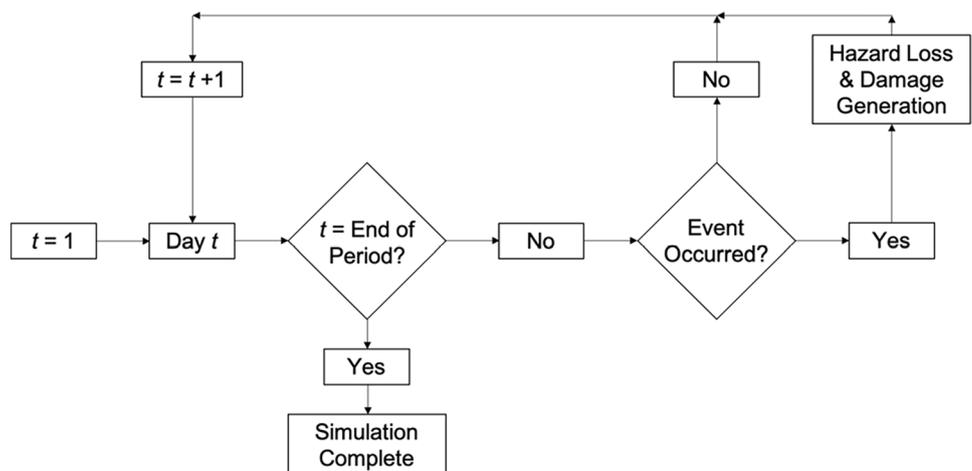
The top 10% of the data was fitted following the same process as previously described, using the Kolmogorov-Smirnov test to determine the best fit for the distribution. In many cases, the end tail distribution required truncation prior to distribution fitting in order to replicate the historical distribution. This distribution was then tested by randomly generating samples to ensure that the fit is a realistic representation of the data due to frequent cases of heavy tailed distributions resulting in disaster costs greater than the wealth in the represented area. The fitting process also occasionally required more samples than the number of samples in the top 10%, in which case the threshold was expanded to include additional events.

Once the two distributions were created, one to represent the lower 90% of the data and another for the top 10%, these distributions were placed into a Monte Carlo simulation (MCS). The simulation uses historical event frequencies to

establish the probability of occurrence of a specific extreme weather event type in the area of interest by dividing the number of records for each extreme weather event type by the observation period (in days). The historical L&D data along with the fitted distributional form from the previous step were used to generate a random cost outcome from the distribution to represent L&D.

This model was then utilized in simulation to generate samples that represent a 20-year period for the selected location and extreme weather category (Fig. 3). The 20-year period was selected to directly compare to the historical data. Each day (time = t) is represented in the simulation as an extreme weather event either happening or not, which is determined by a random outcome based on the historical event frequency (total historical events divided by the total number of days in the historical data recording period). If an event does not occur, the simulation moves

Fig. 3 Simulation flowchart for a single hazard/location



to the next day ($t = t + 1$). Should an event occur, however, the model then determines the associated costs based on the corresponding event type L&D distribution. Once this is determined, the model moves to the next day in the simulation and continues this process until completion of the 20-year period. This creates a single sample and is repeated 1000 times in the MCS. The sample size of 1000 was determined experimentally as convergence towards a relative mean value was found in the model with reasonable computational demand.

The hazard loss and damage generation step in Fig. 3 uses the fitted distributions from the previous step to generate the cost based on the historical data. This process flow is displayed in Fig. 4. The variable C represents the realized cost in the simulation, $C_{0.90}$ represents the 90th percentile cost of the L&D of the specific extreme weather event type and region. These costs are compared to determine if C is greater than $C_{0.90}$; when this is the case, the end tail distribution is used to re-generate a simulated cost.

After verifying that the simulation results fit the historical data, respective extreme weather event trends can be evaluated. The model can then be applied to future extreme weather event scenarios according to anticipated changes in event frequency/severity and demographics. These results can be compared with the status quo (base case) scenario to estimate future L&D with or without the introduction of risk mitigation strategies.

4 Model Validation

The aforementioned methodology was applied as separate simulations for 201 unique region and extreme weather event type combinations, each run 1000 times over a 20-year period. The results from each individual simulation were first compared to normalized historical data in order to understand how well they represented historical observations. This was accomplished by calculating the percentage difference via the following equation for each of the 1000 runs:

$$\% \text{mean difference} = \frac{[(\text{Historical Mean}) - (\text{Simulated Mean})]}{\text{Historical Mean}} \quad (3)$$

The average p value of the historical data following the Kolmogorov–Smirnov test of these 201 combinations was 0.680 with a variance of 0.088. The distribution fits were good considering the variety of distributions presented with the historical data.

By using the mean value, both the individual trial performance and the overall simulation performance can be examined since the mean is both a function of the total events that take place and a measure of the center of the data. Overall, 187 of the 201 simulations showed a percent mean difference below 10%. Despite a few outliers that may be addressed in further testing or re-fitting distributions, generally the simulation was able to reasonably reproduce the L&D associated with historical data representing extreme event types occurring in the areas tested.

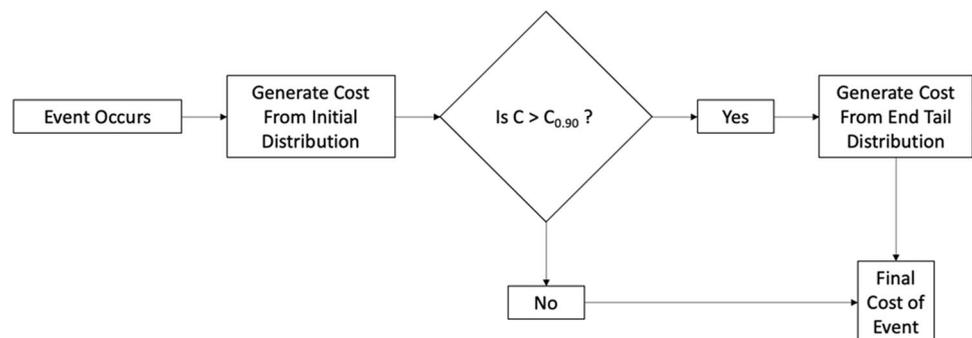
5 Discussion

The efficacy of this hybrid simulation approach, as expected, relies on a sufficient sample size of historical data, which is the primary input for the model. The average percent mean difference of the combined 201 simulations was 5%. Slightly better results were achieved when the input sample size exceeded 1000 historical events. This does not mean the simulation cannot be useful when applied to sample sizes below 1000 historical events, only that it may somewhat reduce confidence in the results.

As seen in Fig. 5, the number of historical events is not the only variable impacting model performance. The variance represents the overall range of possible L&D that is incurred in an event. As the variance increases, there is less consistency in the event outcomes, making the ultimate modeling and data fitting process more difficult. Aside from a few clear outliers, as the overall historical cost variance increases, the percent mean difference also increases.

It can be seen that the variability in the historical data propagates into the model and corresponding cost outputs.

Fig. 4 Generation of damage cost for a specific hazard type. C is the realized cost in the simulation, and if C is greater than the 90th percentile threshold it is replaced with a cost generated from the end tail distribution.



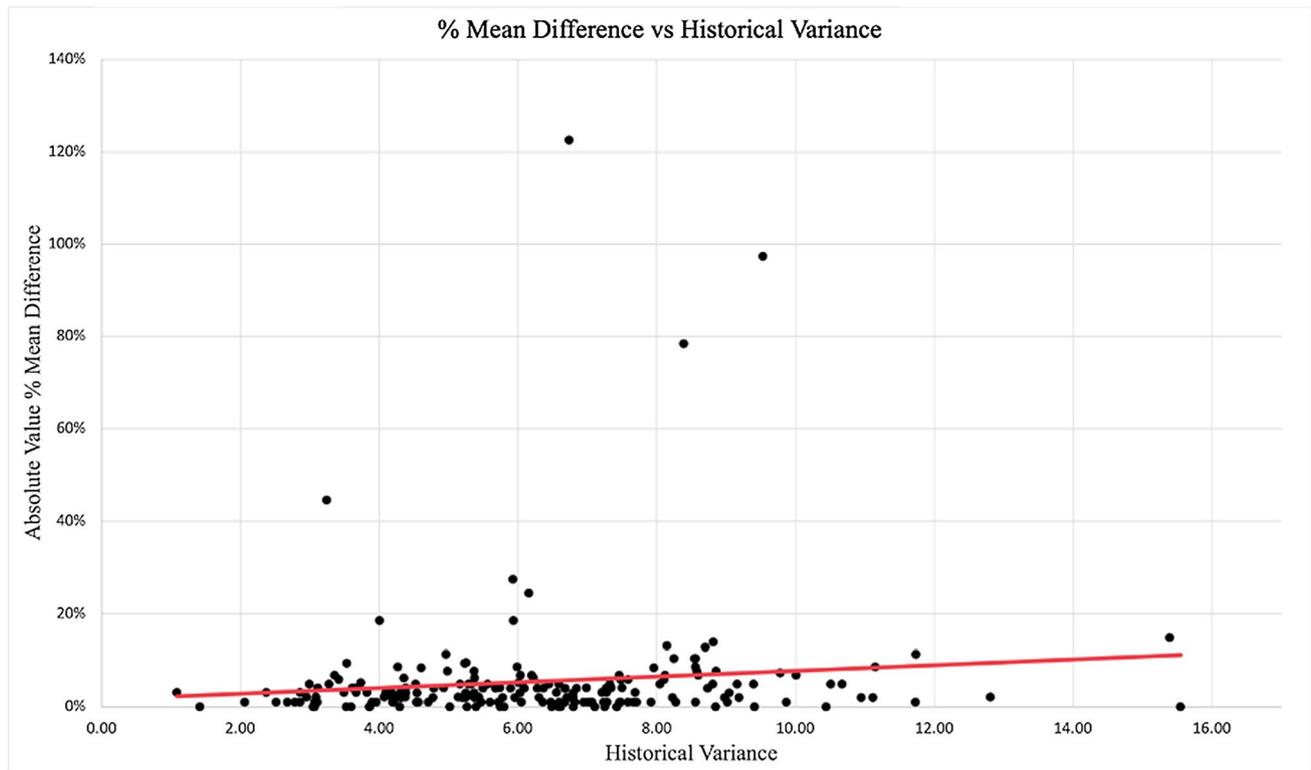


Fig. 5 Percent mean difference as a function of the variance in historical cost of an event. The red line shows the trend.

This relationship is further explained in Table 1, where the extreme weather types with the most events and smallest variability lead to the smallest percent mean differences.

This is even more evident when examining the empirical distributions of the historical L&D data. For example, when comparing the historical data of convective storms in the Midwest and tropical cyclones in Texas, the percentage mean differences are 1% and 15%, respectively. Midwest convective storms had roughly 22,000 events and a historical cost variance of 5.60 compared to Texas tropical storms that consisted of 120 events and a historical cost variance of 15.39. These two combinations were selected to highlight the range of differences between region and

state sample sizes in the historical data and to show inherent differences in the range of possible L&D for different event types. Although tropical cyclones have a much larger range of outcomes compared to convective storms, the sample size for Texas tropical storms is so much smaller than for Midwest convective storms, creating greater difficulty in confident data modeling.

However, even though the model is not as close to the mean in cases of smaller sample sizes, it does not necessarily imply that the results are not useful, only that the estimated results are less precise. This is a common phenomenon in any extreme weather analysis where sparse data exist for low probability, high consequence events.

Table 1 Model performance comparison by event type. The highlighted rows include the smallest average mean difference or the smallest variance.

	Total events	Average mean difference (%)	Historical variance of event costs (log transformed)	No. of test samples
Convective storm	393,598	5	4.95	57
Flooding	83,167	6	7.10	56
Tornado	26,423	4	6.34	33
Tropical cyclone	3759	13	7.95	13
Wildfire	853	6	6.89	4
Winter convective storm	29,191	4	5.87	36

6 Conclusion

By developing a better understanding of L&D associated with different extreme weather event types in various climate regions, it provides a basis for addressing what to expect in terms of L&D based on future climate projections. The model and methodology described herein can help guide decision making in future infrastructure adaptation investments by creating a cost comparison of inaction versus implementation of candidate risk mitigation strategies. It is based on the premise that practitioners desire to make more risk-informed investment decisions using a simple, practical framework given limited time and resources.

As previously discussed, output comparison of the end tail values, specifically the maximum cost events in the simulated period, presents a unique challenge due to the limited historical data for extreme weather events. This is further amplified if an extreme event is not present in the historical data. This issue can be potentially addressed by using an expanded data set, perhaps with a sample size threshold, to act as a benchmark for the normalized maximum costs.

Furthermore, one must be mindful that L&D model estimation is presently limited to only property damage effects of extreme weather events; intangible and indirect damages must be quantified to develop a more complete picture of the L&D impacts of an event on an area. When these empirical data become available, they can be incorporated into the approach espoused.

This methodology is unique in its approach to modeling extreme weather loss and damage. Through development of this methodology, L&D data can be used as another tool in adaptation decision making while limiting the data demands of most hazard cost projection and analysis. The added benefit of this approach is that it can become an even more impactful resource over time as the databases gain more entries from the occurrence of future events.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Blackwell, C. 2014. Power law or lognormal? Distribution of normalized hurricane damages in the United States, 1900–2005. *Natural Hazards Review* 16(3). [https://doi.org/10.1061/\(ASCE\)NH.1527-6996.0000162](https://doi.org/10.1061/(ASCE)NH.1527-6996.0000162).
- Botzen, W.J.W., and J.C.J.M. Van Den Bergh. 2009. Managing natural disaster risks in a changing climate. *Environmental Hazards* 8(3): 209–225.
- Botzen, W.J.W., O. Deschenes, and M. Sanders. 2020. The economic impacts of natural disasters: A review of models and empirical studies. *Review of Environmental Economics and Policy* 13(2): 167–188.
- Choi, D., D.O. Kasdan, and D.K. Yoon. 2019. Analyzing disaster loss trends: A comparison of normalization methodologies in South Korea. *Risk Analysis* 39(4): 859–870.
- Cutter, S., B. Boruff, and L. Shirley. 2003. Social vulnerability to environmental hazards. *Social Science Quarterly* 84(2): 242–261.
- Doktycz, C., and M. Abkowitz. 2019. Loss and damage estimation for extreme weather events: State of the practice. *Sustainability* 11(15): Article 4243.
- Flanagan, B.E., E.W. Gregory, E.J. Hallisey, J.L. Heitgerd, and B. Lewis. 2011. A social vulnerability index for disaster management. *Journal of Homeland Security and Emergency Management* 8(1): Article 3.
- French, J., P. Kokoszka, S. Stoev, and L. Hall. 2018. Quantifying the risk of heat waves using extreme value theory and spatio-temporal functional data. *Computational Statistics and Data Analysis* 131: 176–193.
- Integrated Research on Disaster Risk. 2014. *Peril classification and hazard glossary*. IRDR data publication No. 1. Beijing: Integrated Research on Disaster Risk.
- Kahn, M.E. 2005. The death toll from natural disasters: The role of income, geography, and institutions. *Review of Economics and Statistics* 87(2): 271–284.
- Karl, T.R., and W.J. Koss. 1984. *Regional and national monthly, seasonal, and annual temperatures weighted by area, 1895–1983*. Historical Climatology Series 4-3. Asheville, NC: National Climatic Data Center.
- Katz, R.W. 2020. Statistical issues in detection of trends in losses from extreme weather and climate events. In *Evaluating climate change impacts*, ed. V. Lyubchich, Y.R. Gel, K.H. Kilbourne, T.J. Miller, N.K. Newlands, and A.B. Smith, 165–186. London: Chapman and Hall/CRC.
- Kunruether, H. 2008. Reducing losses from catastrophic risks through long term insurance and mitigation. *Social Research: An International Quarterly* 75(3): 905–930.
- Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, et al. 2021. *Climate change 2021: The physical science basis. Contribution of Working Group I to the sixth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press.
- McNeil, A.J. 1997. Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bulletin: The Journal of the IAA* 27(1): 117–137.
- Nordhaus, W. 2010. The economics of hurricanes and implications of global warming. *Climate Change Economics* 1(1): 1–20.
- Shen, G., L. Zhou, Y. Wu, and Z. Cai. 2018. A global expected risk analysis of fatalities, injuries, and damages by natural disasters. *Sustainability* 10(7): Article 2573.
- Tarling, H.A. 2017. Comparative analysis of social vulnerability indices: CDC's SVI and SoVI®. Masters thesis, Lund University, Lund, Sweden. <http://lup.lub.lu.se/student-papers/record/8928519>. Accessed 16 May 2022.
- Weinkle, J., C. Landsea, D. Collins, R. Musulin, R.P. Crompton, P.J. Klotzbach, and R. Pielke Jr. 2018. Normalized hurricane damage in the continental United States 1900–2017. *Nature Sustainability* 1: 808–813.

Wood, E., M. Sanders, and T. Frazier. 2021. The practical use of social vulnerability indicators in disaster management. *International Journal of Disaster Risk Reduction* 63: Article 102464.

Zimbidis, A., N. Frangos, and A. Pantelous. 2007. Modeling earthquake risk via extreme value theory and pricing the respective

catastrophe bonds. *ASTIN Bulletin: The Journal of the IAA* 37(1): 163–183.