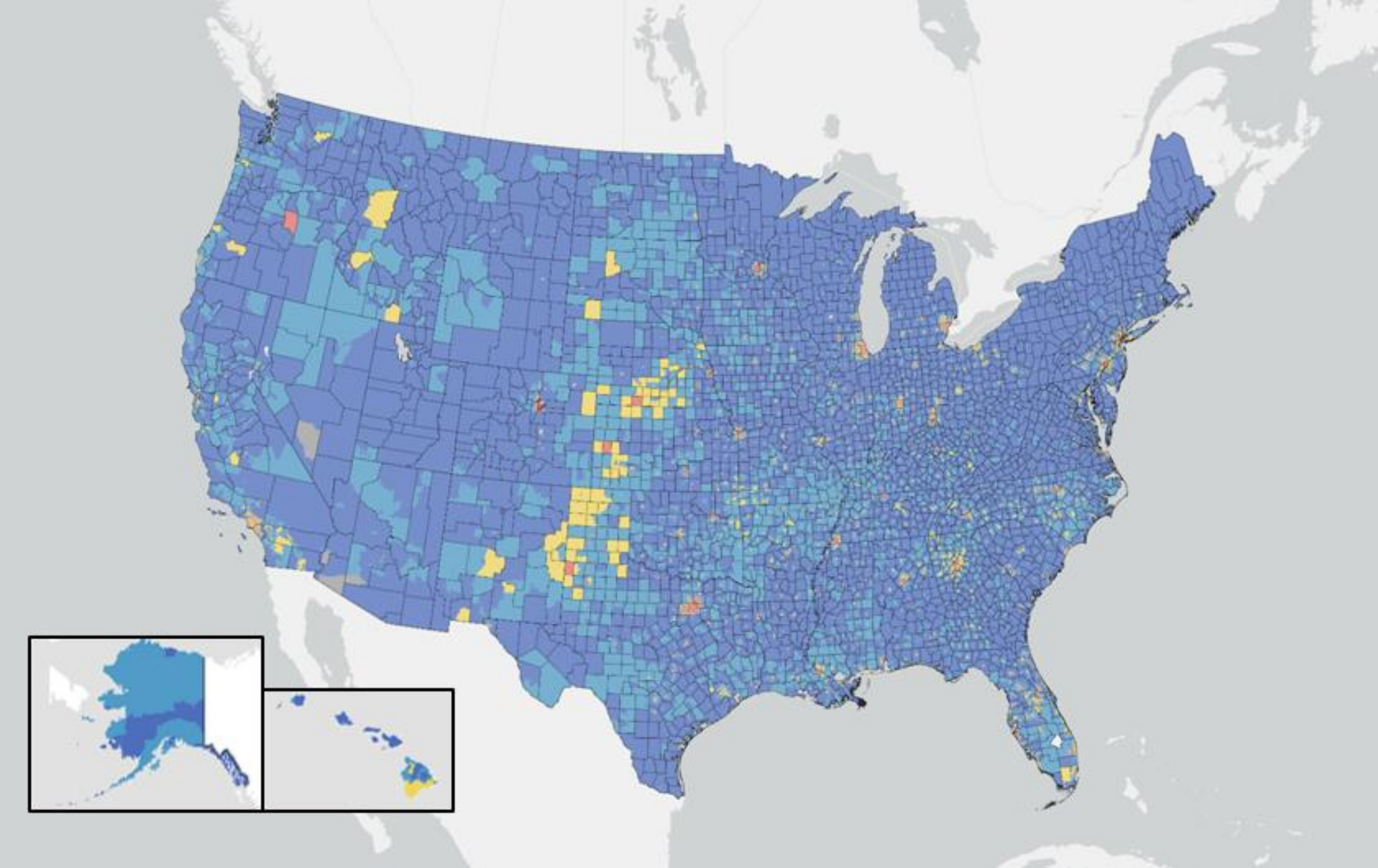

National Risk Index

Primer

December 2020

FEMA

This page intentionally left blank

Table of Contents

1. Introduction

The National Risk Index (NRI) is a dataset and an application that help identify communities most at-risk for natural hazards. The NRI leverages available source data for 18 natural hazards, social vulnerability, and community resilience to develop a baseline relative risk measurement for each United States county and Census tract. The NRI is intended to help users better understand the natural hazard risk of their communities or assigned areas. Intended users include planners and emergency managers at the local, regional, state, and federal levels, as well as other decision makers and interested members of the general public. Specifically, it can support decision-making to:

- Update emergency operations plans
- Enhance hazard mitigation plans
- Prioritize and allocate resources
- Identify the need for more refined risk assessments
- Encourage community-level risk communication and engagement
- Educate homeowners and renters
- Support enhanced codes and standards
- Inform long-term community recovery

This report provides a detailed overview of the National Risk Index, including its background, data sources, and processing methodologies. It describes the high-level concepts used to develop the NRI and calculate its components. The methodologies for computing each hazard's Expected Annual Loss (EAL) are also explained in depth in the NRI Technical Documentation.

2. Background

All communities in the United States experience natural hazards, and there is a wide range of environmental, social, and economic factors that influence each community's risk to natural hazards. The likelihood that a community may experience a natural hazard can vary drastically, as can the associated consequences. Additionally, a community's risk is influenced by many social, economic, and ecological factors. FEMA, along with numerous federal, state, and local governments, academic institutions, nonprofit groups, and private industry (see Figure 1), collaborated to develop the National Risk Index as a baseline risk assessment application.

Beginning in 2016, FEMA's Natural Hazards Risk Assessment Program (NHRAP) started work on the NRI by adopting an established vision for a multi-hazard view of risk that combines the likelihood and consequence of natural hazards with social factors and resilience capabilities. The goal was to take a broad, holistic view and create a nationwide baseline of natural hazard risk. Through various partnerships and working groups, FEMA developed a methodology and procedure to create the NRI dataset, and then researched, designed, and built the NRI website and application.

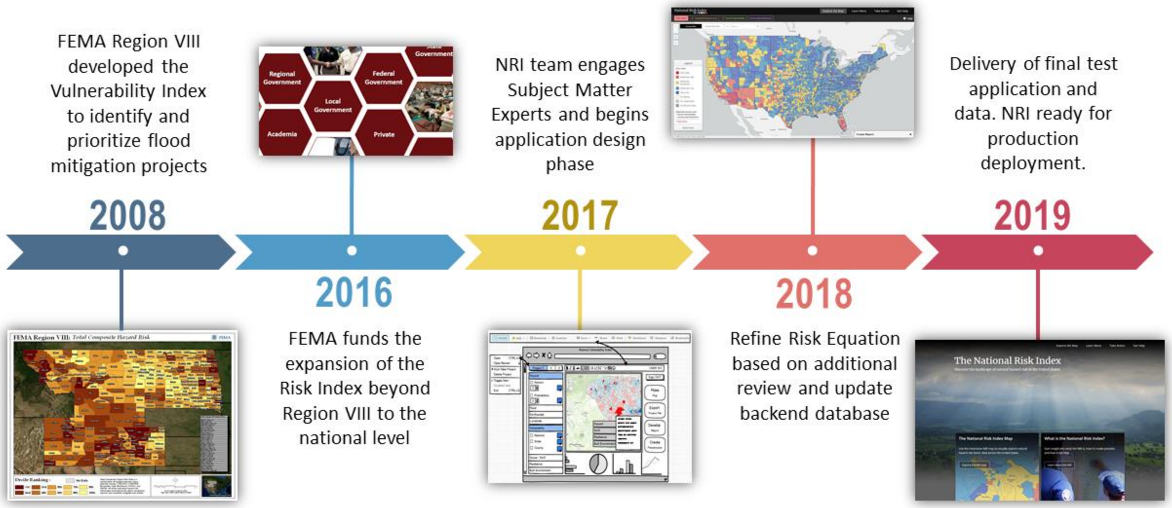

Figure 1: Timeline of the Development of the National Risk Index

The NRI Team conducted multiple workshops and sessions to discuss and determine the methodologies for translating raw source data into natural hazard risk factors for input into the NRI. The key objective of these exercises was to ensure that a vetted risk model or equation was leveraged throughout all methodological development and that certain factors were not being interpreted inconsistently across the 18 natural hazards.

2.1. Natural Hazard Selection

Natural hazard exposure across the country varies from location to location. The 18 natural hazards evaluated by the NRI were chosen after reviewing FEMA-approved State Hazard Mitigation Plans for all 50 states. Tribal hazard mitigation plans were not available at the time of the analysis, and island territories were excluded from the hazard selection process since data for most NRI hazards are not available. Note that Washington, DC, was initially excluded from the hazard selection analysis process; however, it was added to the project scope in 2017 after the hazard selection.

Natural hazards that were included in at least half of the FEMA-approved state plans, or those that were deemed to be of regional significance, were selected to the NRI (see Figure 2). A regionally significant hazard is defined as having the capacity to cause widespread, catastrophic damage, such as Hurricanes, Tsunami, and Volcanic Activity, but otherwise affected fewer than 25 states. It should be noted that one natural hazard, Subsidence, fit these criteria, but could not be evaluated by the NRI as there was no reliable, nationwide dataset cataloging this type of hazard event.

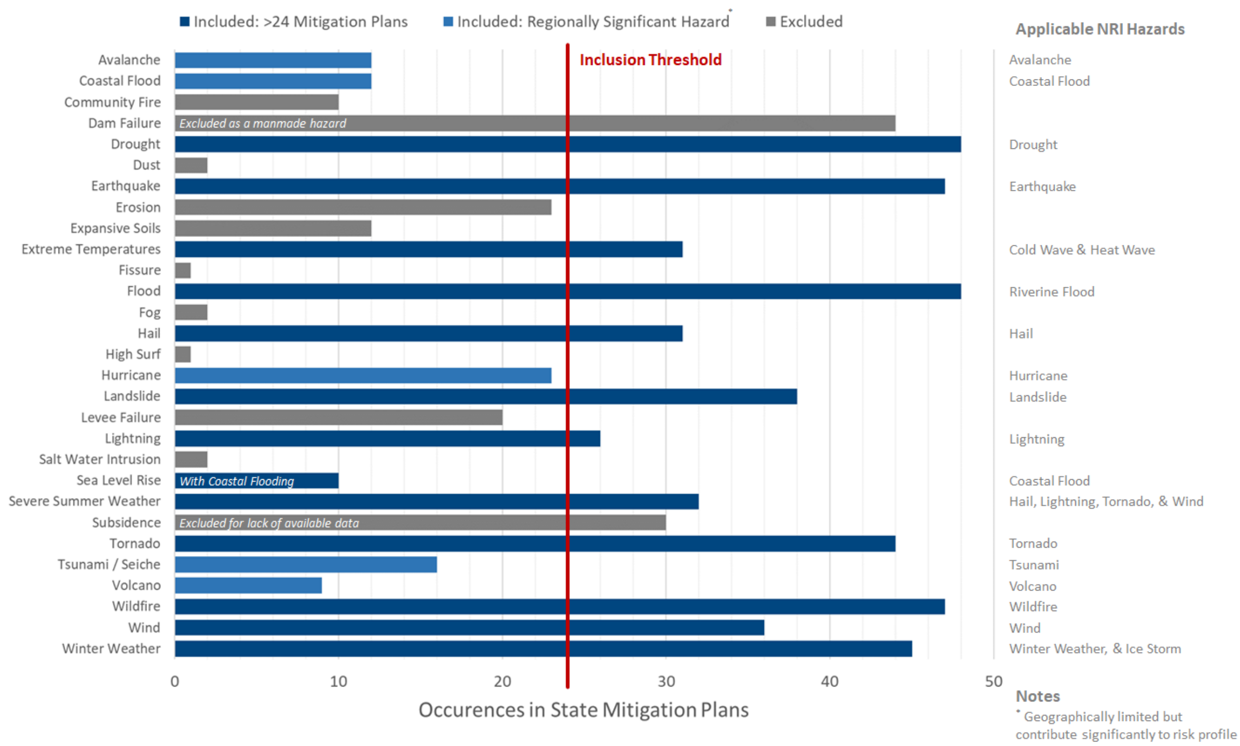

Figure 2: Determination of Hazard Inclusion Based on State Hazard Mitigation Plans

The State Hazard Mitigation Plan review revealed that both Dam Failure and Levee Failure hazards are profiled by many states, but the datasets needed to develop the EAL component of NRI are not nationally or publicly available. Levees may be incorporated into the riverine or coastal flood components if these manmade features are not included on floodplain maps or reflected in NOAA storm surge and coastal flood analysis. These hazards should not be discussed from traditional risk assessment. The State Hazard Mitigation Plan hazard analysis was completed in early 2016 and was limited to the FEMA-approved State Hazard Mitigation Plans. No territorial or tribal plans were reviewed due to their limited availability.

2.2. Working Groups

After a detailed literature review and hazard analysis, the NRI Team convened three working groups made of intended users, subject matter experts (SMEs), and interested stakeholders from all levels of government, private industry, nonprofits, and academia. Each working group was responsible for an aspect of the NRI's development and methodology. Experts in each group helped guide the NRI data and application development.

The Natural Hazards Working Group assessed and recommended datasets associated with the identified 18 natural hazards selected (as well as Subsidence prior to its recommended removal) and determined the best ways to incorporate associated data into the NRI.

The Social Vulnerability and Community Resilience Working Group reviewed and evaluated existing efforts to measure social vulnerability and community resilience to understand which components were most important (vulnerability or resilience, or both) and which indices should be used in the NRI. As a result, both Social Vulnerability and Community Resilience are components of the NRI.

The Data Analytics Working Group oversaw the spatial processing, normalization, and aggregation of data to arrive at a risk indexing methodology and calculation procedure that integrated the datasets identified by the other two working groups.

Together, the groups discussed and developed the National Risk Index, including the datasets and indices to incorporate, definitions of index components, data management strategies and metadata requirements, data processing and index creation methodologies, and the data visualization and interactive web mapping application requirements.

2.3. Literature Review

The NRI's project team reviewed literature in the fields of hazard mitigation, emergency management, hazard risk science, and other related fields. Centering around a search for natural hazard and exposure variables, the literature review identified multiple datasets, risk indices, research reports, methodologies, indicator lists, and existing risk assessment at national and global scales.

The team identified important risk indicator categories and specific indicators during the review (see Table 1).

Table 1: Literature Review Risk Indicators and Categories

Risk Indicator Categories	Individual Risk Indicators	
▪ Social	▪ Income	▪ Road Systems
▪ Economic	▪ Age	▪ Economic Productivity
▪ Environmental	▪ Illnesses	▪ Housing
▪ Infrastructure	▪ Hospitals	▪ Community Revenue

After review, the team concluded the NRI would involve three components: natural hazard risk (likelihoods and consequences), social vulnerability, and community resilience.

2.4. Subject Matter Expert Review

Extensive development of the NRI began in 2017 and proceeded through the end of 2019. Over this period, the NRI team continually iterated on their data processing and risk calculation methodologies, and engaged with SMEs throughout. A full list of organizations whose members contributed to the SME reviews is available in the NRI Technical Documentation.

At major milestones, the team paused development to engage in broader, more comprehensive SME review periods. The first major milestone arrived in January 2019 where teams of SMEs were tasked to evaluate two competing draft methodologies: “Methodology 1,” which relied on unitless standardization of EAL, and “Methodology 2,” which standardized EAL to a dollar value measurement. Over the course of two weeks and many meetings, dozens of SMEs provided feedback to the NRI team, resulting in a clear consensus that, although both methodologies were valid, Methodology 2 created a more robust measurement of risk and a more valuable dataset for the hazard planning and mitigation communities.

With clear direction on the methodology, the NRI team continued iterating through improvements to data sourcing and processing. From July through September 2019, they conducted a final comprehensive SME review period to focus on the new methodology’s results. More than 40 SMEs participated in over 20 review sessions and helped the team reach concurrence on the validity and value of the dataset. From these sessions, the NRI team was equipped to begin final iterations of the methodology and source data processing.

2.5. Data and Methodologies

Over the course of several years, with the help of hundreds of collaborators and contributors, and through unknown iterations of planning, design and development, the NRI working groups concluded their work by reviewing and providing feedback on an iterative version of the National Risk Index dataset (December 2019).

Briefly stated, the NRI is a first-of-its-kind, nationwide, holistic assessment of baseline risk to natural hazards. Although it is based on extensive research and best practices in the risk assessment fields, the NRI’s methodology is unique and carefully constructed the specific needs of natural hazard risk assessment at both small and large geographic scales. A detailed overview of the risk calculation is available in the Risk Analysis Overview section.

The NRI’s most important and central component, Expected Annual Loss (EAL), is a robust measurement that quantifies the anticipated economic damage resulting from natural hazards each year. Details of its equation and analytical techniques are available in the Expected Annual Loss section. EAL consists of the best-available datasets for 18 natural hazards of national and regional significance, with source data being processed to match the unique nature of each natural hazard. Full processing details for each hazard are available in the NRI Technical Documentation. Per the direction established at initiation, the dataset also includes measurements of social vulnerability and community resilience to quantify overall risk. These key components are detailed fully in the Social Vulnerability and Community Resilience sections.

3. Risk Analysis Overview

Risk, in the most general terms, is often defined as the likelihood (or probability) of a natural hazard event happening multiplied by the expected consequence if a natural hazard event occurs. The generalized form of a risk equation is given in Equation 1.

Equation 1: Generalized Risk Equation

$$Risk = Likelihood\ x\ Consequence$$

3.1. Risk Calculation

In the National Risk Index, risk is defined as the potential for negative impacts as a result of a natural hazard. The risk equation behind the NRI includes three components: a natural hazards component, a consequence enhancing component, and a consequence reduction component. EAL is the natural hazards risk component, measuring the expected loss of building value, population, and/or agricultural value each year due to natural hazards. Social vulnerability is the consequence enhancing component and analyzes demographic characteristics to measure a community's susceptibility of social groups to the adverse impacts of natural hazards. Community resilience is the consequence reduction component and uses demographic characteristics to measure a community's ability to prepare for, adapt to, withstand, and recover from the effects of natural hazards. These three risk components are combined into one risk value using Equation 2.

Equation 2: NRI Risk Equation

$$Risk = Expected\ Annual\ Loss \times Social\ Vulnerability \times \frac{1}{Community\ Resilience}$$

An overall composite Risk Index score and individual hazard Risk Index scores are calculated for each county and Census tract included in the NRI. A composite Risk Index score measures the relative risk of a location considering all 18 natural hazards included in the index. An individual hazard Risk Index score measures the relative natural hazard risk of a location for a single natural hazard. All scores are relative as each Census tract or county's score is evaluated in comparison with all other Census tracts or counties.

3.2. Scores and Ratings

In this NRI Risk Equation, each component is represented by a unitless index value, representing a community's score relative to all other communities. From the three indices, the Risk Index score is calculated to measure a community's risk to all 18 natural hazards. The Risk Index is also a unitless index and represents a community's risk relative to all other communities. The Risk Index and EAL are provided as both composite scores from the summation of all 18 natural hazards, as well as individual-hazard scores where each hazard is considered separately.

All calculations are performed separately at two levels-of-detail—county and Census tract—so scores are relative only within their level-of-detail. It must be stressed that scores are relative, representing a community's relative position among all other communities for a given component and level-of-detail. Scores are not absolute measurements and should be expected to change over time either by their own changing measurements or changes in other communities.

All scores are constrained to a range of 0 (lowest possible value) to 100 (highest possible value). To achieve this range, the values of each component are rescaled using min-max normalization, which preserves their distribution while making them easier to understand. EAL values are heavily skewed by an extreme range of population and property value densities between urban and rural communities. To account for this, a cube root transformation is applied before min-max normalization. By applying cube root transformation, the NRI controls for this characteristic and provides scores with greater differentiation and usefulness. If the minimum value of the EAL is a nonzero number before normalization, an artificial minimum is set to 99% of that value, so that entities expected to experience loss do not receive a 0 EAL score.

For every score there is also a qualitative rating that describes the nature of a community's score in comparison to all other communities, ranging from "Very Low" to "Very High." Because all ratings are relative, there are no specific numeric values that determine the rating. For example, a community's Risk Index score could be 8.9 with a rating of "Relatively Low," but its Social Vulnerability score may be 11.3 with a rating of "Very Low." The rating is intended to classify a community for a specific component in relation to all other communities.

To determine ratings, a methodology known as k-means clustering or natural breaks is applied to each score. This approach divides all communities into groups such that the communities within each group are as similar as possible (minimized variance) while the groups are as different as possible (maximized variance).

In the NRI application's maps and data visualizations, standard color schemes have been applied to the qualitative ratings. Risk Index ratings are represented using a diverging blue (Very Low) to red (Very High) color scheme. Ratings for EAL, Social Vulnerability and Community Resilience are represented using sequential color schemes (e.g., single color at various intensities). According to the NRI, higher EAL, higher Social Vulnerability, and/or lower Community Resilience increase your overall risk. In general, darker shading in the map layers represents a higher contribution to overall risk. When source data is not available or a score cannot be calculated, then additional ratings are used and shown in white or shades of gray. The NRI's standard color schemes are shown in Figure 3.

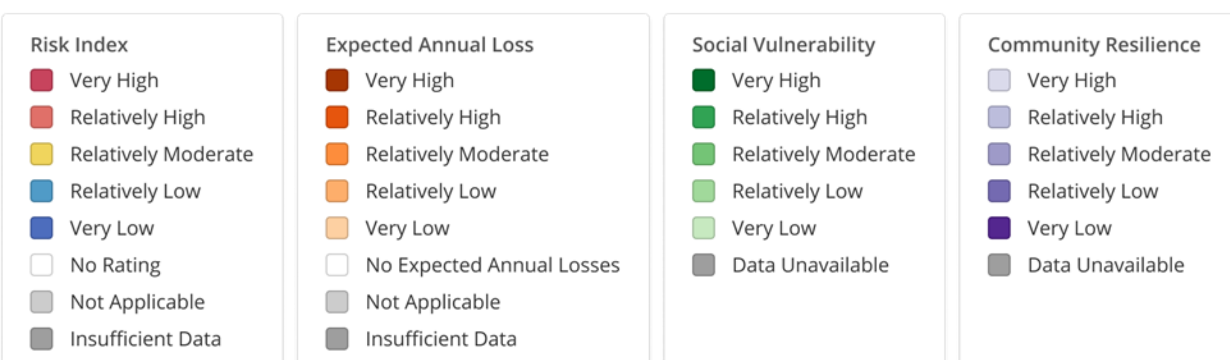

Figure 3: National Risk Index Qualitative Rating Legend

Scores of 0 (zero) or missing values ("nulls") in the EAL components receive ratings that reflect the logic behind the score. A county or tract whose EAL is zero either has no building value, population, or crop value exposed to the hazard, or has a calculated hazard frequency of zero, except for hazards that apply a minimum annual frequency. These areas are displayed in the NRI application as having "No Expected Annual Loss" for the designated hazard.

In collaboration with SMEs most familiar with individual hazards and the source data used in the NRI, a priori definitions of hazard applicability have also been applied to help distinguish between where no hazard risk exists and where the hazard is deemed to be not possible. For example, Coastal Flooding EAL is not computed for inland areas. These areas are displayed in the NRI application as "Not Applicable" for EAL computation for the designated hazard.

Finally, if a component used to calculate the EAL of a Census tract or county for a hazard has a null value, the community is rated as "Insufficient Data." For example, certain hazards, such as Wildfire, Lightning, and Landslide, only have source data used to determine frequency or exposure for the conterminous United States, meaning that both Alaska and Hawaii are rated as "Insufficient Data" to compute the EAL for those hazards. When a hazard is not applicable or there is insufficient data for a community, EAL for that hazard is simply not included in the community's final summation and scoring. A summary of non-numerical ratings is provided Table 2.

Table 2: Definitions of Ratings without Numerical Scores

Rating	Risk Index	Expected Annual Loss	Social Vulnerability	Community Resilience
No Rating	EAL is zero. SoVI and/or HVRI BRIC are not available.	n/a	n/a	n/a
No Expected Annual Loss	n/a	Hazard exposure or frequency is zero.	n/a	n/a
Not Applicable	Location is not considered at-risk for hazard occurrence.	Location is not considered at-risk for hazard occurrence.	n/a	n/a
Insufficient Data	Hazard source data is not available.	Hazard source data is not available.	n/a	n/a
Data Unavailable	n/a	n/a	SoVI is not available.	HVRI BRIC is not available.

3.3. Assumptions and Limitations

The National Risk Index dataset and application are meant for planning purposes only and are intended for use as a tool for broad, nationwide comparisons. Nationwide datasets used as inputs for the NRI are in many cases not as accurate as locally available data. Users with access to local data for each NRI risk factor should consider substituting those data to calculate a more accurate EAL value at the local level.

The NRI does not consider the intricate economic and physical interdependencies that exist across geographic regions. The user should be mindful that hazard impacts in surrounding counties or Census tracts can cause indirect losses in a location regardless of the location's risk profile.

The NRI's most recent source datasets only include a period of record up to 2017. It should be noted that the EAL values represent an extrapolation based on a snapshot in time. Extending source data collection beyond that time may result in varying Census tract or county EAL values due to changes in recorded hazard intensity and frequency, as well as fluctuations in local economic value and/or population density.

Most of the hazards evaluated by the NRI use a frequency model to determine EAL. This makes it difficult to accurately estimate EAL for high consequence, low frequency events. Certain rare hazards (such as Earthquake, Hurricane, Tsunami, and Volcanic Activity) benefit from using a probabilistic model that estimates the likelihood of a hazard event occurring over an extended period of time, which can then be annualized. Of these, only Earthquake has probabilistic source data that is sufficient for accurately estimating EAL.[1]

Best available nationwide data for some risk factors are rudimentary. More sophisticated risk analysis methodologies are available but require more temporally and spatially granular data for hazard exposure, frequency, and historic loss measurements.

The NRI methodology makes various efforts to control for possible discrepancies in source data, but cannot correct for all accuracy problems present in that data. The NRI processing database is a complex system and localized inaccuracies in source data have the potential to propagate. Therefore, the NRI and its components should be considered a baseline measurement and a guideline for determining hazard risk but should not be used as an absolute measurement of risk.

4. Risk Components Overview

The risk score in the NRI is based on three components: Social Vulnerability, Community Resilience, and EAL, with EAL based on Exposure, Annualized Frequency, and Historic Loss components, for a total of five risk factors. Each risk factor contributes to either the likelihood or consequence aspect of risk and can be classified as one of two risk types: either risk based on geographic location or risk

[1] Federal Emergency Management Administration (FEMA). (2017). *Hazus Estimated Annualized Earthquake Losses for the United States: FEMA Publication 366*. Retrieved from https://www.fema.gov/sites/default/files/2020-07/fema_earthquakes_hazus-estimated-annualized-earthquake-losses-for-the-united-states_20170401.pdf

based on the nature and historical occurrences of natural hazards. The five risk factors are summarized in Table 3 and further described in this section.

Table 3: National Risk Index Score Risk Factors

NRI Risk Component	NRI Risk Factors	Risk Factor Description	Risk Equation Bin	Risk Type Assignment
Social Vulnerability	Social Vulnerability	Consequence Enhancer	Consequence	Geographic Risk
Community Resilience	Community Resilience	Consequence Reducer	Consequence	Geographic Risk
Expected Annual Loss	Exposure	Expected Consequence	Consequence	Natural Hazard Risk
Expected Annual Loss	Annualized Frequency	Probability of Occurrence	Likelihood	Natural Hazard Risk
Expected Annual Loss	Historic Loss	Expected Consequence	Consequence	Natural Hazard Risk

4.1. Social Vulnerability

Social vulnerability is broadly defined as the susceptibility of social groups to the adverse impacts of natural hazards, including disproportionate death, injury, loss, or disruption of livelihood. Social vulnerability considers the social, economic, demographic, and housing characteristics of a community that influence its ability to prepare for, respond to, cope with, recover from, and adapt to environmental hazards.

As a consequence-enhancing risk factor, the Social Vulnerability score represents the relative level of social vulnerability for a given county or Census tract. A higher social vulnerability score results in a higher risk score. Because social vulnerability is unique to a geographic location—specifically, a county or Census tract—it is a geographic risk factor.

The Social Vulnerability and Community Resilience Working Group reviewed multiple top-down and bottom-up indices and chose to recommend the University of South Carolina's Hazards and Vulnerability Research Institute (HVRI) Social Vulnerability Index (SoVI).

4.1.1. SOCIAL VULNERABILITY SOURCE DATA

Social Vulnerability source data provider: University of South Carolina's Hazards and Vulnerability Research Institute (HVRI) Social Vulnerability Index (SoVI)

SoVI is a location-specific assessment of social vulnerability that utilizes 29 socioeconomic variables (listed below) deemed to contribute to a community's reduced ability to prepare for, respond to, and recover from hazards.[2]

- Median gross rent for renter-occupied housing units
- Median age
- Median dollar value of owner-occupied housing units
- Per capita income
- Average number of people per household
- % population under 5 years or age 65 and over
- % civilian labor force unemployed
- % population over 25 with <12 years of education
- % children living in married couple families
- % female
- % female participation in the labor force
- % households receiving Social Security benefits
- % unoccupied housing units
- % families with female-headed households with no spouse present
- % population speaking English as second language (with limited English proficiency)
- % Asian population
- % African American (Black) population
- % Hispanic population
- % population living in mobile homes
- % Native American population
- % housing units with no car available
- % population living in nursing facilities
- % persons living in poverty
- % renter-occupied housing units
- % families earning more than $200,000 income per year
- % employment in service occupations
- % employment in extractive industries (e.g., farming)
- % population without health insurance (County SoVI only)
- Community hospitals per capita (County SoVI only)

Data was acquired from HVRI's SoVI website and users looking for more information should consult HVRI.

4.1.2. PROCESSING SOCIAL VULNERABILITY SOURCE DATA FOR THE NRI

For the NRI, the SoVI dataset was incorporated using min-max transformation (0.01-100.00 scale). County-level and Census tract-level Social Vulnerability scores were classified into five qualitative categories, from "Very Low" to "Very High," using k-means clustering. Social Vulnerability scores are available for all counties, but they are absent for 292 Census tracts that have no population. Risk cannot be calculated for tracts without Social Vulnerability scores, so those Census tracts are rated "Insufficient Data."

[2] Cutter, S.L., Boruff, B.J. & Shirley, W.L. (2003). Social vulnerability to environmental hazards. Social Science Quarterly, 84(2): 242-261. Retrieved from https://doi.org/10.1111/1540-6237.8402002

4.2. Community Resilience

Community Resilience is defined by FEMA as the ability of a community to prepare for anticipated natural hazards, adapt to changing conditions, and withstand and recover rapidly from disruptions.[3]

There are multiple, well-established ways to define community resilience at the local level, and key drivers of resilience vary between locations. Because there are no nationally available, bottom-up community resilience indices available, the Social Vulnerability and Community Resilience Working Group chose to utilize a top-down approach. The NRI relies on using broad factors to define resilience at a national level and create a comparative metric to use as a risk factor. The Social Vulnerability and Community Resilience Working Group reviewed multiple top-down indices and chose to recommend the University of South Carolina's Hazards and Vulnerability Research Institute (HVRI) Baseline Resilience Indicators for Communities (HVRI BRIC) index.

The Community Resilience score is a consequence reduction risk factor of the NRI and represents the relative level of community resilience for a given location. A higher Community Resilience score results in a lower Risk score. Because Community Resilience is unique to a geographic location—specifically, a county—it is a geographic risk factor.

4.2.1. COMMUNITY RESILIENCE SOURCE DATA

Community Resilience source data provider: University of South Carolina's Hazards and Vulnerability Research Institute (HVRI) Baseline Resilience Indicators for Communities (BRIC)

Community Resilience data for the NRI is supported by the HVRI BRIC. HVRI BRIC provides a sound methodology for quantifying community resilience by identifying the ability of a community to prepare and plan for, absorb, recover from, and more successfully adapt to the impacts of natural hazards. The HVRI BRIC dataset includes a set of 49 indicators that represent six types of resilience: social, economic, community capital, institutional capacity, housing/infrastructure, and environmental. It uses a local scale within a nationwide scope, and the national dataset serves as a baseline for measuring relative resilience. This data can be used to compare one place to another and determine specific drivers of resilience, and a higher HVRI BRIC score indicates a stronger and more resilient community.

4.2.2. PROCESSING COMMUNITY RESILIENCE SOURCE DATA FOR THE NRI

For the NRI, the HVRI BRIC dataset was in incorporated using min-max transformation (0.01-100.00 scale). Because HVRI BRIC has a potential range of 0.0 to 6.0, but the full range does not exist in the dataset, the normalized score for Community Resilience ranges from 41.2 to 64.7. HVRI BRIC is only available at the county-level, so Community Resilience scores were inferred from counties to Census tracts by assigning each Census tract the value of its parent county. Community Resilience scores

[3] National Institute of Standards and Technology (NIST). (2020). *Community Resilience*. Retrieved from: https://www.nist.gov/topics/community-resilience

were classified into five qualitative categories, from "Very Low" to "Very High," using k-means clustering.

For more information on the creation of the HVRI BRIC, please refer to HVRI's BRIC website or the geographies of community disaster resilience paper published by Cutter, Ash, and Emrich (2014). [4,5]

4.3. Expected Annual Loss (EAL)

The EAL for each Census tract or county is the average economic loss in dollars resulting from natural hazards each year. EAL is computed for each hazard type and only quantifies loss for relevant consequence types (i.e., buildings, people, or agriculture). For example, most natural hazards only significantly impact buildings and population, so the loss to agriculture is not included in the computation. However, the EAL for Drought only quantifies the damage to crops and livestock (agriculture) in its computation. A consequence type is only included in the EAL computation for a hazard if at least 10% of the total reported economic loss due to the hazard (see the Natural Hazard Historic Loss Ratio section) is of that consequence type.

All loss is quantified as a dollar amount. While building and agriculture loss are quantified in dollars in the source data, population loss is quantified as the number of fatalities and injuries and must be converted to ensure all EAL values use a common unit of measurement. Population loss is monetized using the value of statistical life approach in which each fatality or ten injuries is treated as $7.4 million of economic loss, an inflation-adjusted Value of Statistical Life (VSL) used by FEMA.[6] To adjust for inflation, all historic losses are converted to 2016 dollars.

4.3.1. CALCULATING EXPECTED ANNUAL LOSS

EAL is calculated using a multiplicative equation that considers the consequence risk factors of natural hazard exposure and historic loss, and the likelihood risk factor of natural hazard frequency for 18 natural hazards. The EAL value for each consequence type is calculated by multiplying the total exposure value of an area by the estimated annual frequency of a natural hazard event and by the historic loss ratio (see Equation 3). See the Natural Hazard Expected Annual Loss Components section for further explanation of these EAL components and how they are computed. EAL values are computed at the Census block level (or for some hazards, the Census tract level) for each relevant consequence type and summed to produce a composite EAL for each hazard (see Equation 4). A cubic root transformation is applied to each hazard-specific EAL value to address skew. The resulting transformed values are then min-max normalized (0.00 – 100.00 scale) to produce an EAL score for

[4] Cutter, S.L., Ash, K.D., & Emrich, C.T. (2014). The geographies of community disaster resilience. *Global Environmental Change*, *29*, 65-77. https://doi.org/10.1016/j.gloenvcha.2014.08.005

[5] See also Mitigation Framework Leadership Group (MitFLG), Federal Emergency Management Agency (FEMA). (2016). Draft Interagency Concept for Community Resilience Indicators and National-Level Measures. Washington, DC: Department of Homeland Security (DHS). Retrieved from https://www.fema.gov/media-library-data/1466085676217-a14e229a461adfa574a5d03041a6297c/FEMA-CRI-Draft-Concept-Paper-508_Jun_2016.pdf

[6] Federal Emergency Management Agency (FEMA). (2016). *Benefit-cost sustainment and enhancements: baseline standard economic value methodology report.* Retrieved from https://www.caloes.ca.gov/RecoverySite/Documents/Benefit%20Cost%20Sustainment.pdf

each hazard. A total EAL is also summed from all hazard EALs for the area and a total EAL score is calculated using the same cubic root transformation and min-max normalization process.

Hazard-specific Risk Index scores are calculated using individual hazard EAL scores. Overall Risk Index scores are calculated using the composite EAL score.

Equation 3: Hazard-Specific Expected Annual Loss by Consequence Type

$$\begin{aligned} \mathit{Expected\ Annual\ Loss}_{\mathit{Hazard}_{\mathit{Consequence\ Type}}} &= \mathit{Exposure}_{\mathit{Hazard}_{\mathit{Consequence\ Type}}} \times \mathit{Frequency}_{\mathit{Hazard}} \\ &\times \mathit{Historic\ Loss\ Ratio}_{\mathit{Hazard}_{\mathit{Consequence\ Type}}} \end{aligned}$$

Equation 4: Composite Hazard-Specific Expected Annual Loss

$$\begin{aligned} \mathit{Expected\ Annual\ Loss}_{\mathit{Hazard}_{\mathit{Total}}} &= \mathit{Expected\ Annual\ Loss}_{\mathit{Hazard}_{\mathit{Building\ Value}}} \\ &+ \mathit{Expected\ Annual\ Loss}_{\mathit{Hazard}_{\mathit{Population\ Value}}} \\ &+ \mathit{Expected\ Annual\ Loss}_{\mathit{Hazard}_{\mathit{Agriculture\ Value}}} \end{aligned}$$

While each hazard uses the same components to calculate EAL, these computations require different approaches due to the varying nature of the hazards and the differences in source data format. A set of common analytical techniques (see the Expected Annual Loss section) are leveraged to achieve the best possible normalization between all hazards for accurate NRI calculation. The process for computing the EAL and its components for each individual hazard are described in the hazard-specific sections of the NRI Technical Documentation.

See Table 4 for a simplified example of a county-level EAL calculation for the hazard Hail. All three consequence types are included in the calculation of Hail EAL. By multiplying the county's consequence exposure, hazard frequency, and consequence-specific historic loss ratio, an EAL value for that consequence type is determined. The values for each consequence are summed to produce the composite EAL for the county. This composite EAL is used to derive the hazard's EAL score for that county. This computation includes a min-max normalization using the hazard-specific composite EAL values of all counties in the nation. The composite EAL for Hail is summed with the composite EAL values for the 17 other hazards to calculate the total EAL, which is scored in the same way.

Table 4: Example of a County-Level EAL Calculation for Hail

EAL Component	Building Value	Population	Agriculture Value
Exposure	$23.14 M	182,265 people or $1.35 T	$120,000
Frequency	9.7 events/year	9.7 events/year	9.7 events/year
Historic Loss Ratio	1.6e-8	3.2e-8	1.4e-7

EAL Component	Building Value	Population	Agriculture Value
Expected Annual Loss	$3,478	0.054 people or $399,954	$156

4.3.2. ANALYTICAL TECHNIQUES

Arriving at a dollar value representing the EAL due to each of the 18 hazards for every county and Census tract in the United States requires multiple analytical techniques utilized across all hazards to ensure the most accurate representation of loss.

NRI Processing Database

To support the processing of the NRI, a dedicated SQL Server database environment was established. Using a relational database to store and analyze each dataset used to compute the NRI provides a variety of benefits. The database allows for computational efficiencies when calculating the components of the EAL for more than 11 million Census blocks in the United States. Grouping and aggregation functions can be used easily to roll these values into the Census tract and county level values displayed in the NRI application. Implementation of NRI methodologies in stored procedures allows for application and adaptation of complex business logic and spatial analysis. The NRI processing database also makes quality control easier by allowing complex calculations to be processed in steps with output for each step accessible in its own table. Records for each Census block can be checked to identify outliers and any possible problems with the methodology or algorithms. Additionally, repeatable processes can be modified and run in smaller portions, cutting down on processing time as methodology is adapted. For example, a change in source data for a hazard only requires the replacement of hazard-specific source data tables and for the re-processing of a single hazard to be executed. The NRI processing database also supports version control and allows backups of each version to be stored securely.

Most spatial functions, such as buffering and intersection, are performed within the NRI processing database. However, some processes, such as land use tabulation necessitate the use of ArcGIS tools and functions. The output of these externally performed processes is transferred and stored within the NRI processing database where it is used to compute the components of the EAL.

Geographic/Administrative Layers

EAL components may be calculated at three different administrative layers: Census block, Census tract, and county. The most granular level is the Census block and, when possible, values are calculated at this level and then aggregated. The source of the boundaries for these layers is the US Census Bureau's 2017 TIGER/Line shapefiles.[7] The shapefiles include US territories and some large bodies of water. These are either manually removed or clipped based on a County boundary

[7] US Census Bureau. (2017). *Cartographic Boundary Shapefiles* [cartographic dataset]. Retrieved from https://www.Census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.2017.html

shapefile provided by Esri.[8] All spatial layers use the North America Albers Equal Area Conic projection. Figure 4 provides examples of block, tract, and county boundaries.

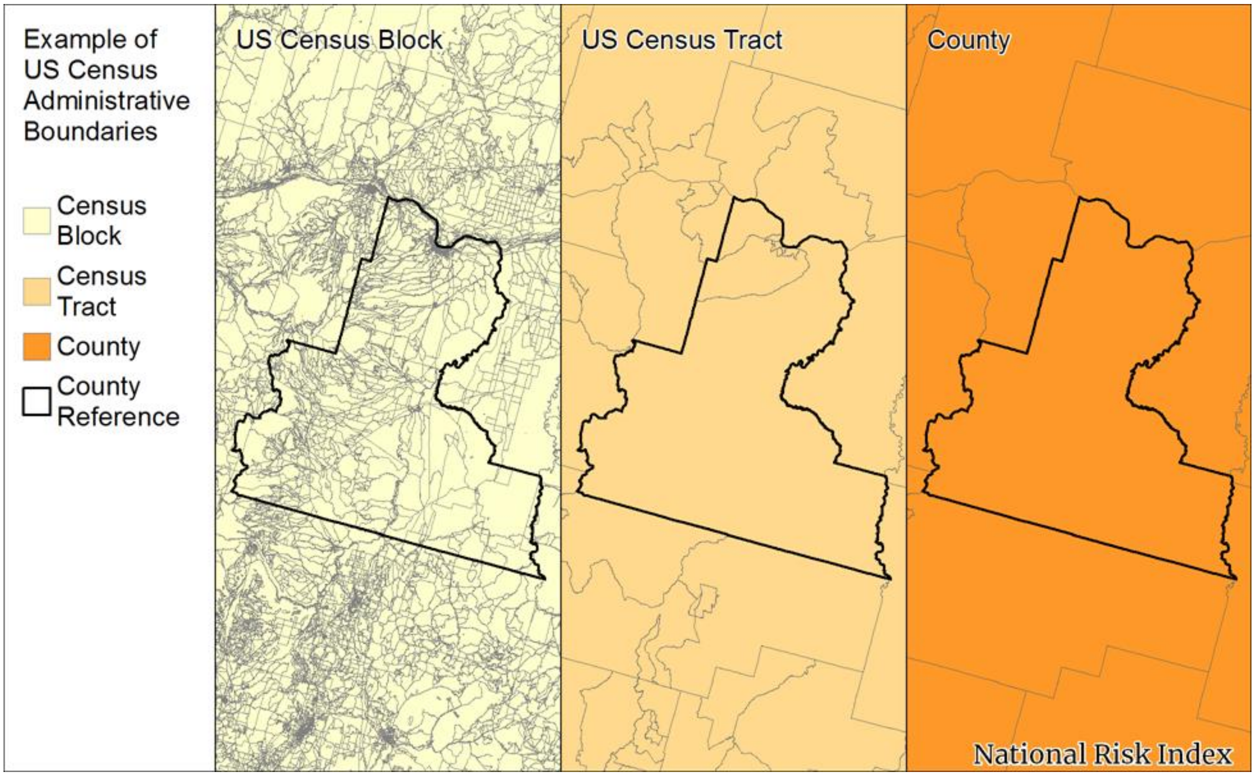

Figure 4: Example of County, Census Tract, Census Block Shapes

Determining County-Level Possibility of Hazard Occurrence

Not all hazards are able to occur in all areas. For example, Coastal Flooding cannot occur in Kansas and Avalanches cannot occur on flat terrain. The NRI logically differentiates areas where a given hazard is unlikely or has never occurred from areas where that hazard is impossible using a control table in the database that designates where each hazard can occur. This table is based on counties that intersect past hazard event polygons generated through spatial processing or which have some possibility of occurrence as identified by probabilistic or susceptibility source data or which have recorded loss due to hazard occurrence.

Base Calculation and Aggregation

One of the NRI's strengths is that it determines the EAL for an area at the lowest geographical level deemed appropriate, predominantly the Census block level. EAL is determined by assessing the combination of a specific location's frequency of occurrence and associated consequence if it were to occur (for example, how often Riverine Flooding occurs in the area and what buildings, population

[8] Esri, TomTom North America, Inc., & US Census Bureau. (2012). *USA County Boundaries* [cartographic dataset]. Retrieved from https://www.arcgis.com/home/item.html?id=f16090f6d3da48ec8f144a0771c8fec4

and crops are potentially affected). For many hazard types, frequency and exposure can be highly localized. Modeling the event frequency in coordination with its exposure provides the best assessment of its expected impact.

The Census block is currently the lowest administrative level at which population and building value data are nationally, consistently, and publicly available. By performing the EAL calculation at the Census block level, the NRI is more accurately assessing EAL by looking at specific frequency and exposure combinations at the lowest possible resolution. The NRI provides the most relevant aggregations to its users, namely EAL values at the Census tract and county levels. For most hazards, Census tract and county level exposure and frequency are calculated by “rolling up” or aggregating values from the Census block level.

Representation of Hazards as Spatial Polygons

EAL components for each hazard are derived from one or more sources of spatial hazard information. This can include identified hazard-susceptible areas, spatiotemporal records of past hazard occurrences, and countywide records of economic loss due to a hazard event. The format of spatial source data varies by hazard. Frequency and exposure calculations typically require spatiotemporal records of past hazards or probabilistic modelling. To achieve a uniform level of accuracy, any spatial hazard source data were converted to vector polygon format and intersected with the Census blocks or tracts.

Necessary conversions are performed either with tools available in Esri’s ArcGIS software or with SQL Server’s spatial operations. Common methods of hazard conversion used for NRI calculation are the buffering of points and lines to form polygons, and raster-to-polygon conversion.

Point and line representations of hazard events or hazard-susceptible areas are buffered by different distances depending on the hazard. Point buffers allow for better representation of event coverage or area of possible impact. Path representations, such as those for Tornado and Hurricane, are included in the source data as a series of points with a common identifier (e.g., StormID). These are connected by a line or multi-segmented line. The line is then buffered by a distance depending on the intensity of the Tornado (Enhanced-Fujita scale) or Hurricane (Saffir-Simpson scale) event. See the spatial processing discussion in the hazard-specific sections of the NRI Technical Documentation for more detail on buffering techniques.

Conversion from raster to polygon vector format is performed by using ArcGIS’s Create Fishnet tool to form a grid of rectangular cells that match the extent and dimensions of the original raster and then using the Extract Values to Table tool to insert the cell values of the raster into the corresponding fishnet polygon’s attribute table. In vector format, attributes from the source raster data can be used to filter or select the data needed for hazard specific methodology calculations.

Intersection

Determining areas of spatial intersection between hazard events or susceptible areas and the various levels of reference layers is an essential function used in calculating EAL. The results of these intersections are stored in the NRI processing database and used for multiple purposes. For many hazards, the quantification of a hazard's exposure is done at the Census block level. This requires the computation of intersecting areas of exposure. Figure 5 provides an example of a hazard event shape intersecting a Census block.

Frequency computations also typically involve counting the number of hazard event polygons that intersect the Census block. Widespread hazards, like Hurricanes, often require a larger administrative layer to more accurately represent the frequency of Hazard events. For these types of hazards, the intersection is performed with a 49-by-49 km fishnet grid and the count of the fishnet grid cell is inherited by the Census blocks it encompasses, using an area-weighted value when a Census block intersects more than one cell.

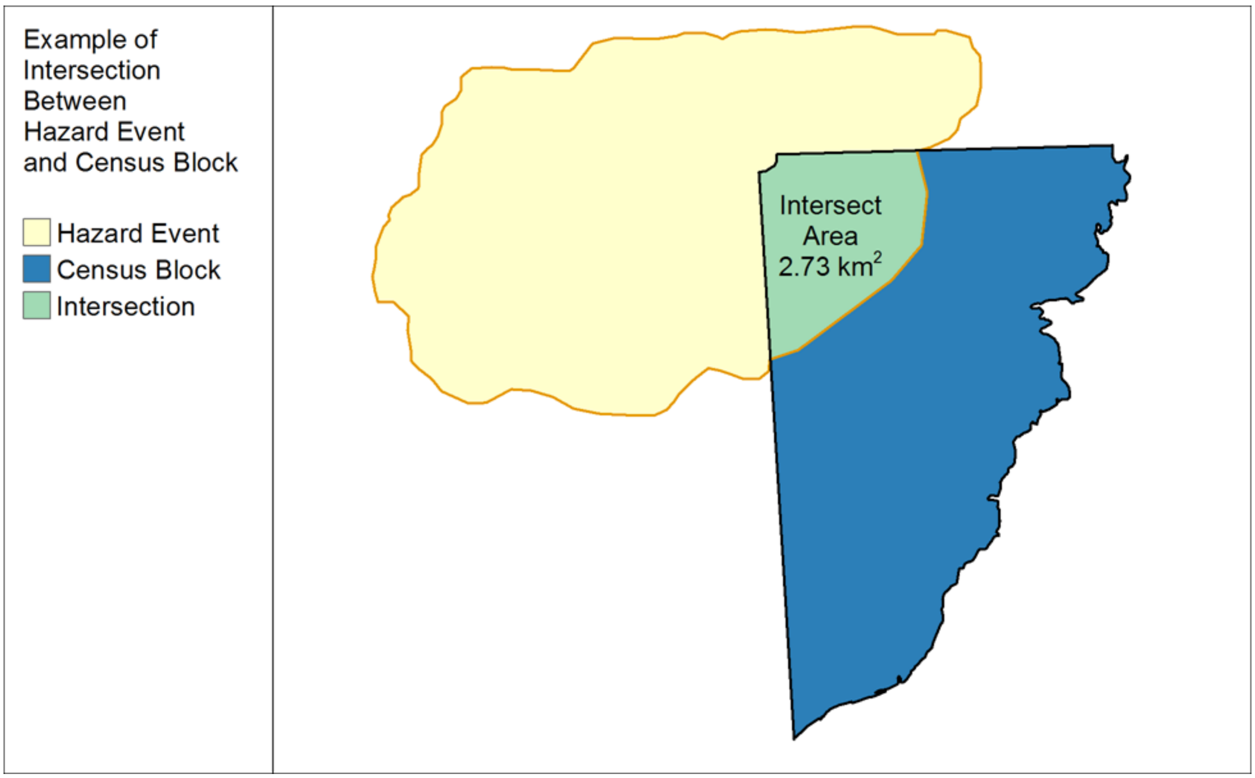

Figure 5: Example of Intersection Between Hazard Event and Census Block

The 49-by-49 km grid cell size was used because of analysis conducted early in the project which roughly estimated the average Census tract size to be 4,900 m^2 (or 70-by-70 m) and the average county size to be 2,500 km^2 (or 50-by-50 km), which was reduced slightly to 49-by-49 km to ensure the county size was a multiple of the tract size. Though the use of a grid at the average Census tract resolution was discarded, the use of the 49-by-49 km fishnet grid was maintained for the calculation of frequency for widespread hazards.

Tabulation

Tabulation refers to the process of calculating the composition of a vector shape by overlaying it on a raster layer inside a GIS. The GIS computes the area of raster cells completely contained within the vector shape by raster value.

The land use tabulation process is performed by using the Tabulate Area tool in Esri's ArcGIS software. All spatial layers use the North America Albers Equal Area Conic projection. A layer containing county boundaries is tabulated against the 2017 CropScape raster file[9], which describes the land use of the conterminous United States in 30-by-30-m cells using 132 distinct raster values. The output layer contains a record for each county (by county FIPS code) with fields for each class (crop types, developed areas, etc.) displaying the area (in square meters) of each type of land use within the county. There are five classes of developed area (Developed, Developed Open Space, and Developed Low, Medium, and High Intensity) which can be summed to get the total developed area of the county. The area values of all crop classes can be summed to give a total agricultural area. This same tabulation is performed at the Census tract and Census block level to support the computation of developed area densities at these levels. The EAL calculations for most hazards utilize the developed area density values at the Census block level (see the Approach 1. Developed Area Density Concentrated Exposure).

The CropScape layer only contains information for the conterminous United States. For Alaska and Hawaii, a similar tabulation process is carried out substituting the 2016 National Land Cover Database (NLCD) raster files[10] for both states. NLCD uses the same classification types for developed land as CropScape. It has two classifications for agricultural land: Pasture/Hay and Cultivated Crops.

Primary tabulation involves summing the total area of interest (e.g., developed land use) and dividing by the total area of raster cells contained. The shape area (e.g., Census block, Census tract, or county) is multiplied by this developed area percent to calculate the developed area (in square kilometers). To speed up calculations, the intersected shapes are classified as whether they completely contain the Census block, tract, or county (for which developed area and crop/pasture area had already been calculated). For such shapes, the values were transferred over without tabulation. Tabulated areas are approximations based on the cell size of the source raster and can exceed the area of the shape being tabulated. In these cases, the total area of the shape is set as the ceiling of the tabulation area results.

Very small intersections of hazard event shapes with Census blocks can be too small to tabulate against 900-m^2 raster cells. If not, all shapes are tabulated using the primary method, secondary methods are pursued. Secondary methods are hazard-specific. For example, secondary tabulation of

[9] US Department of Agriculture (USDA), National Agricultural Statistics Service (NASS). (2017). *Published crop-specific data layer* [online dataset]. Retrieved from https://nassgeodata.gmu.edu/CropScape/

[10] Multi-Resolution Land Characteristics Consortium. (2016). *National Land Cover Database (NLCD)* [online dataset]. Retrieved from https://www.mrlc.gov/data

Drought-Census tract shapes involves extracting the raster value at the centroid of the shape. The entire area of the shape is classified as the raster value extracted at the centroid. On the other hand, Riverine Flooding shapes, as many administrative boundaries are drawn using rivers, are winding and narrow (see the shape on the right in Figure 6). A centroid-based approach is not the most accurate. For this reason, raster cell centroids representing developed areas were exported. SQL Spatial routines then calculated whether a developed land-use was within 42 meters (the hypotenuse distance of a 30-by-30 m raster cell). If so, the entire shape was deemed developed. If not, the shape was considered to have zero developed area.

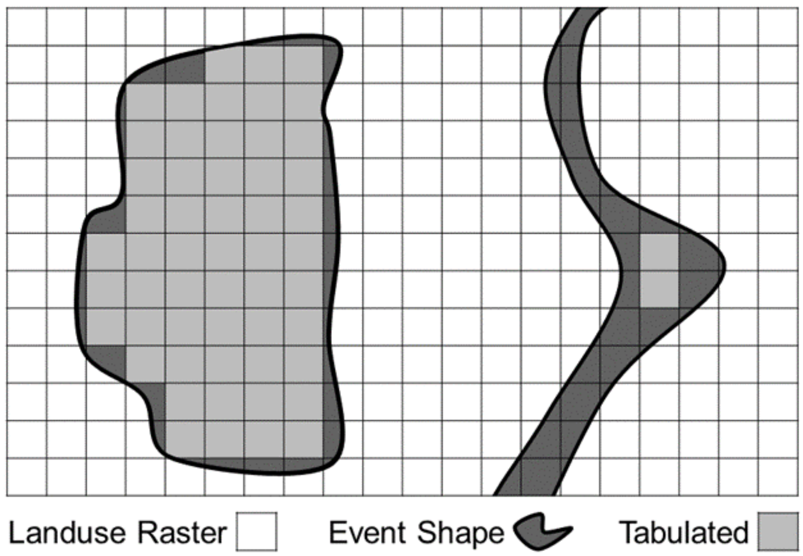

Figure 6: Land Use Raster Tabulation

(This section intentionally left blank.)

(This section intentionally left blank.)

5. Natural Hazard Expected Annual Loss Components

The NRI represents natural hazard in terms of EAL, which incorporates data for natural hazard exposure, annualized frequency, and historic loss. A single "mental model" was leveraged throughout all methodological processes in calculating these EAL components, so that certain risk factors were not being interpreted inconsistently across the 18 natural hazards.

5.1. Natural Hazards

Natural hazards are defined as environmental phenomena that have the potential to impact societies and the human environment. These should not be confused with other types of hazards, such as manmade hazards. For example, a flood resulting from changes in river flows is a natural hazard, whereas flooding due to a dam failure is a considered manmade hazard by the NRI.

Natural hazard events can induce secondary natural hazard events. For example, Landslides can be caused by an Earthquake. Natural hazards are distinct from natural disasters. A natural hazard is the threat of an event that will likely have a negative impact. A natural disaster is the negative impact following an actual occurrence of the natural hazard in the event that it significantly harms a community. The NRI only considers primary natural hazard events and not their results or after-effects.

The NRI considers 18 natural hazards, including: Avalanche, Coastal Flooding, Cold Wave, Drought, Earthquake, Hail, Heat Wave, Hurricane, Ice Storm, Landslide, Lightning, Riverine Flooding, Strong Wind, Tornado, Tsunami, Volcanic Activity, Wildfire, and Winter Weather. These hazards are listed below and described in more detail in the NRI Technical Documentation.

5.2. Natural Hazard Annualized Frequency

The annualized natural hazard frequency is defined as the expected frequency or probability of an event happening per year. Frequency is derived either from the number of recorded events each year over a given period or the modeled probability of an event occurring each year. The NRI considers that natural hazards can occur in places where they may have not yet been recorded to-date and that hazards may have occurred in locations without being recorded. Therefore, the NRI has built-in minimum representative frequency values for certain geographical areas and hazards, such as Hurricane, Ice Storm, Landslide, Tornado, and Tsunami.

5.2.1. SELECTING SOURCE DATA

Annualized frequency data are derived from multiple sources and depend on the natural hazard. Data sources were identified through public knowledge, guidance by SMEs, and research. Examples of selected data sources include the National Weather Service (NWS), the National Oceanic and Atmospheric Administration (NOAA), the US Geological Survey (USGS), the US Army Corps of Engineers (USACE), the Smithsonian databases, and the US Department of Agriculture (USDA). See the hazard-specific sections in the NRI Technical Documentation for more information on spatial data sources.

5.2.2. ANNUALIZED FREQUENCY METHODOLOGY

The natural hazard annualized frequency is the expected frequency for a given hazard event and measures the actual or expected number of events or event days each year. Not all events are considered relevant for frequency calculation. SMEs established that some hazards meet certain criteria to be included as a hazard event capable of causing damage e.g., Hail size of diameter greater than 0.75 in. (see the hazard-specific sections for more information on these criteria). Annualized frequency can be defined as the number of historical occurrences of a natural hazard within a known period of record per geographic area, as seen below in Equation 5:

Equation 5: Annualized Frequency Equation

$$Annualized\ Frequency = \frac{Number\ of\ Recorded\ Events}{Period\ of\ Record}$$

In some cases, as with Wildfire and Earthquake, the best available source data consists of probabilistic statistics contained in raster files which are used to compute an annualized frequency. In these cases, the frequency value represents the probability of a hazard event occurring in a given year.

For hazards that track actual hazard occurrences, the historical event count quantifies either the number of distinct hazard events that have occurred (e.g., Hurricanes to hit the area) or the count of days on which a hazard has occurred (e.g., on how many days a Hail event was reported). The determination of whether hazard occurrence was defined by event-days or discrete events was based on SME review of the source data. This determination depended on how hazard occurrence was recorded as well as how economic loss was reported. Table 5 gives the frequency basis (event or event-day) for each hazard.

Table 5: Geographic Level of Event Count Determination and Hazard Occurrence Basis

Natural Hazard	Geographic Level of Historic Event Count Determination	Hazard Occurrence Basis
Avalanche	County	Distinct events
Coastal Flooding	No event count	No event count
Cold Wave	Census Block	Event days
Drought	Census Tract	Event days
Earthquake	No event count	No event count
Hail	49-km Fishnet	Event days
Heat Wave	Census Block	Event days
Hurricane	49-km Fishnet	Distinct events
Ice Storm	49-km Fishnet	Event days
Landslide	Census Tract	Distinct events
Lightning	4-km Fishnet (Source raster cell)	Distinct events
Riverine Flooding	County	Distinct events
Strong Wind	49-km Fishnet	Event days
Tornado	49-km Fishnet	Distinct events
Tsunami	Census Tract	Distinct events

Natural Hazard	Geographic Level of Historic Event Count Determination	Hazard Occurrence Basis
Volcanic Activity	Census Block	Distinct events
Wildfire	No event count	No event count
Winter Weather	Census Block	Event days

While the NRI application reports information at the Census tract and county level, often the data used to determine this information is captured at either a lower or higher level. Predominantly, EAL components are assessed at the Census block level, so the number of hazard events (or event-days) which have historically occurred is determined for each Census block.

Depending on the nature of the hazard and its source data, the event count used to calculate frequency can be initially captured at the Census block, Census tract, county, or 49-by-49 km fishnet grid cell level (see each hazard's frequency section in the NRI Technical Documentation for specific hazard event count methodology). Table 5 provides the geographic level at which event count information is determined for use in frequency calculations for each hazard.

For large geographic areas and areas with a statistically significant number of events recorded, the logic supporting Equation 5 is sound and is used as one approach for calculating annualized frequency in the NRI for some natural hazards. However, for hazards with few events historically recorded, due to urban bias and varying demographics across the country, this equation is not always accurate or representative. Additionally, as geographic boundaries are partitioned into much smaller regions (counties, Census tracts, and Census blocks), further challenges are uncovered resulting from the fact that geographic areas that have not been historically impacted by a hazard and/or recorded hazard events are being calculated as having no risk from that hazard (since the EAL and NRI risk equation is multiplicative, and therefore any individual factor of zero results in a total NRI score of 0).

Consider an example (Figure 7) where four Tornadoes hit a single Census tract (e.g. "Tract A") near its geographic border. Using Equation 5, the annualized frequency for "Tract A" would be calculated using a 4 in the numerator. However, given the Tornado event locations (specifically, their proximity to the neighboring tracts), these four events could easily have occurred within, say, "Tract B". Therefore, "Tract B" should not be represented as having no (zero) risk, and, yet, it would be zero if annualized frequency was deemed to be zero based on the fact that no Tornado has historically occurred in "Tract B". Natural hazard events cannot be expected to respect arbitrarily drawn political boundaries, so, in evaluating risk, hazard occurrence definition should account for events in nearby Census blocks or tracts that easily could have impacted a given area.

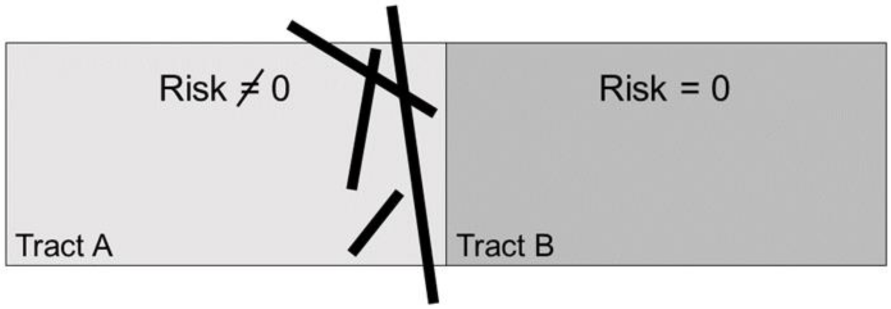

Figure 7: Example of the Issues with a Simplistic Annualized Frequency Methodology

Three main solutions were incorporated to spread the area of hazard influence used to calculate frequency and/or exposure. Hazard-specific frequency methodologies may use some or all of these approaches:

1. **Hazard Event Counting Using a 49-by-49 km Fishnet Grid:** This approach involves creating a fishnet grid covering the United States and counting the number of events (or event-days) of hazard occurrence within each cell. Areas within the cell inherit the event count (or receive an area-weighted event count when intersecting multiple cells; see the Data Aggregation section) and frequency is then calculated according to Equation 5. Hazards using this approach include Hail, Hurricane, Ice Storm, Strong Wind, and Tornado.
2. **Minimum Annual Frequency:** A minimum annual frequency (MAF) is assigned to areas which have not experienced a hazard occurrence recorded by the source data, but are determined to be at some risk due to their location (see the Determining County-Level Possibility of Hazard Occurrence section). Appropriate MAF values were identified by natural hazard SMEs. The estimated values were typically low, given the fact that historic events had never been recorded over the period of record, which sometimes dated back multiple centuries. Minimum values were typically defined in the format of “once in the period of record,” or similar. Hazards using this approach include Avalanche, Hurricane, Ice Storm, Landslide, Riverine Flooding, Tornado, and Tsunami.
3. **Hazard Event Shape Buffering:** Hazards with widespread and/or unpredictable event locations are buffered using SME-determined distances to create more representative areas with potential exposure to natural hazards. Buffering also allows events with relatively small surface areas to be smoothed together into general representative shapes to eliminate gaps that may exist between historically recorded hazard events (see Figure 8). Hazards using this approach include Hail, Hurricane, Strong Wind, Tornado, Tsunami, and Volcanic Activity.

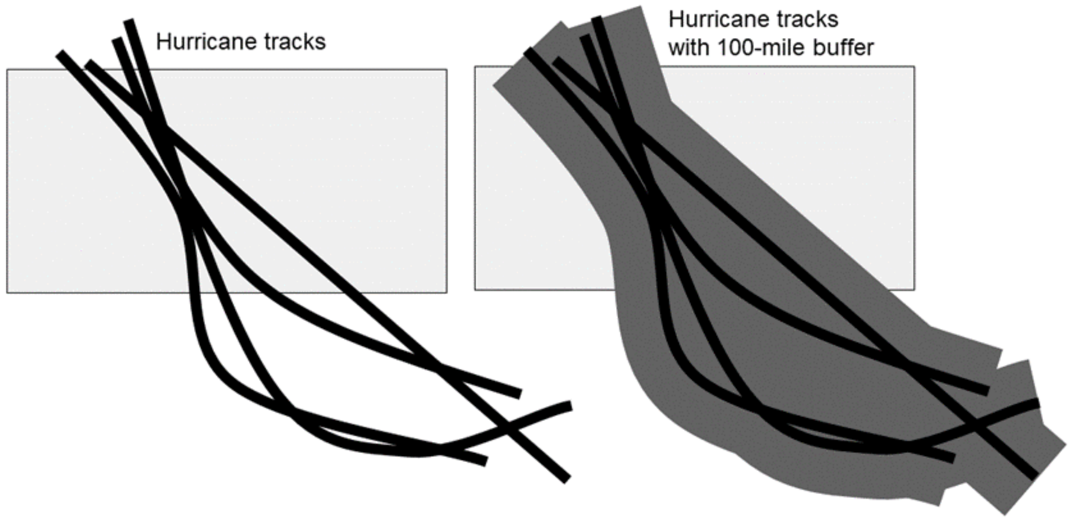

Figure 8: Example of Buffering Hazard Events to Determine Areas Applicable to Minimum Frequency Values

Some hazards do not require any of these solutions due to the nature of the source data or the widespread prevalence of the hazard. For example, the spatial data for Cold Wave, Heat Wave, and Winter Weather events cover areas the size and shape of NWS forecast zones and counties. These events can occur across the entire United States, so it is not necessary to spread the hazards' area of influence any further.

5.2.3. DATA AGGREGATION

In most instances, annualized frequency is calculated first at the Census block level. In cases where the event count is evaluated at the fishnet level (see Table 5), the Census block inherits the event count from the fishnet cell that encompasses it, performing an area-weighted count if a Census block intersects multiple fishnet cells, as computed in Equation 6. Applying this equation to the example in Figure 9 results in a Census block event count of about 22. This fishnet-aggregated count is used to calculate the Census block frequency.

Equation 6: Census Block Area-Weighted Fishnet Event Count Calculation

$$Census\ Block\ Event\ Count = \frac{\sum(Fishnet\ Event\ Count \times Area\ of\ Fishnet\ Intersection\ )}{Area\ of\ Census\ Block}$$

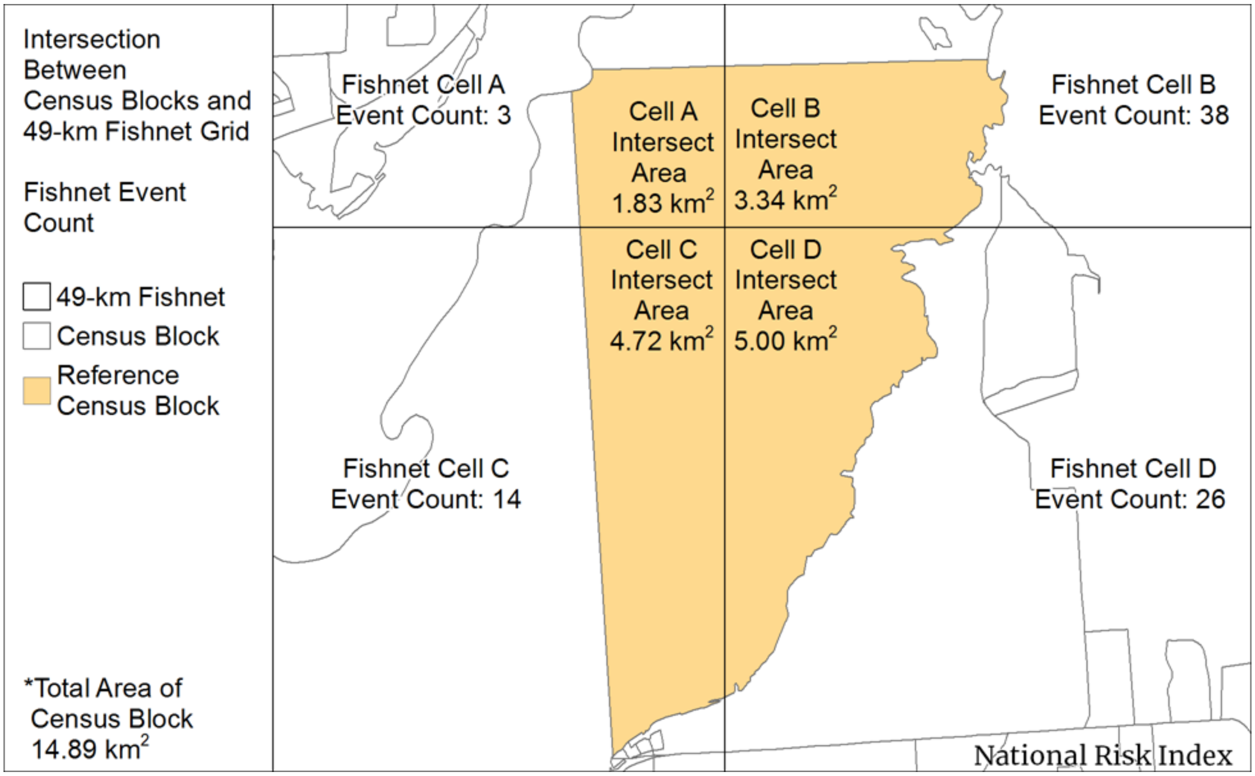

Figure 9: Aggregation from Fishnet Cell to Census Block Example

The NRI rolls up data from the Census block to the Census tract and county level, usually by leveraging area-weighted aggregation as computed in Equation 7. These Census tract and county level frequency values may not exactly match that of dividing the Census tract and county level number of historical hazard events by the period of record, as they are based on an area-weighted aggregation.

Equation 7: Census Tract and County Frequency Aggregations

$$Census\ Tract\ Frequency = \frac{\sum(Census\ Block\ Frequency \times Area\ of\ Census\ Block)}{Area\ of\ Census\ Tract}$$

$$County\ Frequency = \frac{\sum(Census\ Block\ Frequency \times Area\ of\ Census\ Block)}{Area\ of\ County}$$

For a few natural hazards (typically those that are widespread, such as Tsunami or Drought), annualized frequency is calculated at the Census tract level, after which the Census block simply inherits the value of its parent tract (see Table 5). Only annualized frequency of the Avalanche and Riverine Flooding natural hazards are calculated at the county level directly, where the Census tracts and blocks inherit the value of their parent county.

5.3. Natural Hazard Exposure

Natural hazard exposure is defined as the representative value of buildings, population, or agriculture potentially exposed to a natural hazard event. Data sources with the best available national-level data for each hazard were selected to perform a spatial analysis and compute areas of exposure.

5.3.1. SELECTING SOURCE DATA

The initial spatial processing of the source data for each hazard is used to identify areas of natural hazard exposure. Data sources were selected for their accuracy, long period of record, and spatial component, based on the best available, national-level data per natural hazard. Sources were identified through public knowledge, subject matter expert recommendations, and research. Providers of natural hazard exposure data include:

- National Oceanic & Atmospheric Administration (NOAA)
- USC Hazards & Vulnerability Research Institute (HVRI)
- Spatial Hazard Events & Losses Database for the United States (SHELDUS)
- United States Army Corps of Engineers (USACE)
- United States Geological Survey (USGS)
- United States Department of Agriculture (USDA)
- National Weather Service (NWS)
- Federal Emergency Management Agency (FEMA)

5.3.2. CONSEQUENCE TYPES

A natural hazard consequence is defined in the NRI as economic loss or bodily harm to individuals that is directly caused by a natural hazard event. Consequences of natural hazard events are categorized into three different types: buildings, population, and agriculture.

Buildings

Building exposure is defined as the dollar value of the buildings determined by the source data to be exposed to a hazard according to a hazard-specific methodology. The maximum possible building exposure of an area (Census block, Census tract, or county) is its building value as recorded in Hazus 4.2, Service Pack 01 (SP1),[11] which provides 2018 valuations of the 2010 Census.[12]

Population

Population exposure is defined as the estimated number of people determined by the source data to be exposed to a hazard according to a hazard-specific methodology. The maximum possible population exposure of an area (Census block, Census tract, or county) is its population as recorded

[11] Federal Emergency Management Agency (FEMA). (2018). Hazus 4.2, Service Pack 01 Release. Retrieved from https://msc.fema.gov/portal/resources/hazus

[12] US Census Bureau. (2010). 2010 Census. Retrieved from http://www.Census.gov/2010Census/data/

in Hazus 4.2 SP1. The Value of Statistical Life (VSL) was used to express population exposure in terms of dollars.

Agriculture

Agriculture exposure is defined as the estimated dollar value of the crops and livestock determined by the source data to be exposed to a hazard according to a hazard-specific methodology. This is derived from the USDA 2017 Census of Agriculture[13] county-level value of crop and pastureland.

5.3.3. EXPOSURE METHODOLOGY

Natural hazard exposure is typically calculated at the Census block level and then aggregated to the tract and county level by summing the block exposure values within the parent tract or parent county. See each hazard's exposure section for more information.

Some hazard exposure areas are represented as polygons in the source data, while others are represented as points, lines, or raster cells. Hazard exposure is based on either historic event locations or areas of identifiable risk, e.g., Tsunami inundation zones. Eventually, every relevant record in the source data is processed into a polygon via a hazard-specific methodology. This polygon represents an area of exposure to the hazard.

To calculate the natural hazard's representative size for a given area, the NRI leverages a few techniques, such as using subject matter expertise to define a single representative hazard size, calculating historical average event occurrence sizes, or defining the size of probabilistic/susceptible zones for hazards within the area of interest using existing source data (Figure 10).

[13] US Department of Agriculture. (2017). 2017 Census of Agriculture. Retrieved from https://www.nass.usda.gov/Publications/AgCensus/2017/index.php

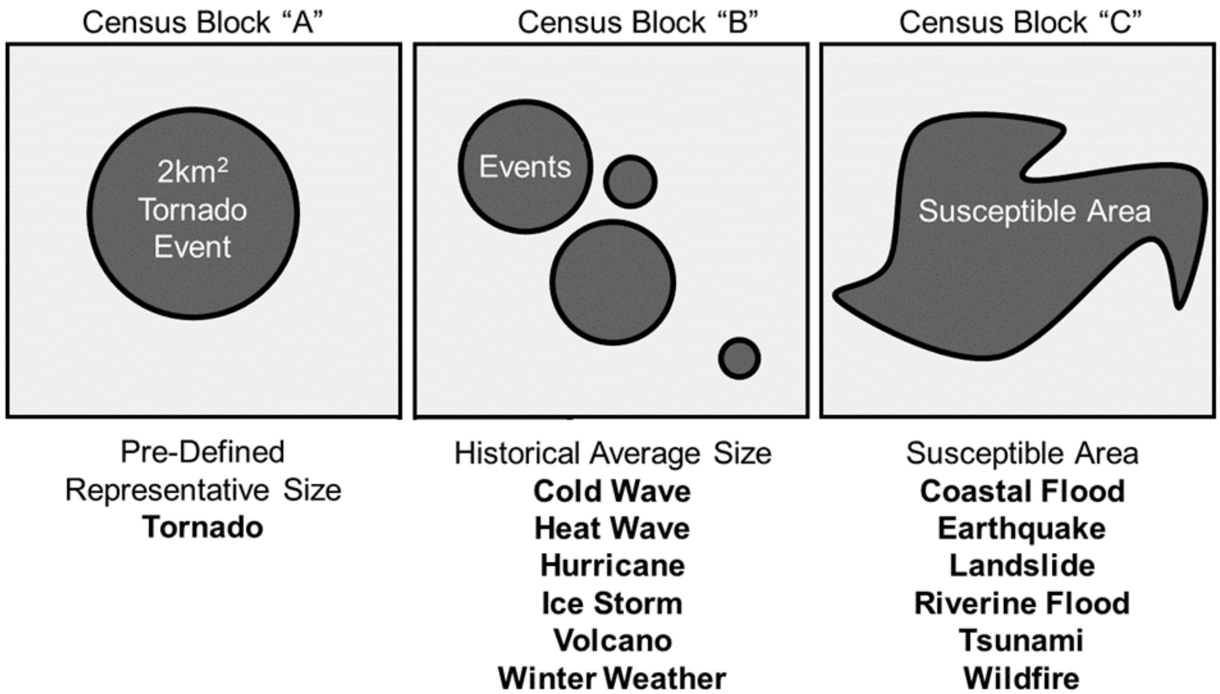

Figure 10: Examples of Representative Hazard Size

To estimate exposure, the hazard event or susceptible area polygons are intersected with the appropriate administrative layer polygons and the resulting intersect shape defines the area of hazard exposure. Once the area of exposure is defined, one of three generalized approaches are executed within the NRI processing database to estimate the exposure value within the administrative area. The approach used for a natural hazard was determined by the hazard's recorded historic events, hazard susceptibility maps, and subject matter expertise. The type of exposure method used for each of the 18 hazards is described further in the NRI Technical Documentation. The general approaches to modeling exposure include:

1. **Developed Area/Agricultural Area Density Concentrated Exposure.** The NRI determined area of hazard exposure intersected with the administrative area is multiplied by the density of either the population or building value within the developed land of the area to calculate the worst-case concentration of hazard consequence. To estimate agriculture exposure, this method uses the density of crop and livestock value within the agricultural land of the area.
2. **Widespread Hazard Event Exposure.** The entire Census block is considered to be exposed. This approach is leveraged for hazards whose extent likely spans the entire area of interest and whose boundaries are indefinable.
3. **Pre-Defined Representative Exposure.** Subject matter experts defined a default, representative exposure value for areas of interest deemed at risk of natural hazard events.

Approach 1. Developed Area Density Concentrated Exposure

Exposure is calculated for most of the natural hazards using the developed area density approach. This approach uses the area of the hazard event exposure shape (intersection of hazard shape with the administrative area) multiplied by the developed area density of the administrative area to generate the worst-case representative property damage or population that could result from a future natural hazard event within the area.

The Hazus 4.2 SP1 data provides building value and population estimates at each administrative reference layer (Census block, Census tract, and county). For certain hazards, a density estimate was needed for the hazard's exposure calculation. Rather than only calculating an average density value for each administrative layer (i.e., by dividing the population of a Census block by the area of the Census block), effort was made to refine the density estimate by first estimating where people and buildings might exist within an area. Using the USDA CropScape 2017 raster, which categorizes land types and use (see Figure 11), a spatial tabulation process was used to derive an estimate of the developed area within each administrative reference layer. This same tabulation process was used to estimate the crop and pasture area as well (see the Tabulation section).

With an estimate of the developed area and crop and pasture area for each record of the administrative reference layers, densities were then calculated. Using the Hazus data's Building Stock Value and Population estimates for each administrative layer, the ratio of developed area within an administrative reference over its whole area was used to calculate the building value and population densities. These densities represent an assumption that population and the presence of buildings are concentrated in developed areas rather than being equally distributed across an administrative area.

Note that, in cases where the Hazus data reports population and or building value and the tabulation process did not identify any developed land area, the record was assigned an average density value calculated as the building value (or population) divided by the total area of the record. For cases where the tabulation process identified developed area but the Hazus data did not report any population or building values, the densities where set to 0. This ensures that the tabulation process, which can be spatially imprecise due to the resolution of the source rasters, does not count adjacent developed area as developed area within the administrative area when Hazus data does not consider it populated or developed.

To compute the building and population value densities, the building and population values of the administrative layer (Census block, Census tract, or county) are divided by the total developed area (determined for the tabulation process) of the administrative layer, as in Equation 8.

Equation 8: Census Block Building and Population Value Density

$$BldgValueDen_{CB} = \frac{BldgValue_{CB}}{DevArea_{CB}}$$

$$PopDen_{CB} = \frac{Pop_{CB}}{DevArea_{CB}}$$

where:

$BldgValueDen_{CB}$	is the building value density calculated at the Census block level (in dollars per square kilometer)
$BldgValue_{CB}$	is the total building value of the Census block, as recorded in Hazus 4.2 (in dollars)
$DevArea_{CB}$	is the total developed area of the Census block, tabulated from CropScape or NLCD raster files (in square kilometers)
$PopDen_{CB}$	is the population density calculated at the Census block level (in people per square kilometer)
Pop_{CB}	is the total population of the Census block, as recorded in Hazus 4.2

For agriculture, the USDA 2017 Census of Agriculture provides an estimated dollar value of crop and livestock within each county. The county value is divided by the total agricultural area of the county to find its crop value density (see Equation 9). The county level agricultural value density is inherited by any Census tracts or Census blocks that contain crop or pastureland.

Equation 9: County Crop Value Density

$$AgValueDen_{Co} = \frac{AgValue_{Co}}{AgArea_{Co}}$$

where:

$AgValueDen_{Co}$	is the agricultural value density calculated at the county level (in dollars per square kilometer)
$AgValue_{Co}$	is the total crop and livestock value of the county, as reported in the 2017 Census of Agriculture (in dollars)
$AgArea_{Co}$	is the total agricultural area of the county (in square kilometers

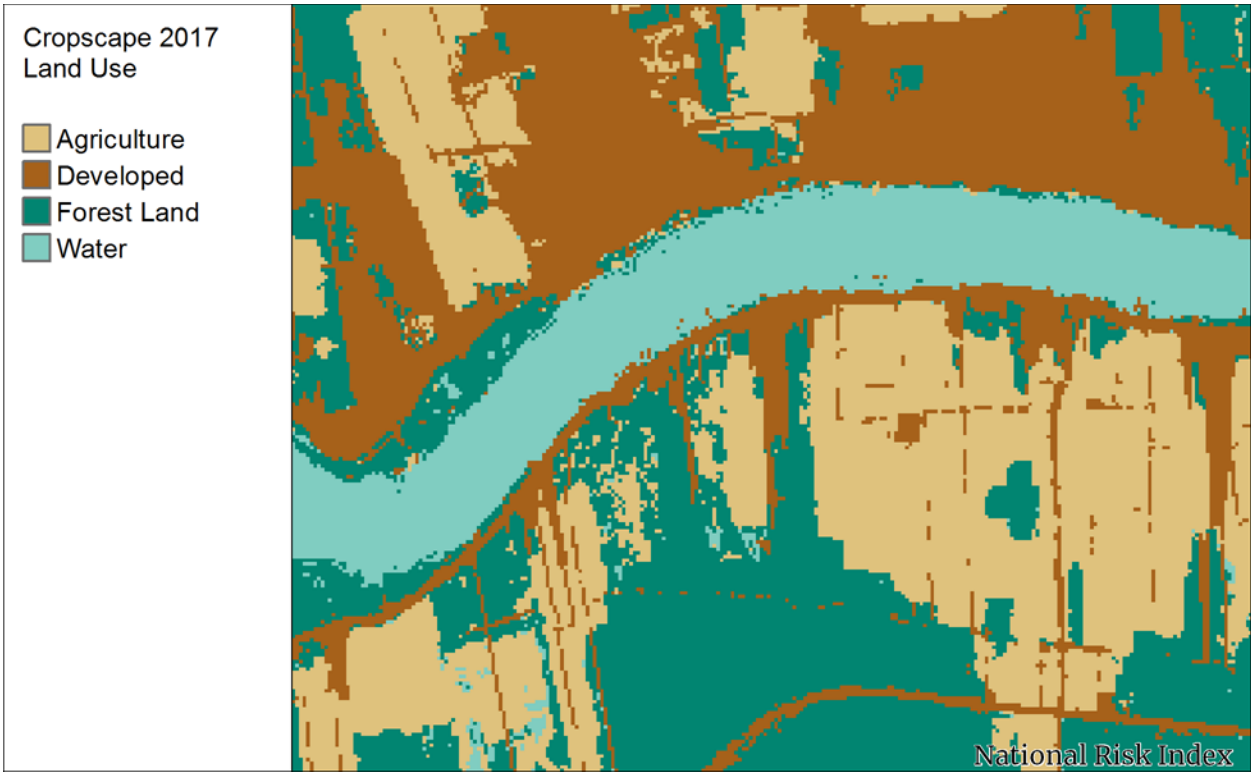

Figure 11: CropScape Developed Land Layer

Approach 2. Widespread Hazard Event Exposure

For certain natural hazards whose extent is widespread with indefinable boundaries, the entire area of interest is considered exposed. For these natural hazards, exposure values are defined to be the entire area of interest's building value, crop and livestock value, or population as recorded by Hazus 4.2 SP1 or the 2017 Census of Agriculture.

Approach 3. Hazard-Specific Representative Exposure

Avalanche and Tornado each have a unique method of calculating exposure. For Avalanche, a single exposure value, defined by SMEs, is pre-determined and assigned to all areas deemed at risk of Avalanche events. A review of the source data found that 98% of historical Tornado events impact an area of 50 km^2 or less, with the average damage area being 2.07 km^2, so a 2 km^2 area was used to estimate an average area impacted by a Tornado. This representative footprint area is multiplied by the average building or population density of the Census tract to find exposure.

5.3.4. DATA AGGREGATION

Natural hazard exposure is calculated at the Census block level and then is aggregated to the tract and county level by summing the block exposure values within the parent tract or parent county (with the exception of Avalanche, Drought, Earthquake, and Tornado which are initially calculated at the tract level). Detailed methodologies per hazard are explained in the hazard-specific sections of the NRI Technical Documentation.

5.4. Natural Hazard Historic Loss Ratio

The Historic Loss Ratio (HLR) is an area-specific estimate of the percentage of the exposed consequence type (building value, population, or agriculture value) expected to be lost due to a single hazard occurrence. In concept, it is the average of the loss ratios associated with past hazard events and is used to estimate the potential impact of a future hazard events. To begin the determination of this value, a Loss Ratio per Basis (event or event-day) (LRB) is calculated for each historical loss-causing hazard occurrence (for each relevant consequence type) as the value of the loss divided by the exposed consequence value.

A Bayesian credibility analysis is then performed with the individual LRBs at multiple geographic levels (county, surrounding area, regional, and/or national) to better balance historic loss accuracy with geographic precision and characteristics. The resulting HLR (by consequence type) is a Bayesian-adjusted ratio that is the summed weighted average of various geospatial groupings of the consequence LRBs at the relevant geographic levels for the hazard. This Bayesian-adjusted resulting HLR value, computed for each County-Hazard-Consequence type combination, serves as a prediction of the ratio of loss to exposed consequence value that can be expected from a single hazard occurrence. Computation of the HLR also considers hazard events which resulted in no loss prior to performing the Bayesian credibility spatial modelling analysis. This ensures that HLR can be multiplied by frequency within the risk equation without over-inflating the EAL value.

5.4.1. SELECTING SOURCE DATA: SHELDUS

Historic Losses source data provider: Arizona State University, Spatial Hazard Events and Losses Database of the United States (SHELDUS)[14]

Arizona State University's Spatial Hazard Events and Losses Database of the United States (SHELDUS) loss data were used for most natural hazards. SHELDUS provides county-level data that correspond to nearly all of the natural hazards represented by the NRI. It offers a further degree of description by identifying events by peril as well as hazard. SHELDUS aggregates property damage, crop losses, injuries, and fatalities due to a peril by month, year, and county since 1960. Most of this data, at the event level, were collected by NOAA and published in the monthly Storm Data and Unusual Weather Phenomena report, though information for some hazards is extracted from additional resources.

SHELDUS represents the best available data on economic, population, and agricultural losses due to natural hazards. However, there are many cases where the geographic precision of the recorded loss is imperfect. In these cases, the exact location of injuries and fatalities may be unknown due to regional reporting from the source data interpreted by SHELDUS, often based on a forecast zone that

[14] Center for Emergency Management and Homeland Security, Arizona State University. (2017). Spatial Hazard Events and Losses Database for the United States, Version 16.0. [online database]. Retrieved from https://cemhs.asu.edu/sheldus

covers multiple counties. For example, in Table 6, an Ice Storm injury is recorded as 0.5 for two neighboring counties and both have the same level of property damage. This signifies that the precise location of the damage associated with this event could not be determined between the two counties, so the damage is split evenly between them. The NRI utilizes SHELDUS data as it is compiled and does nothing to alter the source information.

Table 6: Sample SHELDUS Data, Aggregated by Peril, County, and Year-Month

County FIPS	Year	Month	Peril	Number of Records	Duration Days	Crop Damage (2016 $)	Property Damage (2016 $)	Injuries	Fatali-ties
01001	1996	4	Hail	1	1	3,115.02	18,690.11	0	0
32003	1996	6	WindVortex	2	1	0	7,787.55	0	1
05007	2009	1	Ice	1	3	0	17,643,421.41	0.5	0
05143	2009	1	Ice	1	3	0	17,643,421.41	0.5	0

Data were downloaded at the peril level, aggregated to a county-month level, and mapped via a control table in the NRI processing database to the appropriate NRI-defined natural hazards. Peril data were downloaded because natural hazard types as defined in SHELDUS do not directly map into the natural hazard definitions utilized in the NRI. For example, SHELDUS classifies all flooding perils under the hazard Flood while the NRI explores two flooding hazards (Coastal and Riverine) and classifies the different flooding perils accordingly (see Table 7).

Table 7: NRI Hazard to SHELDUS Peril Mapping

NRI Hazard	SHELDUS Perils
Avalanche	Avalanche, AvalancheDebris, AvalancheSnow, SnowSlide
Coastal Flooding	Coastal, CoastalStorm, FloodCoastal, FloodTidal
Drought	Drought
Earthquake	Earthquake, Fire-following Earthquake, LandslideFollowingEQ, Liquefaction
Hail	Hail
Heat Wave	Heat, HeatWave

NRI Hazard	SHELDUS Perils
Hurricane	CycloneExtratropical, CycloneSubtropical, CycloneUnspecified, HurricaneTropicalStorm, NorEaster, StormSurge, TropicalDepression, TropicalStorm
Ice	Ice Storm
Landslide	Landslide, LandslideSlump, MudFlow, Mudslide, RockSlide
Lightning	FireStElmos, Lightning
Riverine Flooding	FloodFlash, FloodIceJam, Flooding, FloodLakeshore, FloodLowland, FloodRiverine, FloodSmallStream, FloodSnowmelt
Strong Wind	Derecho, Wind, WindStraightLine
Tornado	FireTornado, Tornado, Waterspout, WindTornadic, WindVortex,
Tsunami	Tsunami, TsunamiSeiche
Volcanic Activity	Ashfall, Lahar, LavaFlow, PyroclasticFlow, Vog, Volcano
Wildfire	FireBrush, FireBush, FireForest, FireGrass, Wildfire
Winter Weather	Blizzard, StormWinter, WinterWeather

5.4.2. SELECTING SOURCE DATA: NWS STORM EVENTS DATABASE

National Weather Service, Storm Events Database[15]

Unlike the other natural hazards included in the NRI, the loss information for Cold Wave is derived from the NWS's Storm Events Database. Loss data for property damage and crop damage is recorded in the same manner as the SHELDUS data, much of which originates from the Storm Events Database. Unlike SHELDUS, the Storm Events Database includes natural hazard events with no reported loss.

Dollar amounts in the Storm Events Database are not inflation-adjusted, so these were converted to 2016 dollars using the Bureau of Labor Statistics Consumer Price Index[16] to correspond with the SHELDUS inflation-adjusted dollar amounts, using Equation 10.

[15] National Weather Service. (2017). *Storm Events Database, Version 3.0.* [online database]. Retrieved from https://www.ncdc.noaa.gov/stormevents/

[16] Bureau of Labor Statistics. (2019). Consumer Price Index for all urban consumers [online dataset]. Retrieved from https://www.bls.gov/data/

Equation 10: Conversion to 2016 Dollars

$$V_{Mo_{2016}} = V_{Orig} \times \frac{CPI_{Mo_{2016}}}{CPI_{Mo_{Year}}}$$

where:

$V_{Mo_{2016}}$ is the dollar value in 2016 dollars

V_{Orig} is the original dollar value (assumed dollar value at the time of the loss event)

$CPI_{Mo_{2016}}$ is the Consumer Price Index for the month of the loss event in 2016

$CPI_{Mo_{Year}}$ is the Consumer Price Index for the month/year of the loss event

Some loss records in the Storm Events Database are designated with a forecast zone rather than a county, so each must be joined to a county via a county-zone correlation table with data that is also provided by the NWS. Cold Wave events also have beginning and end dates recorded, so the number of event-days can be computed. Cold Wave events extracted from the Storm Events Database use the same date range as most of the data utilized from SHELDUS, 1/1/1995 to 12/31/2016. The resulting extracted records mimic the structure of the SHELDUS data in that all records are aggregated by county, peril, year, and month.

5.4.3. CONSEQUENCE TYPES

The consequence types in the loss data sources are treated as direct corollaries to consequence types measured for NRI Hazard exposure.

Property

Property loss is defined as the SHELDUS -- or NWS -- reported damage to property caused by the hazard event in 2016 dollars. In the calculation of HLR, property loss is treated as the equivalent of building value recorded in Hazus 4.2 SP1. However, SHELDUS property damage can include other types of property, like vehicles or infrastructure, which would not be reported in the Census data used by Hazus to estimate building value. This is a caveat to consider when working with this data. SHELDUS and Hazus data remain the best available estimates of loss and value that could be utilized for the NRI.

Population

Population loss is defined as the SHELDUS -- or NWS -- reported number of fatalities and injuries caused by the hazard event. To combine fatalities and injuries for the computation of population loss value, an injury is counted as one-tenth (1/10) of a fatality.

The NWS Storm Events Database classifies injuries and fatalities as direct or indirect. For the purposes of the NRI, both direct and indirect injuries and fatalities are counted in the population loss value.

Agriculture

Agriculture loss is defined as the SHELDUS – or NWS – reported damage to crops and livestock caused by the hazard event in dollars. SHELDUS also tracks crop indemnity payments for USDA-insured crop loss, however the total crop/livestock damage value was considered to be more inclusive and the crop indemnity data is not used.

5.4.4. HISTORIC LOSS RATIO METHODOLOGY

Conceptually, the Historic Loss Ratio (HLR) is the representative percentage of a location's hazard exposure area that experiences loss due to a Hazard, or the average rate of loss associated with the occurrence of a hazard.

This could be computed as the average of the individual occurrence loss rates (referred to here as Loss Ratios per Basis). However, HLR cannot be calculated in these simple terms and be considered accurate. Many counties which have not experienced a loss-causing event during the time period captured from SHELDUS may be in close proximity to counties which share similar characteristics that have experienced loss to the hazard. For example, it may be inaccurate to say that a county's likely loss ratio to Hurricane is zero just because it has not experienced a loss-causing Hurricane event during the 22-year window, especially if it borders counties which have experienced loss to Hurricanes. A better approximation of the HLR is achieved by applying a Bayesian spatial weighting matrix to smooth the loss ratio data spatially and ensure that historic loss is represented in a rational way without allowing anomalous Hazard events to distort the data. To implement Bayesian credibility weighting, loss ratio averages and variances need to be computed for spatial groupings of national, surrounding, county and, for some hazards, regional levels. The nature of the source data requires some pre-processing within the database to ensure that all historical hazard events are included in the loss ratio calculations, including per-basis record expansion of the native SHELDUS records and the insertion of records representing hazard occurrences which did not result in economic loss. See the Limitations and Assumptions in Historic Loss Ratio Methodology section for more information.

Loss Record Expansion to per Basis Records

Native SHELDUS and NWS records represent loss aggregated on a county, year, month, and peril basis. Each row includes the number of reported loss-causing peril events for the month in the county and the total duration days of the events. For example, the January 2009 Ice Storm event in Table 8 lasted three days. The basis of Ice Storm occurrences is the event-day as this definition better captures the variability in duration for Ice Storm events. Without the resolution of knowing which event-day the damage occurred on, the loss is divided among the days so that each event-day record has an equal portion of the total loss (see Table 9). In this example, the three event-day records replace the native SHELDUS record. Similarly, a single native SHELDUS peril month record for an event-based hazard like Hurricane could describe two separate events. This native record would be replaced with two records, each representing a single event with half the loss of the native aggregated record. Because SHELDUS does not specify the amount of loss associated with each of the events, each SHELDUS record is expanded based on the occurrence basis (Number of Records for event basis and Duration Days for event-day basis) if the basis count is greater than one (see Table 5). This record count expansion process is performed because loss ratios will ultimately be computed for each event (or event-day) record. Having a record for each hazard occurrence per basis unit better supports the process of determining loss ratio averages and variance.

Table 8: Native SHELDUS Loss Records

County FIPS	Year	Month	Peril	Event Records	Duration (Days)	Crop Damage (2016 $)	Property Damage (2016 $)	Injuries	Fatalities
5007	2009	1	Ice	1	3	0	17,643,421.41	0.5	0
1097	1998	9	Hurricane Tropical Storm	2	5	681,464.65	23,749,724.65	0	1

Table 9: Expanded SHELDUS Loss Records

County	Year	Month	Peril	Native Loss Record Expanded per Basis	Crop Damage (2016 $)	Property Damage (2016 $)	Injuries	Fatalities
5007	2009	1	Ice	EventDay	0	5,881,140.47	0.1666	0
5007	2009	1	Ice	EventDay	0	5,881,140.47	0.1666	0
5007	2009	1	Ice	EventDay	0	5,881,140.47	0.1666	0
1097	1998	9	Hurricane, Tropical Storm	Event	340,732.33	11,874,862.33	0	0.5
1097	1998	9	Hurricane, Tropical Storm	Event	340,732.33	11,874,862.33	0	0.5

Loss Ratio per Basis Calculation

After this expansion of records to convert the loss data to loss per single event or event-day is performed, the Loss Ratio per Basis (LRB) is calculated for each event or event-day occurrence for each consequence type (building, population, or agriculture) according to Equation 11.

Equation 11: Loss Ratio per Basis Calculation

$$LRB_{Haz_{Co_{CnsqType}}} = \frac{Loss_{Haz_{Co_{CnsqType}}}}{HLRExposure_{Haz_{Co_{CnsqType}}}}$$

where:

$LRB_{Haz_{Co_{CnsqType}}}$ is the Loss Ratio per Basis (event or event-day) representing the ratio of loss to exposure experienced by a specific county due do the occurrence of a specific Hazard event, performed for each relevant consequence type (building, population, and agriculture)

$Loss_{Haz_{Co_{CnsqType}}}$ is the Loss (by consequence type) experienced from the Hazard event (or event day) documented to have occurred in the county (in dollars)

$HLRExposure_{Haz_{Co_{CnsqType}}}$ is the total value (by consequence type) estimated to have been exposed to the event or event-day Hazard occurrence (in dollars)

The definition of the HLR exposure variable in the LRB formula does not always match the definition of the exposure component utilized in the EAL formula. For hazards which can occur almost anywhere or affect large geographic areas, the HLR exposure is the entire county's building, population, or agriculture value. Hazards which only occur in certain susceptible areas, such as floodplains and tsunami inundation zones, use the HLR exposure value associated with those areas. Tornado HLR exposure is defined by the area footprint of specific historical Tornado paths. Avalanche is a unique case which requires the use of default exposure values. The HLR exposure types utilized for each hazard can be seen in the table below. Specific methods of determining HLR exposure in the LRB calculation can be found in the HLR section for each hazard. Table 10 lists the exposure types used in each hazard's LRB calculation.

Table 10: HLR Exposure Types Used in Loss Ratio per Basis Calculation

Natural Hazard	HLR Exposure Type Used in Loss Ratio per Basis Calculation
Avalanche	Default Value
Coastal Flooding	Value Defined by Hazard Intersect
Cold Wave	Total County Value
Drought	Total County Value
Earthquake	Total County Value
Hail	Total County Value

Natural Hazard	HLR Exposure Type Used in Loss Ratio per Basis Calculation
Heat Wave	Total County Value
Hurricane	Total County Value
Ice Storm	Total County Value
Landslide	Value Defined by Hazard Intersect
Lightning	Total County Value
Riverine Flooding	Value Defined by Hazard Intersect
Strong Wind	Total County Value
Tornado	Historical Footprint Matched to Specific SHELDUS Loss
Tsunami	Value Defined by Hazard Intersect
Volcanic Activity	Value Defined by Hazard Intersect
Wildfire	Value Defined by Hazard Intersect
Winter Weather	Total County Value

Non-Loss Causing Hazard Occurrence

Hazards may occur without resulting in recorded loss to buildings, population, or agriculture. For example, Lightning may strike with a high frequency, but have few loss-causing events. SHELDUS does not record events in which no loss was reported. In an effort to capture events that do not cause loss, a count of historic year-month events is produced from hazard source data and compared to a count of loss-producing events from SHELDUS. When the hazard historic event source records more events than SHELDUS, a number of zero-loss records are inserted into the set of Loss Ratios per Basis to make up the difference between historic events and loss-causing events from SHELDUS so that the event counts for both metrics are equal.

Computing loss ratio averages and variances without including the zero-loss records produces very different results than when they are included. For example, a county with 100 historical Lightning strikes may only have two loss-causing events, one causing $40,000 in damage to buildings and the other causing $60,000. If the building exposure value is $10M, the loss ratios for each loss-causing event would be 0.004 and 0.006, respectively. If only the LRBs for two loss-causing events were considered, the average would be 0.005. Including the 98 Lightning strikes that did not result in loss lowers the average to 0.0001, a more accurate approximation of the average Lightning strike's impact on the county as not every Lightning strike is a loss-causing event.

The output of the Loss Ratio per Basis calculation (see Equation 11) and all corrective record insertion is stored in the LRB table within the NRI processing database, and are then used to compute Bayesian metrics and calculate the weighting factors that are applied to find the Bayesian-adjusted HLR for each consequence type for the county. Table 11 illustrates the content of the LRB database table after the corrective record insertions. Notice the loss ratios for three Ice Storm event-days in one county in January of 2009. These have been expanded from a single SHELDUS record based on duration days and consequence types. Also, one zero-loss record for each relevant consequence type has been inserted to recognize an Ice Storm event-day which occurred within the county (based on the historical event source data) but resulted in no economic loss. These records can then be used to calculate loss ratio averages and variance.

Table 11: Sample Data from the Loss Ratio per Basis Table

Hazard	Peril	Basis	Year	Month	Conseq. Type	Conseq. Exposure	Conseq. Loss per Basis	Conseq. Ratio per Basis Unit	Record Type
Ice Storm	Ice	Event Day	2009	1	People	221339	0.01666667	7.53E-08	Peril Basis Expansion
Ice Storm	Ice	Event Day	2009	1	People	221339	0.01666667	7.53E-08	Peril Basis Expansion
Ice Storm	Ice	Event Day	2009	1	People	221339	0.01666667	7.53E-08	Peril Basis Expansion
Ice Storm	Ice	Event Day	2009	1	Property	2.3138E+10	5881140.47	0.00025	Peril Basis Expansion
Ice Storm	Ice	Event Day	2009	1	Property	2.3138E+10	5881140.47	0.00025	Peril Basis Expansion
Ice Storm	Ice	Event Day	2001	11	People	221339	0	0	SHELDUS Native Record
Ice Storm	Ice	Event Day	2001	11	Property	2.3138E+10	310468.525	0.0000134	SHELDUS Native Record
Ice Storm	Inserted Zero-Loss Record	Event Day			People	221339	0	0	Inserted Zero-Loss Record
Ice Storm	Inserted Zero-Loss Record	Event Day			Property	2.3138E+10	0	0	Inserted Zero-Loss Record

Bayesian Credibility

To apply Bayesian credibility weighting factors and balance Historic Loss accuracy with geographic precision in areas where small sample sizes result in volatile Historic Loss estimates, LRB averages and variance may be calculated at the level of: county, surrounding 196-by-196-km fishnet grid cell,[17] region, and national. These geographic levels define which spatial grouping (or set) of LRBs are used to calculate the average and variance values. The county level grouping includes all LRBs for the county, the surrounding grouping includes LRBs for all counties that intersect the same 196-by-196-km fishnet cell, and national includes all LRBs. The formulas in Equation 12 illustrate the computation of the loss ratio average and variance.

Equation 12: Geographic Level Consequence Ratio Average and Variance Computations

$$avgLRB_{Haz_{Level_{CnsqType}}} = \frac{\sum LRB_{Haz_{Level_{CnsqType}}}}{CountEvents_{Haz_{Level_{CnsqType}}}}$$

$$varLRB_{Haz_{Level_{CnsqType}}} = \frac{\sum(LRB_{Haz_{Level_{CnsqType}}} - avgLRB_{Haz_{Level_{CnsqType}}})^2}{CountEvents_{Haz_{Level_{CnsqType}}}}$$

where:

$avgLRB_{Haz_{Level_{CnsqType}}}$	is the Average value of all Loss Ratio per Basis (event or event-day) (LRB) records of the consequence type for the geographic level due to the Hazard
$LRB_{Haz_{Level_{CnsqType}}}$	is the Loss Ratio per Basis (event or event-day) of the consequence type within the geographic area due to the Hazard occurrence basis
$CountEvents_{Haz_{Level_{CnsqType}}}$	is the total number of records of Hazard events or event-days occurring in the geographic area (includes any non-loss causing events/event-days identified)
$varLRB_{Haz_{Level_{CnsqType}}}$	is the consequence LRB variance of the geographic level due to the Hazard

Credibility increases as a function of sample size and decreased LRB variance. In other words, the higher the credibility at a given geographic level, the higher the contribution to the location's calculated Historic Loss value. Figure 12 illustrates possible loss ratio variance in neighboring

[17] The 196-by-196 km fishnet grid cell is roughly the area of four average counties. See the Intersection section for more information on the use of the 49-by-49 km fishnet resolution to represent average county area.

counties. Weighting factors in the Bayesian credibility calculation are what determines the contribution of each geographic level to the final HLR value.

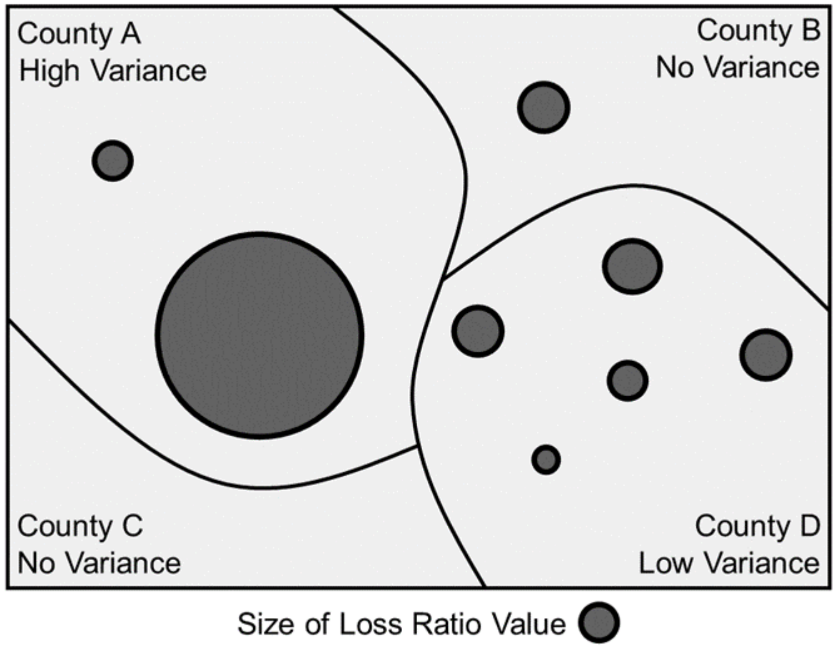

Figure 12: Example of Variance in County Loss Ratio Values

Weighting factors are derived from the variance values (calculated using Equation 12) at each geographic level according to Equation 13. For the surrounding fishnet level, if the county intersects more than one fishnet grid cell, the cell with the lowest LRB variance value is used as this provides the data with the best fit.

Equation 13: HLR Bayesian Weighting Factor Calculation

$$Wt_{Denom} = \frac{1}{varLRB_{Haz_{Ntl_{CnsqType}}}} + \frac{1}{varLRB_{Haz_{Reg_{CnsqType}}}} + \frac{1}{varLRB_{Haz_{Sur_{CnsqType}}}} + \frac{1}{varLRB_{Haz_{Co_{CnsqType}}}}$$

$$Wt_{Haz_{Ntl_{CnsqType}}} = \frac{1/varLRB_{Haz_{Ntl_{CnsqType}}}}{Wt_{Denom}}$$

$$Wt_{Haz_{Reg_{CnsqType}}} = \frac{1/varLRB_{Haz_{Reg_{CnsqType}}}}{Wt_{Denom}}$$

$$Wt_{Haz_{Sur_{CnsqType}}} = \frac{1/varLRB_{Haz_{Sur_{CnsqType}}}}{Wt_{Denom}}$$

$$Wt_{Haz_{Co_{CnsqType}}} = \frac{1/varLRB_{Haz_{Co_{CnsqType}}}}{Wt_{Denom}}$$

where:

Wt_{Denom}	is the sum of the inverted variances calculated at each geographic level, used as a denominator for the level weighting factors
$Wt_{Haz_{X_{CnsqType}}}$	is the weighting factor to be applied to the average consequence LRB for the Hazard at X level (national, regional, surrounding, county)
$varLRB_{Haz_{X_{CnsqType}}}$	is the consequence LRB variance for the Hazard at X level (national, regional, surrounding, county)

For several hazards, regional Bayesian HLR weighting supplies a more accurate estimation of historic loss for areas which have not experienced economic loss due to hazard events during the hazard's period of record. This is especially true for areas whose hazard frequency and severity are dependent on their geographic location and climate. For example, Ice Storm, Winter Weather, and Cold Wave will have a very different degree of impact on the Northeast than on the Southwest. For this reason, the Bayesian spatial weighting incorporates regional weighting rather than national for these hazards.

Most hazard-specific HLR region definitions are derived from the FEMA administrative region definitions, the only difference being that FEMA Regions I, II, and III are merged to form a region whose size is closer to that of the other regions (see Figure 13). The definition of regions for Hurricane utilizes the FEMA administrative region definitions, but further divides them into coastal regions (for the East and Gulf coasts) and inland regions along a county-level boundary which approximates the hurricane prone regions identified in the American Society of Civil Engineers (ASCE) 7-05, Minimum Design Loads for Buildings and Other Structures (see Figure 14).[18]

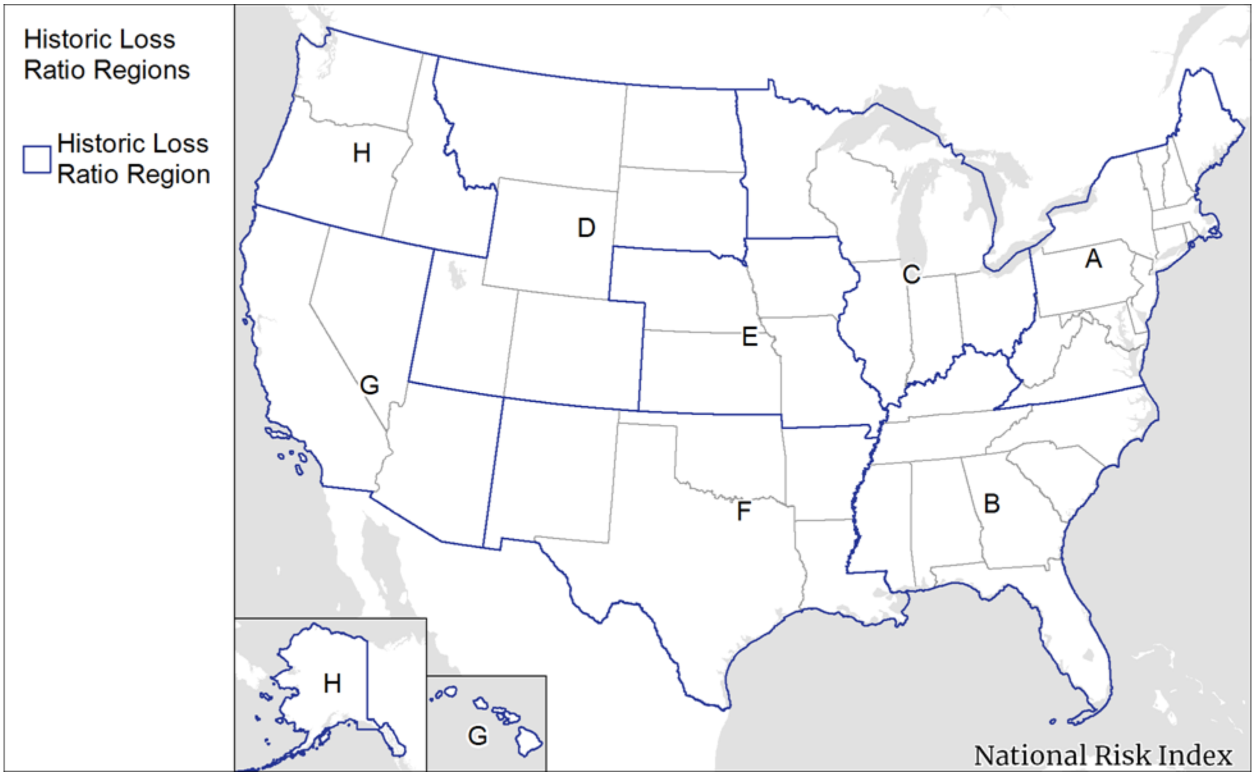

Figure 13: Historic Loss Ratio Region Definitions

[18] American Society of Civil Engineers. (2005). Minimum design loads for buildings and other structures (ASCE/SEI 7-05). Reston, VA: American Society of Civil Engineers.

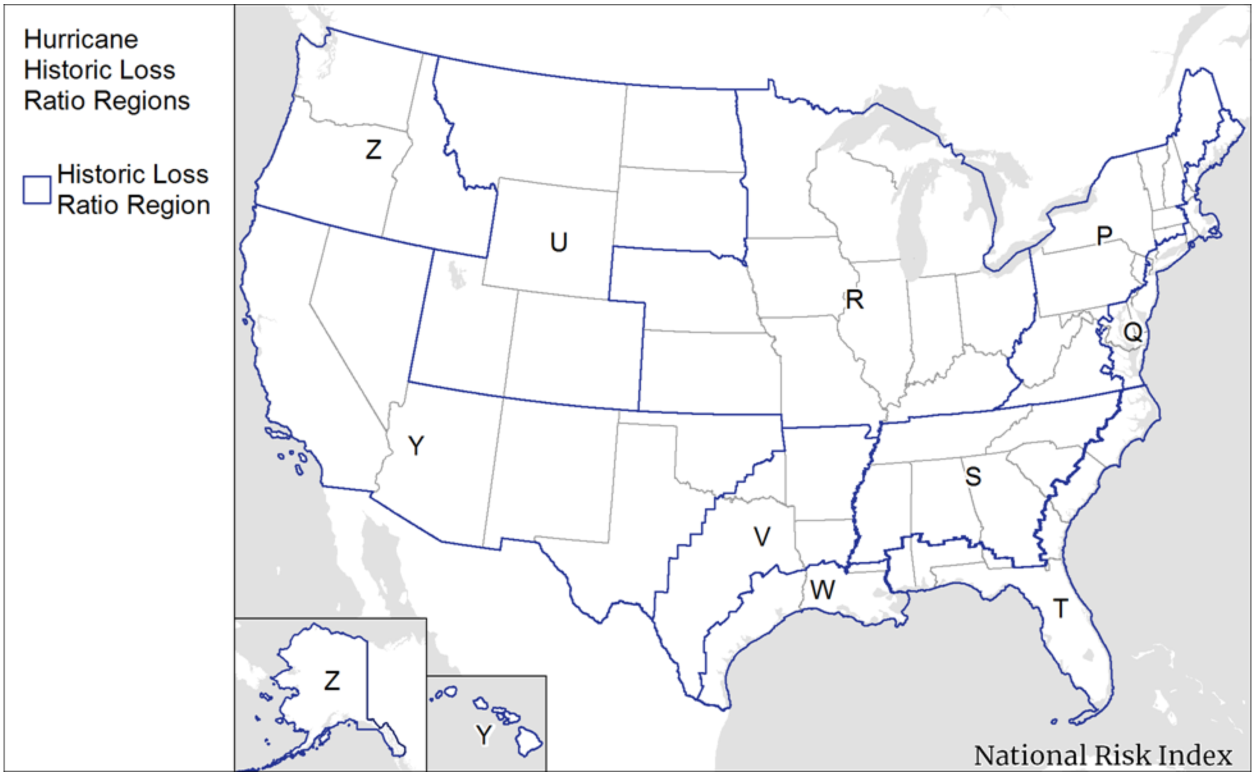

Figure 14: Hurricane Historic Loss Ratio Region Definitions

The Historic Loss Ratio for each relevant consequence type is calculated as the sum of its weighted average county, surrounding fishnet, regional, and national average LRB (see Equation 14). Levels not used for a specific hazard are removed from the computation.

Equation 14: County Bayesian-Adjusted HLR Calculation

$$\begin{aligned} HLR_{Haz\,Co_{CnsqType}} &= \left(avgLRB_{Haz\,Ntl_{CnsqType}} \times Wt_{Haz\,Ntl_{CnsqType}}\right) \\ &+ \left(avgLRB_{Haz\,Reg_{CnsqType}} \times Wt_{Haz\,Reg_{CnsqType}}\right) \\ &+ \left(avgLRB_{Haz\,Sur_{CnsqType}} \times Wt_{Haz\,Sur_{CnsqType}}\right) \\ &+ \left(avgLRB_{Haz\,Co_{CnsqType}} \times Wt_{Haz\,Co_{CnsqType}}\right) \end{aligned}$$

where:

$HLR_{Haz\,Co_{CnsqType}}$ is the Historic Loss Ratio for the Hazard at the county level, by consequence type

$avgLRB_{Haz_{X_{CnsqType}}}$ is the average LRB by consequence type for the Hazard at X level (national, regional, surrounding, county)

$Wt_{Haz_{X_{CnsqType}}}$ is the weighting factor applied to the LRB by consequence type for the Hazard at X level (national, regional, surrounding, county)

HLR Inheritance

The Bayesian-adjusted county Historic Loss Ratio is inherited by the Census blocks and Census tracts within the county when used in the NRI EAL calculations, as in Equation 15.

Equation 15: Census Tract and Census Block HLR Inheritance

$$HLR_{Haz_{Co_{CnsqType}}} = HLR_{Haz_{CT_{CnsqType}}} = HLR_{Haz_{CB_{CnsqType}}}$$

where:

$HLR_{Haz_{Co_{CnsqType}}}$ is the Bayesian-adjusted Historic Loss Ratio, a hazard-county-consequence type specific value

$HLR_{Haz_{CT_{CnsqType}}}$ is the Inherited Historic Loss Ratio for the Hazard at the Census tract level

$HLR_{Haz_{CB_{CnsqType}}}$ is the Inherited Historic Loss Ratio for the Hazard at the Census block level

5.4.5. LIMITATIONS AND ASSUMPTIONS IN HISTORIC LOSS RATIO METHODOLOGY

Several factors are not entirely accounted for in the calculation of HLR. Certain processes, such as Bayesian credibility adjustments, attempt to correct some of these limitations. This section addresses some of the assumptions that are intrinsic within the current methodology and how these can limit the accuracy of the calculation.

Evaluating historic economic loss from SHELDUS over a relatively brief period of time and comparing it to a static HLR exposure value does not account for changes in development patterns over these years. For example, a hazard event in 1995 may have a low HLR when its loss is compared to its 2010 Hazus-derived exposure value, though because of increased development and population influx over the years, its HLR would be much higher if the same loss were compared to the actual 1995 exposure value. There is an inherent assumption in the methodology that all buildings, population, and agriculture exposed to the hazard are static in economic value and quantity over the data period. Additionally, the SHELDUS loss values are inflation-adjusted to 2016 dollars while

Hazus-derived exposure values are in 2018 dollars based on 2010 valuations and there is an assumption that these dollar values are comparable.

Since the HLR calculation is based on historical events, it does not project reductions due to enhanced mitigation efforts and improved building standards that have changed over time (i.e., a seawall being built after a destructive flooding event may reduce the damage caused by subsequent flooding events).

Characterizing agriculture losses from events is highly complex and can vary based on a number of factors, including supply and demand, substitution effects, crop rotation, and seasonality. The simplified HLR calculations use crop and livestock distribution and values based on agriculture data from CropScape and the Census of Agriculture.

There are many cases where the geographic precision of the recorded loss is imperfectly captured in SHELDUS. The regional reporting data used to compile SHELDUS may mention multiple counties for a loss-causing event. In these cases, the loss is spread equally over the counties where the hazard occurred, though the loss may have only occurred in one county. Also, loss may only occur in a portion of the county, yet the HLR will apply to the entire county due to loss not being recorded with any granularity below the county level.

5.5. Validating Expected Annual Loss Estimates to Historical Losses

The diversity of the hazards and source data included in the calculation of the NRI presents a significant challenge to provide accurate and meaningful results for the variety of potential lenses through which the results may be viewed, such as:

- Hazard EAL rankings within a county;
- County EAL rankings within a hazard;
- County EAL rankings across all hazards;
- Hazard EAL rankings all counties.

As an attempt to validate the EAL, historic loss from SHELDUS for the period from 1995 to 2016 was aggregated for the entire nation for each hazard and divided by the period of record (22 years) to give a rough nationwide hazard annualized loss estimate.[19] This value was compared to the aggregated EAL estimate calculated for the NRI for its corresponding hazard. All but two (Earthquake and Volcanic Activity) of the natural hazard EALs are within the same order of magnitude as the experienced historic losses and half of the hazards are within a factor of 2.

When evaluating the historical record, losses for some hazards are driven by relatively few events. For example, from 1960 to 2016, 50% of all hurricane consequences were caused by only 8 storms. Similarly, from 1995 to 2016, 50% of all riverine flooding consequences were caused by only 48 events. The same pattern applies to Earthquakes and Volcanic Activity. These events are statistical

[19] For Cold Wave, the historic loss data was aggregated from the NWS Storm Events Database for 1996 to 2016 and divided by the 21-year period of record.

outliers where high-value urban areas have been impacted by severe hazard events. For Wildfire and Earthquake, annualized frequency uses probabilistic statistics to compute an annualized frequency. Use of probabilistic data to calculate EALs for these hazards account for the probability that the outlier event may occur. Reliance on historical data alone for the other hazards will generally underestimate the EALs for hazards where losses are driven by the rare catastrophic events. For this reason, Hurricane EALs are significantly lower (~75%) than their historical losses. This is because, for every severe hurricane that directly strikes a major city, there may be dozens of glancing blows from minor hurricanes or tropical storms that cause minimal damage. The HLR approach calculates an average value; so, HLRs are weighted toward the more common, lower loss events rather than the rare catastrophic events.

Despite these outliers, a relatively high level of agreement between the NRI-calculated EAL and the historical loss records serves as an indication that the NRI estimated annual hazard loss is fairly aligned with actual recorded historic loss.

6. Using the National Risk Index

The NRI is available to the public through https://www.fema.gov/nri. FEMA provides access to the NRI data and information through multiple venues, including a website, an interactive map and data exploration tool, tabular and spatial dataset files, and GIS-based REST services.

6.1. The NRI Website

The NRI website is the hub of access to the vast array of information in the National Risk Index. The website provides an overview of the NRI and links to documentation with important details about the source data and source data providers, descriptions of the methodology, and guidance on interpreting the results. With the interactive mapping tool, users can visually explore components of the NRI dataset and then delve into any location and examine its risk factors.

6.2. Downloadable and Online Datasets

File-based versions of the NRI dataset can be retrieved from the Data Download feature of the NRI website. Tabular and spatial formats are provided for both Census tract-level and county-level datasets. Tabular data, provided as CSV files, can be used in a wide variety of applications, and shapefiles are available for spatial applications.

The NRI dataset can also be used from web services that are hosted in the FEMA Hazards Geoplatform and accessible through ArcGIS Online. These services are a convenient way to explore the data with online tools other than the NRI website, and developers can leverage the REST services to integrate NRI data into their own applications.