Technical Methodology: Hazard Explorer

Hazard Data Methodology

Ref: Hazard Explorer: Technical Methodology
Date: 2026-02-28
Version: v1.0

Prepared by:
D.J. Rasmussen, PhD
Degree Day, LLC
dj.rasmussen@degreeday.org
https://www.degreeday.org

Degree Day

This document describes the data sources and methods underlying the Degree Day Hazard Explorer.

Contents

1 Overview

The Hazard Explorer provides a consistent framework for assessing the exposure of assets worldwide to both climate-related and geological hazards. Included perils span flooding, extreme heat, drought, wildfire, extreme winds, hail, earthquakes, landslides, and land subsidence. Rather than developing all datasets from scratch, the methodology curates and integrates dozens of peer-reviewed, globally recognized datasets to ensure that information is scientifically credible and transparently documented. These data sources include historical weather observations, reanalysis datasets, land cover and elevation data, and specialized models for hazards such as flooding and drought. Data are typically represented on global grids, in some cases down to 90 m resolution for flood inundation.

The **Hazard Explorer** presents these data through an interactive mapping interface for global hazard screening. Each hazard layer is scored on a 1–10 scale with accompanying hazard categories. The application is designed for exposure screening only; it does not model vulnerability, loss, or resilience.

Table 1 summarizes the hazard metrics presented in the Hazard Explorer, including the physical quantity, statistical basis, and units for each layer.

Table 1: Summary of hazard metrics in the Hazard Explorer.

Hazard Layer	Metric	Units (SI)
Wildfire	Annual burn probability	% per year
Extreme Heat	Annual max. temperature, 10-year return level	°C
Human Heat Stress	Annual max. wet-bulb temperature, 10-year return level	°C
Human Cold Stress	Annual min. temperature, 10-year return level	°C
Drought	Composite drought hazard score	Index (integer)
River & Coastal Flood	Total inundated land area	km^2
Surface Water Flooding	Annual max. daily precipitation, 100-year return level	mm
Extreme Wind	3-second wind gust, 1 000-year return level	km/h
Lightning	Annual thunder hours	hours/yr
Large Hail	Annual probability of large hail day ($>$5 cm)	% per year
Tornado (U.S.)	Annual probability of EF2+ tornado	% per year
Earthquake	Peak ground acceleration, 475-year return period (rock site)	g
Land Subsidence	Land subsidence rate	mm/yr
Landslide (U.S.)	Susceptibility score	Index (integer)
Landslide (Global)	Susceptibility score	Index (integer)
Heating Demand	Annual heating degree days (base 18.3 °C)	°C-day
Cooling Demand	Annual cooling degree days (base 18.3 °C)	°C-day

Table 2 describes the practical relevance of each metric and the rationale for the statistical basis or return period chosen.

Table 2: Metric relevance and rationale.

Hazard Layer	Relevance
Wildfire	Annual burn probability quantifies the likelihood that a location experiences wildfire in any given year, informing land-use planning, insurance underwriting, and defensible-space requirements.
Extreme Heat	The 10-year return level of annual maximum temperature captures rare but realistic hot extremes that stress infrastructure (pavement buckling, transformer failures, power-grid strain) and affect outdoor worker safety and productivity.
Human Heat Stress	Wet-bulb temperature integrates heat and humidity, reflecting physiological heat-stress limits more accurately than air temperature alone. A 10-year return level identifies locations where dangerous heat-stress episodes affect outdoor worker safety and productivity.
Human Cold Stress	The 10-year return level of annual minimum temperature captures cold extremes that drive pipe bursting, heating-demand spikes, and hypothermia risk. This return period is relevant for outdoor worker safety and mechanical-system design.
Drought	A composite drought hazard score integrates meteorological, hydrological, and agricultural drought dimensions into a single index, supporting water-supply planning and supply-chain resilience analysis.
River & Coastal Flood	Flooded land area directly quantifies the spatial extent of potential inundation at a location, making it well suited for screening property exposure and prioritizing sites for detailed hydraulic study.
Surface Water Flooding	The 100-year daily rainfall depth indicates the intensity of rare precipitation events that can overwhelm drainage systems.
Extreme Wind	The 3-second gust is the standard metric for structural wind loading in building codes worldwide. The 1 000-year return period aligns with design standards for critical and essential facilities.
Lightning	Annual thunder hours indicate the frequency of thunderstorm activity, relevant for designing electrical-system protection, assessing outdoor-activity safety, and evaluating wildfire ignition risk from lightning strikes.
Large Hail	Hailstones exceeding 5 cm can cause significant damage to roofs, facades, vehicles, solar panels, and crops. The annual probability of a large-hail day informs material selection, insurance pricing, and maintenance planning.
Tornado (U.S.)	EF2+ tornadoes generate winds exceeding 180 km/h, capable of significant structural damage. The annual probability informs safe-room requirements, shelter planning, and insurance risk. Coverage is limited to the contiguous United States.
Earthquake	Peak ground acceleration on rock is the standard seismic-hazard metric used in building codes worldwide. The 475-year return period corresponds to a 10 % probability of exceedance in 50 years, the most widely adopted seismic design level.
Land Subsidence	Gradual ground subsidence damages foundations, buried pipelines, roads, aviation runways, and rail infrastructure over time. The annual rate indicates long-term structural and maintenance risk.
Landslide (U.S.)	A susceptibility score reflects terrain vulnerability to landslides based on slope, geology, soil properties, and land cover, supporting development siting and slope-stability screening.

Continued on next page

Table 2 continued

Hazard Layer	Relevance
Landslide (Global)	Provides the same landslide susceptibility information as the U.S. layer but with global coverage, enabling consistent screening of international portfolios.
Heating Demand	Heating degree days quantify cumulative cold exposure that drives space-heating energy consumption. Higher values indicate greater heating costs, fuel demand, and mechanical-system sizing requirements.
Cooling Demand	Cooling degree days quantify cumulative warm exposure that drives space-cooling energy consumption. Higher values indicate greater cooling costs, peak electricity demand, and refrigeration infrastructure needs.

For each hazard, notable advantages and limitations are summarized. Because appropriate use cases vary widely, expert judgment is required to interpret results in the context of a specific application. General considerations and limitations are provided at the end of this document.

2 Flooding

Flood hazard is assessed from both inland and coastal sources. Inland flooding encompasses riverine (fluvial) and stormwater (pluvial) mechanisms, while coastal flooding captures sea-level rise, tides, and storm surge. Each is evaluated separately due to distinct causal drivers, physical mechanisms, and modeling approaches.

2.1 Inland Flooding

Inland flooding is assessed by analyzing its two primary drivers: riverine and stormwater flood hazards.

2.1.1 Riverine Flood

Riverine flooding occurs when rivers, streams, or other natural channels overflow their banks due to excessive water input. The riverine (fluvial) flood hazard layers are taken from a published global dataset of riverine flood hazard at 3 arcsecond resolution (∼90 m; Alfieri et al 2023). The dataset was developed using a sequential chain of meteorological, hydrological, and hydraulic models, each building on the outputs of the previous stage. Hydrological processes are simulated using a global hydrological model that resolves runoff, soil moisture, and snow dynamics through a geomorphological approach informed by digital elevation data. Climate inputs are taken from a bias-adjusted version of ERA5. Model calibration uses observed discharge, soil moisture, and evapotranspiration data through a multi-site optimization procedure that minimizes a cost function across locations. Hydraulic modeling applies a 1D flow model using cross-sections derived every kilometer from a global digital elevation model and hydrography map and computes inundation using the Manning equation. Flood defenses, such as levees, are not represented. As a result, the riverine flood inundation extents represent undefended flood hazard.

Advantages: The modeling framework is both global and consistent, facilitating comparisons across regions.
Limitations: The model is not locally calibrated, defenses are not incorporated, and the simple, 1D hydraulic approach does not explicitly model surface water spreading.

2.1.2 Stormwater Flood

Stormwater (pluvial) flooding occurs when rainfall overwhelms local drainage systems and water ponds on streets, parking lots, and other developed surfaces. Unlike river flooding, stormwater flooding does not require a nearby stream or river. It is especially common in urban areas where pavement and rooftops prevent water from soaking into the ground. Short-duration, high-intensity rainfall events are the primary driver of this hazard.

Because it is not feasible to simulate drainage systems and local runoff processes at global scale, flood depth is not modeled directly. Instead, stormwater hazard is assessed based on the frequency and magnitude of extreme rainfall.

2.1.3 Extreme Rainfall Methodology

Extreme precipitation statistics are estimated using the Simplified Metastatistical Extreme Value (SMEV) framework of Hoch et al (2025). SMEV is an advanced extreme value method designed specifically for rainfall. Rather than relying only on the single largest storm each year (the traditional approach), SMEV evaluates all statistically independent rainfall events. This allows both how often storms occur and how intense they are to be modeled together.

In this implementation, SMEV is applied to observed daily precipitation totals. Independent rainfall events are identified using an event-separation criterion consistent with Hoch et al (2025). From these events, model parameters are estimated and used to calculate rainfall amounts associated with specific return periods (for example, the rainfall expected on average once every 10 or 100 years).

To produce spatially continuous global coverage, SMEV parameters are regionalized to a $0.1^\circ \times 0.1^\circ$ grid using the bottom-up regionalization and machine-learning framework described in Hoch et al (2025). This approach uses precipitation climatology, atmospheric variables, elevation, and climate classification information to estimate extreme rainfall behavior in locations without dense gauge coverage.

Advantages: Grounded in observed rainfall data; uses a rainfall-specific extreme value framework that incorporates both storm frequency and intensity; provides globally consistent return-period estimates suitable for screening and comparative risk analysis.

Limitations: Results represent rainfall hazard only and do not simulate runoff generation, drainage system performance, or flood depth. The analysis uses daily precipitation totals, which aggregate rainfall over 24 hours and do not resolve the sub-hourly intensity peaks (typically 15–60 min) that are the primary driver of urban drainage failures; consequently, the metric may underestimate stormwater flood hazard where short-duration bursts dominate.

2.2 Coastal Flooding

Coastal flooding is assessed using a model that employs multiple datasets to estimate the effects of tides and storm surge from tropical cyclones and other coastal storms on coastal inundation (Deltares, 2021). For the expected magnitude of storm surge and tides, the Global Tide and Surge Model (GTSM) is used (Muis et al, 2020). GTSM is a depth-averaged global hydrodynamic model built using the DELFT-3D hydrodynamic model. The model was run with a coastal resolution of 2.5 km (1.25 km in Europe) and a deep-ocean resolution of 25 km. The GTSM-ERA5 dataset spans 1979–2018 and was developed by forcing GTSM with hourly ERA5 wind speeds and atmospheric pressure at 10 m (Hersbach et al, 2020). GTSM-ERA5 has a 10-minute temporal resolution and provides a time series at locations approximately every 50 km along the coastline (10 km in Europe). Validation carried out by Muis et al (2020) shows that the dataset performs well against observations of annual maximum water level. Inundation modeling is done with a non-hydrodynamic and non-mass-conservative modified planar flood model that uses a constant energy dissipation term to reduce storm surge and tides linearly with inland propagation (Vafeidis et al, 2019). Coastal elevation is taken from the NASADEM digital elevation model ($\sim$90 m). Coastal flood defenses, such as levees and storm surge barriers, are not represented. As a result, the coastal inundation extents represent undefended coastal flood hazard.

Advantages: Globally consistent, incorporates multiple drivers of coastal floods (sea-level rise, tides, and storm surge).

Limitations: GTSM tends to underestimate extreme storm surge magnitudes in regions affected by intense tropical cyclones, which can lead to a systematic low bias in coastal flood extent in such areas. The planar inundation approach simplifies complex hydraulics and is not mass-conservative. Local compound flood dynamics (e.g., river–coast interactions) may be under-resolved. Reliance on NASADEM ($\sim$90 m) introduces elevation-dependent bias: in low-lying coastal terrain, small vertical errors in the DEM can substantially alter modeled inundation boundaries, and known DEM artefacts in vegetated or urban areas may further degrade accuracy.

3 Temperature and Energy Demand

This section describes methods for assessing extreme temperature hazards and their implications for building energy demand.

3.1 Human Cold and Heat Stress

Extreme temperature hazards are assessed using a bottom-up regionalization framework conceptually consistent with Hoch et al (2025), adapted for temperature-based extremes. The analysis focuses on two metrics relevant for infrastructure performance and human health: (1) extreme cold temperature and (2) extreme wet-bulb temperature.

Observational Data and Quality Control

All model fitting is based on observed station data from NOAA's Global Historical Climatology Network Daily (GHCNd) and Global Historical Climatology Network Hourly (GHCNh) archives. These datasets provide long-term surface air temperature observations from meteorological stations worldwide.

Because extreme value analysis is highly sensitive to data quality, observations were subjected to an extensive quality control (QC) process prior to modeling. The QC procedure included:

- Removal of flagged or suspect observations;

- Physical range and internal consistency checks;
- Homogeneity screening for discontinuities;
- Minimum record-length requirements;
- Filtering of stations with excessive missing data or unrealistic variance.

Wet-bulb temperature was calculated from cleaned air temperature and humidity observations using physically consistent psychrometric relationships.

Extreme Value Framework

Extreme cold and extreme wet-bulb temperatures were modeled using the Gumbel distribution, a special case of the Generalized Extreme Value (GEV) distribution with shape parameter $\xi = 0$.

Preliminary fitting using the full GEV distribution showed that estimated shape parameters were generally very close to zero across stations. The Gumbel formulation therefore provides a parsimonious and numerically stable representation of tail behavior while reducing parameter uncertainty.

Parameter estimation was performed within a non-stationary Bayesian framework. In this formulation, the location and scale parameters are allowed to vary as a function of global mean surface temperature, improving robustness in the presence of temporal trends driven by large-scale climate variability. Bayesian inference yields full posterior distributions for return levels, enabling uncertainty quantification rather than single-point estimates.

For each station, annual block maxima (for wet-bulb) or minima (for cold extremes) were extracted and used to estimate return levels associated with specified recurrence intervals (e.g., 10-year, 50-year, 100-year events).

Spatial Regionalization

To generate spatially continuous global fields, station-derived Gumbel parameters were regionalized using a machine learning framework consistent with the bottom-up philosophy of Hoch et al (2025). Rather than interpolating return levels directly, the model predicts distribution parameters at ungauged locations.

Predictor variables include:

- CHELSA high-resolution (30 arcsecond; ~1 km) climatologies;
- Elevation and terrain metrics;
- Latitude and large-scale climate classification variables;
- Additional atmospheric and moisture-related predictors relevant to wet-bulb behavior.

The trained machine learning model estimates Gumbel distribution parameters on a 30 arcsecond (~1 km) grid, enabling calculation of return levels at ungauged locations globally.

Advantages: Grounded in observed station data; uses a parsimonious extreme value formulation supported by empirical diagnostics; employs Bayesian inference to quantify uncertainty; produces high-resolution (~1 km) global hazard estimates; explicitly models wet-bulb temperature to capture combined heat and humidity stress.

Limitations: Results depend on station density and record length. Wet-bulb calculations rely on the accuracy of humidity observations. Regionalization assumes that predictor variables adequately capture spatial controls on extremes. The analysis characterizes meteorological extremes only and does not model exposure, vulnerability, or infrastructure response.

3.2 Energy Demand

Building-related energy demand is assessed using Heating Degree Days (HDD) and Cooling Degree Days (CDD), calculated from daily near-surface air temperature fields derived from CHELSA (Climatologies at High Resolution for the Earth's Land Surface Areas) at 30 arcsecond (~1 km) spatial resolution for the period 2004–2023.

Methodology

CHELSA is a high-resolution global climate dataset that provides gridded temperature, precipitation, and derived bioclimatic variables. It is produced using a quasi-mechanistic statistical downscaling approach that incorporates orographic precipitation processes, wind fields, valley exposure, and boundary-layer dynamics to improve representation of terrain-driven climate variability. Unlike simple interpolation products, CHELSA explicitly accounts for topographic influences on precipitation formation and temperature gradients, producing spatially consistent climate fields at approximately 1 km resolution globally. This terrain-informed structure improves representation of mountainous regions, coastal gradients, and other areas where coarse-resolution reanalysis products often smooth critical spatial variability.

Cooling Degree Days (CDD) assume that no mechanical cooling is required when outdoor temperature is below 18.3 °C (65 °F). For each day, CDD are defined as the positive difference between the mean daily temperature (average of daily maximum and minimum temperature) and 18.3 °C; if the mean temperature is below this threshold, no CDD are recorded.

Heating Degree Days (HDD) assume that no mechanical heating is required when outdoor temperature exceeds 18.3 °C (65 °F). For each day, HDD are defined as the positive difference between 18.3 °C and the mean daily temperature; if the mean temperature exceeds this threshold, no HDD are recorded.

The 18.3 °C (65 °F) baseline represents a widely used balance-point temperature at which many buildings transition between heating and cooling demand. While actual balance points vary by building design, insulation levels, internal gains, and occupancy patterns, this standardized threshold enables consistent spatial comparison across regions.

Daily HDD and CDD values are aggregated to monthly and annual totals to characterize climatological energy demand patterns and interannual variability.

Advantages: High spatial resolution ($\sim$1 km) enables terrain-informed assessment of heating and cooling demand; standardized balance-point methodology ensures comparability across regions; daily resolution preserves variability relevant for seasonal and annual demand characterization.

Limitations: Degree-day metrics approximate temperature-driven demand only and do not explicitly account for humidity, solar gains, building efficiency, electrification trends, or behavioral adaptation. The fixed 18.3 °C balance point may not represent all building types or climate regimes.

4 Wildfire

Estimating current and future wildfire hazard is inherently complex due to the many interacting drivers of fire behavior. These include atmospheric conditions conducive to burning, ignition sources, fuel availability (both natural vegetation and human-origin materials such as structures), and the effects of fire suppression or vegetation management. Wildfires are also resource-depleting phenomena: once fuel is consumed, additional fire cannot occur at that location until regeneration occurs. Modeling fire spread introduces further complexity, and limited observations of large, high-intensity wildfires constrain calibration and validation efforts. As a result, explicitly modeling global wildfire hazard remains subject to significant uncertainty.

Methodology

The wildfire hazard model is a machine learning framework built upon high-resolution (30 arcsecond, $\sim$1 km) global environmental predictors. The methodology integrates 19 spatially explicit variables designed to characterize climatic forcing, fuel aridity, surface conditions, and ecological context.

The climatic foundation is derived from CHELSA V2.1 bioclimatic variables (see Section 3.2 for a description of CHELSA), including annual precipitation totals, minimum monthly precipitation, driest-quarter averages, temperature seasonality, and mean temperatures of the warmest and driest quarters.

To represent fuel moisture dynamics, the model incorporates Annual Maximum Potential Evapotranspiration and the Annual Minimum Monthly Climatic Moisture Index (CMI), which captures the balance between precipitation and evaporative demand during the most moisture-limited period of the year. Surface forcing variables include 10 m wind speed from the Global Wind Atlas and snow metrics (snow cover duration and annual snow water equivalent) from CHELSA.

Ecological context is represented through Köppen–Geiger climate classifications and Global Biomes. Biomes serve as a proxy for fuel type, structure, and continuity.

Model Architecture and Training

The modeling framework employs an Extreme Gradient Boosting Regressor (XGBoost), trained on one million spatial samples across Alaska, Hawaii, and the Contiguous United States. Ground-truth hazard labels were derived from U.S. Forest Service Burn Probability datasets.

To mitigate spatial autocorrelation and ensure geographic robustness, a Spatial Block 3-fold cross-validation strategy was implemented. The domain was partitioned into 2.0° geographic blocks, with training conducted on subsets of blocks and validation performed on entirely unseen regions. This spatial validation approach substantially improved model generalization.

The transition to gradient boosting and spatial tuning increased model performance from an R^2 of 0.77 to 0.877, with a Mean Absolute Error of 0.001.

Topographic–Moisture Interaction Framework

A central innovation of this methodology is the introduction of topographic–moisture interaction terms to eliminate spatial biases in humid high-elevation regions. Interaction features were engineered between Terrain Steepness, Elevation, Wind Speed, and Temperature metrics with the Annual Minimum Monthly CMI.

Importantly, the non-interacted raw predictors were excluded from the final training set. This forces the model to learn wildfire hazard as a conditional function of fuel aridity, preventing the assignment of artificially high hazard probabilities to high-altitude but moisture-rich tropical regions.

Global Inference and Spatial Masking

Global inference is performed at 30 arcsecond resolution using a high-performance windowed processing architecture. Interaction terms and latitudinal grids are dynamically generated during inference to maintain strict feature consistency with the training dataset.

To ensure physical realism, predictions are spatially masked using Copernicus Land Use and Land Cover data. Hazard is restricted to burnable land classes, including closed forests, shrublands, herbaceous vegetation, wetlands, and tundra. Non-burnable surfaces such as open water and urban centers are excluded.

This masking process, combined with moisture-conditioned feature engineering, produces a refined global wildfire hazard surface that highlights areas of legitimate burn potential while minimizing false positives in climatically unsuitable regions.

Advantages: High-resolution (1 km) global coverage; terrain-informed climate inputs; spatial cross-validation for robustness; physically constrained inference via moisture-conditioned feature engineering; explicit masking of non-burnable land classes.

Limitations: The model is trained exclusively on U.S. burn probability data (Alaska, Hawaii, and the contiguous United States). Global application assumes that the statistical relationships between environmental predictors and burn probability learned in U.S. ecosystems transfer to other continents, but fire regimes, ignition patterns, suppression practices, and fuel structures differ substantially worldwide; transferability has not been formally validated outside the training domain. The framework does not explicitly simulate ignition frequency, human suppression, or dynamic fire spread.

5 Drought

5.1 Precipitation Drought Assessment

Meteorological drought is assessed using basin-scale precipitation deficits derived from observation-constrained reanalysis data. This approach quantifies multi-year water availability stress through percentage deviations from climatological norms and extreme value statistics to estimate design drought severity for infrastructure planning.

Methodology Overview

Precipitation drought metrics are computed at two complementary spatial scales: hydrologically-defined watersheds (HydroSHEDS Level 3 basins) where available, and first-order administrative regions (Admin 1) for complete global coverage including islands and high-latitude areas not represented in watershed databases.

Daily precipitation from ERA5 reanalysis (1940–present, 0.25° resolution), which assimilates global precipitation observations, is aggregated to 48-month accumulations to capture multi-year persistence relevant to water supply systems, reservoir storage, and hydrological drought. These accumulations are area-weighted across each basin or administrative region, then normalized as percentage anomalies (expressed as percent deviation from the long-term mean) relative to a 1991–2020 climatological baseline following WMO guidelines.

Extreme value analysis using the Generalized Extreme Value (GEV) distribution is applied to annual minima of rolling 48-month precipitation accumulations to estimate return levels for 10-, 25-, 50-, and 100-year design events. These are expressed both as percentage deficits from normal precipitation and as absolute precipitation amounts to support infrastructure design decisions.

Composite Drought Hazard Score

Drought is measured using a composite drought hazard score (0–100, where higher values indicate greater vulnerability) that is designed to answer a simple physical question:

> *After a severe drought, how much water is left in the system?*

Rather than measuring drought as a relative departure from local climatology, the score emphasizes **absolute hydrological scarcity** under a design-level drought event. This framing reflects the fact that water management failures occur when physical water buffers are exhausted, not when anomalies exceed a statistical threshold.

Remaining Water Calculation

For each spatial unit, remaining water is defined as the precipitation available after accounting for a 25-year return-period drought deficit over a multi-year accumulation window:

$$R = \overline{P}_{\text{clim}} - D_{25} \tag{1}$$

where:

- R is the remaining water (mm),
- $\overline{P}_{\text{clim}}$ is the mean climatological precipitation accumulated over the reference period, and
- D_{25} is the estimated precipitation deficit associated with the 25-year return-period drought.

This quantity represents the *"bottom-of-the-tank"* hydrological condition—the amount of precipitation expected to remain during a severe but plausible drought event relevant to long-lived infrastructure and water supply systems.

Percentile-Based Scoring

Raw remaining water values exhibit a strongly right-skewed (heavy-tailed) distribution due to the presence of humid and tropical regions with extremely high precipitation totals. Direct linear scaling of R would therefore obscure meaningful differences among water-limited regions.

To address this, remaining water is transformed into a globally comparable drought hazard score using an empirical percentile-ranking approach:

$$S = (1 - F_R(R)) \times 100 \tag{2}$$

where:

- S is the composite drought hazard score (0–100),
- $F_R(\cdot)$ is the empirical cumulative distribution function (CDF) of remaining water computed across all regions.

Under this transformation:

- Regions in the lowest percentiles of remaining water (least water left) receive scores approaching 100,
- Regions with large residual water buffers receive scores approaching 0.

This percentile-based scoring "flattens" the fat-tailed precipitation distribution into a near-uniform score distribution, increasing contrast in the critical transition zone between semi-arid and arid climates where water scarcity risk is most acute.

Interpretation

The resulting score is **relative but physically grounded**. Although the score represents a global rank rather than an absolute threshold, it is anchored in a meaningful physical quantity (remaining precipitation). As a result, identical scores correspond to equivalent levels of global water scarcity regardless of geographic location. For example, a score of 80 indicates that a region has less remaining water than approximately 80% of all regions evaluated worldwide.

Advantages: Emphasizes physical water availability rather than statistical anomaly magnitude relative to local climate (e.g., SPI and SPEI); enables consistent comparison across wet and dry climates using a common global reference; aligns naturally with infrastructure planning horizons through use of a 25-year design-level drought.

Limitations: The score reflects meteorological water availability only and does not incorporate groundwater storage, reservoir operations, interbasin transfers, or adaptive water management. Water supply demand is not considered. Spatial aggregation smooths local variability and does not capture within-region heterogeneity in drought exposure or demand. Reanalysis-based precipitation may poorly estimate precipitation climatology. The metric reflects historical design-level statistics and does not explicitly account for compound-event risk.

In practical terms, the composite drought hazard score ranks regions by how close they come to exhausting their precipitation-based water supply during a severe drought. By combining a physically interpretable scarcity metric with a globally consistent percentile-based transformation, the score provides a robust and intuitive measure of drought vulnerability.

6 Wind and Convective Storms

This section describes methods for assessing hazards related to extreme wind and convective storm activity, including tropical cyclone winds, thunderstorm frequency, severe thunderstorm environments, large hail, and tornadoes.

6.1 Extreme Wind

Annual maximum 3-second wind gust estimates are produced at 30 arcseconds ($\sim$1 km) over land and nearshore regions. Gusts capture both tropical-cyclone (TC) and non-TC drivers. TC wind fields are derived from the STORM dataset (Bloemendaal et al, 2020, 2022), based on historical records and projections. Data are smoothed and resampled to 0.25°, converted to 3-second gusts via a gust factor, and adjusted for surface roughness using ESA-CCI land cover (Thøgersen, 2021) and roughness lengths (Floors et al, 2021) via the Harris & Deaves model (Harris and Deaves, 1981).

Non-TC gusts are estimated from ERA5 (1979–2023) using the parametric gust profile methodology of van den Brink (2019). A multiple linear regression model is calibrated against in situ gust observations in Europe; predictors include squared ERA5 gust forecasts, squared shear-derived gusts (100 m to 10 m), and station elevation. The approach provides an effective parametrization of gust profiles during severe wind conditions, yielding smooth fields in flat terrain and resolving fine-scale variability in mountainous regions.

Advantages: The downscaling procedure improves local ERA5 gust estimates.
Limitations: Highly localized phenomena (downbursts, derechos, tornadic winds) are not captured. The non-TC gust regression is calibrated against European in situ observations only; application outside Europe assumes that the parametric relationship between ERA5 predictors and observed gusts generalizes to other wind climates and terrain types, which has not been independently validated. Roughness correction does not account for long-fetch decay.

6.2 Thunderstorm Frequency

Global thunderstorm frequency is assessed using a long-term, satellite-calibrated lightning detection dataset that provides climatological records from 2013 onward (Kaplan and Lau, 2021, 2022). Thunderstorm occurrence is represented using the *thunder hour* metric: any one-hour period with at least two lightning strokes detected within a 15 km radius of a location is counted as a thunder hour. These hourly data are aggregated to annual totals on a uniform global grid at $\sim$5 km resolution.

Advantages: Near-global coverage and consistent thresholds enable cross-region comparability; multi-year records emphasize robust spatial patterns while reducing sensitivity to interannual variability.
Limitations: Detection efficiency varies by region and stroke intensity; using lightning as a proxy for thunder audibility introduces uncertainty.

6.3 Severe Thunderstorm Potential

Severe thunderstorms are associated with large hail, intense rainfall, localized downbursts/microbursts, and tornadoes. Because these perils are difficult to characterize directly at large scales, indicators of storm environments are used instead. Severe thunderstorm potential is estimated using WMAXSHEAR, a composite parameter combining instability and vertical wind shear (Brooks, 2013; Taszarek et al, 2017). WMAXSHEAR effectively distinguishes between non-severe and severe environments in the U.S. and Europe (Brooks, 2013; Taszarek et al, 2017, 2019; Rodríguez and Bech, 2020).

Hourly WMAXSHEAR is calculated using ERA5 vertical profiles at 0.25° ($\sim$31 km) with 137 levels; 95th percentiles provide a globally consistent measure of severe thunderstorm potential (Taszarek et al, 2020). To ensure the parameter reflects actual storm activity, WMAXSHEAR values are conditioned on global daily lightning time series at 30 arcmin ($\sim$50 km), used as a proxy for convective initiation.

Advantages: WMAXSHEAR has been extensively evaluated and is more geographically universal than U.S.-specific parameters (STP, SHIP, SCP).
Limitations: Performance outside the U.S./Europe is less established due to limited reports; ERA5's coarse resolution requires convective parameterization, affecting vertical profiles; boundary-layer and assimilation biases can affect near-surface conditions.

6.4 Large Hail

The large hail layer is based on a global climatology of very large hail (VLH; ≥ 5 cm / 2 in), based on the additive regressive convective hazard model (AR-CHaMo) applied to ERA5 atmospheric reanalysis data (0.25°, $\sim$31 km; 3-hourly) for 1950–2023 (Battaglioli et al, 2026). The AR-CHaMo framework estimates the probability of very large hail occurrence at each grid point by combining (1) the probability of thunderstorm occurrence and (2) the conditional probability of VLH given a thunderstorm. The model uses atmospheric predictors derived from ERA5 profiles, including measures of instability, vertical wind shear, cloud-base height, low- and mid-level moisture, and other convective parameters calculated from thermodynamic soundings.

The model was trained and evaluated using lightning observations and high-quality hail reports from Europe, the United States, and Australia, and subsequently applied globally. More than 80 billion atmospheric profiles were processed to construct the long-term global VLH climatology. The resulting dataset provides spatially consistent estimates of mean annual VLH frequency, seasonal cycles, and multi-decadal trends.

Globally, the highest VLH frequencies are identified in northern Argentina and the tri-border region of Uruguay, Paraguay, and southern Brazil, followed by the U.S. Great Plains and parts of South Africa. Europe exhibits lower absolute frequencies but shows the strongest positive trends in VLH occurrence over recent decades. In contrast, portions of the Southern Hemisphere, including northern Argentina and South Africa, show decreasing VLH frequency linked to reductions in mid-level moisture and atmospheric instability (Battaglioli et al, 2026).

Advantages: Provides a 74-year globally consistent climatology (1950–2023) of very large hail derived from physically interpretable environmental predictors. Demonstrates skill in reproducing known regional hotspots, seasonal cycles, interannual variability, and observed multi-decadal trends across Europe, North America, and Australia. Enables attribution analysis linking VLH trends to thermodynamic and moisture changes.

Limitations: ERA5 does not explicitly simulate hail; VLH occurrence is inferred from environmental conditions via a statistical model. Model skill may vary in regions lacking dense observational records or where ERA5 exhibits known biases (e.g., parts of the tropics).

6.5 Tornadoes

The tornado layer provides a U.S. climatology of significant tornado activity (EF2+), derived from Storm Prediction Center (SPC) historical tornado records. The metric represents the mean number of EF2+ tornado days per decade occurring within 25 miles (40 km) of a given location, calculated over the 1986–2015 baseline period.

The use of EF2+ tornadoes focuses the analysis on events associated with substantial structural damage and insured losses. Spatial aggregation within a 25-mile radius accounts for track uncertainty and provides a location-relevant exposure metric suitable for infrastructure and portfolio screening.

The climatology highlights a core region of elevated significant tornado activity extending from the southern Great Plains into the Mid-South and lower Mississippi Valley, with secondary maxima across portions of the Midwest. The Southeast exhibits relatively high significant tornado-day frequency compared to the central Plains, reflecting differences in storm mode, population exposure, and seasonal convective environments.

Advantages: Uses an observationally based dataset with standardized EF-scale ratings. Focuses on EF2+ tornadoes, which are more consistently reported over time and most relevant for structural damage and insurance loss. The 25-mile aggregation produces a stable, location-relevant hazard screening metric.

Limitations: Tornado records are influenced by reporting practices, radar deployment history, and population density. The dataset reflects historical frequency (1986–2015) and does not explicitly account for future climate change. Spatial smoothing may underrepresent highly localized tornado corridors or mesoscale clustering effects.

7 Geohazards

7.1 Seismic Hazard

The seismic hazard layer represents peak ground acceleration (PGA) with a 10% probability of exceedance in 50 years (return period of 475 years), expressed in units of g. It is compiled as a global approximately 1 km resolution raster by merging the best available regional probabilistic seismic hazard models with the Global Seismic Hazard Assessment Program (GSHAP) map as a background layer.

7.1.1 Source Models

Where regional models exist, they take precedence over the global background. Table 3 summarizes all source datasets. For all remaining land areas not covered by a regional model, PGA values are drawn from the GSHAP global seismic hazard map (Giardini et al, 2000).

The 2023 USGS National Seismic Hazard Model covers the conterminous United States, Alaska, and Hawaii (Petersen et al, 2023). The 2020 European Seismic Hazard Model (Danciu et al, 2021) and the Earthquake Model of the Middle East (Giardini et al, 2016; Erdik et al, 2012) provide coverage across Europe and the Middle East respectively. Additional USGS models cover Puerto Rico and the U.S. Virgin Islands (Shumway et al, 2025), Haiti (Shumway, 2020), American Samoa (Shumway, 2019a), Guam and the Northern Mariana Islands (Shumway, 2019b), and South America (Petersen et al, 2018).

Further regional coverage is drawn from models distributed through the Global Earthquake Model (GEM) Foundation (Pagani et al, 2020). These include the Geoscience Australia National Seismic Hazard Assessment (Allen et al, 2023), the GNS Science New Zealand National Seismic Hazard Model (Gerstenberger et al, 2024), and regional models for the Arabian Peninsula (Sokolov et al, 2017), Caribbean and Central America, Central Asia (Ullah et al, 2015), Indonesia (Irsyam et al, 2020), Papua New Guinea (Ghasemi et al, 2020), Southeast Asia (Ornthammarath et al, 2020), Sub-Saharan Africa (Poggi et al, 2017), and Taiwan (Chan et al, 2020). South Africa is covered by the Council for Geoscience (CGS) national seismic hazard model (Midzi et al, 2020). India is covered by the probabilistic seismic hazard assessment of Nath and Thingbaijam (2012). All models assume rock or firm-rock site conditions.

Table 3: Source datasets for the merged seismic hazard layer. All models represent PGA at the 475-year return period (10% probability of exceedance in 50 years) for rock site conditions.

Region	Model	Year	Reference
Global (background)	GSHAP	2000	Giardini et al (2000)
American Samoa	USGS	2019	Shumway (2019a)
Arabian Peninsula	GEM / SGS	2018	Sokolov et al (2017)
Australia	GA NSHA	2023	Allen et al (2023)
Caribbean & C. America	GEM / CCARA	2018	Pagani et al (2020)
Central Asia	GEM / EMCA	2018	Ullah et al (2015)
Europe	ESHM20	2021	Danciu et al (2021)
Guam / NMI	USGS	2019	Shumway (2019b)
Haiti	USGS	2020	Shumway (2020)
Indonesia	GEM / PuSGeN	2017	Irsyam et al (2020)
Middle East	EMME14	2016	Giardini et al (2016)
New Zealand	GNS NSHM	2022	Gerstenberger et al (2024)
Papua New Guinea	GEM / GA	2020	Ghasemi et al (2020)
Puerto Rico / USVI	USGS PRVI	2025	Shumway et al (2025)
South Africa	CGS	2017	Midzi et al (2020)
South Asia (India)	PSHA	2012	Nath and Thingbaijam (2012)
South America	USGS	2018	Petersen et al (2018)
Southeast Asia	GEM / EOS	2018	Ornthammarath et al (2020)
Sub-Saharan Africa	GEM / SSAHARA	2018	Poggi et al (2017)
Taiwan	GEM / TEM	2020	Chan et al (2020)
United States	USGS NSHM	2023	Petersen et al (2023)

7.1.2 Processing

All source grids were extrapolated outward by approximately 50 km to fill gaps along coastlines and small islands. Sources were then resampled to a common approximately 1 km global grid. Where two or more regional models overlap, the maximum PGA value is retained. Regional values override GSHAP wherever regional data exists; GSHAP fills all remaining land areas. The output is masked to land areas using a 50 km buffered global land polygon.

This product represents a compromise between global coverage and regional accuracy. The GSHAP background provides uniform worldwide coverage but dates from 2000, while the regional models are more recent and incorporate improved seismological data and methods. Because the source models were developed independently—using different seismicity catalogues, source characterisations, ground-motion models, and computational approaches—sharp discontinuities in PGA values may exist at the boundaries between mosaicked regions.

Advantages: Globally consistent estimates expressed in probabilistic terms aligned with engineering standards; reference rock conditions ensure comparability across regions; merges the best available regional models with global coverage as a fallback.
Limitations: Local site effects (soil amplification) are not captured; secondary hazards (landslides, liquefaction, surface rupture, tsunamis, fires) are not represented; boundary discontinuities may exist between mosaicked regions; GSHAP background dates from 2000 and may not reflect current seismological understanding in gap regions.

7.2 Landslides

For regions outside the United States, landslide hazard is assessed using a global rainfall-driven dataset (Palau et al, 2023). The global model quantifies rainfall-triggered landslide hazard by combining terrain susceptibility with extreme rainfall data. Susceptibility is categorized into five classes (Very Low to Very High) based on global datasets of slope, lithology, vegetation, and soil moisture. Rainfall forcing is represented using daily intensities from historical ERA5 data. For each grid cell, rainfall thresholds are overlaid with susceptibility to estimate the conditional probability of landslide initiation. The methodology produces gridded hazard estimates at 3 arcseconds (∼90 m).

Within the conterminous United States (CONUS), Alaska, Hawaii, and Puerto Rico, landslide susceptibility is instead assessed using the high-resolution slope–relief threshold models developed by the U.S. Geological Survey (USGS) (Mirus et al, 2024). These models were trained on a national compilation of over 600,000 documented landslides and 1/3 arc-second (10 m) digital elevation models. Two parsimonious models are provided: (1) a weighted linear regression model and (2) a quantile nonlinear regression model fitted to the 10th quantile of the data. Model outputs were averaged across multiple runs and subsequently down-sampled from 10 m to 90 m resolution to account for digital elevation model and landslide positional uncertainty (Belair et al, 2022).

Susceptibility values represent the number of susceptible 10 m cells contained within each 90 m grid cell (ranging from 0 to 81), with higher values indicating greater concentration of susceptible terrain. The nonlinear (n10) model captures approximately 98.9% of inventoried landslides while limiting susceptible terrain coverage to approximately 44.6% of mapped area, providing a balance between sensitivity and spatial efficiency (Mirus et al, 2024). County-level aggregated susceptibility metrics (e.g., proportion of susceptible terrain and landslide density within susceptible terrain) are also available to support regional screening and comparative risk assessment.

Advantages: Outside the U.S., the global framework integrates rainfall forcing and terrain susceptibility within a probabilistic structure at approximately 90 m spatial resolution. Within the U.S., the USGS models provide nationally consistent, empirically trained susceptibility estimates calibrated against a comprehensive landslide inventory and high-resolution elevation data, improving reliability for infrastructure screening and planning applications.

Limitations: For the global model, rainfall forcing data remain coarse ($\sim$50 km) and may not capture localized convective extremes or short-duration intensity bursts. Terrain factors are simplified, and validation is limited in regions with sparse landslide inventories. For the USGS susceptibility maps, the models are terrain-based and do not explicitly incorporate dynamic rainfall triggering; they represent relative susceptibility rather than event-specific probability and may not reflect localized anthropogenic modifications or rapidly evolving landscape conditions.

7.3 Land Subsidence

Land subsidence refers to the vertical sinking of the land surface. Subsidence estimates are drawn from a recent global study that quantified near-present-day subsidence rates using a deep learning framework applied to $>$46,000 quality-controlled observation points (Davydzenka et al, 2023). The model used 23 predictors (including groundwater abstraction and recharge, climate, geology, and topography) and a neural network with three hidden layers (740 neurons). Model interpretability with SHAP values (Lundberg and Lee, 2017) identified groundwater abstraction as the dominant driver, followed by seismic hazard, recharge-related climate variables, and sediment thickness. The resulting dataset is provided at 30 arcseconds ($\sim$1 km).

Advantages: Provides globally consistent $\sim$1 km estimates; large, quality-controlled training data and diverse predictors improve robustness and capture key drivers.

Limitations: Local-scale subsidence can vary greatly due to factors not represented in the global predictors. The model represents near-present-day conditions and does not directly project future changes.

8 Hazard Rating Scale

Each hazard layer in the Hazard Explorer is assigned an integer rating from 0 (negligible or below detection) to 10 (extreme). Ratings provide a standardized basis for comparing relative hazard severity across layers and locations. Two scoring approaches are used depending on the physical characteristics of each metric:

- **Percentile-based (decile) scoring:** All global land-area pixels are ranked and divided into ten equal-frequency bins (deciles). A score of 1 corresponds to the lowest decile and a score of 10 to the highest. This approach is applied to metrics whose distributions are continuous and span a wide range of values, ensuring that scores reflect relative global standing regardless of the underlying unit. Some layers exclude zero or near-zero values before ranking to avoid inflating the lowest bins (e.g., heating and cooling demand).
- **Custom-bin scoring:** For metrics with well-established engineering or scientific thresholds, or where the distribution is heavily zero-inflated or categorical, score boundaries are set at physically meaningful breakpoints rather than statistical percentiles. Examples include peak ground acceleration (aligned with seismic design categories), wildfire burn probability, lightning thunder hours, and land subsidence rate.

The choice between percentile and custom-bin scoring is driven by two considerations: (1) whether the metric has widely accepted engineering or regulatory thresholds that would be obscured by a purely statistical ranking, and (2) whether the global distribution is sufficiently continuous for decile binning to produce informative score gradations. Metrics that satisfy both conditions use percentile scoring; those with established absolute thresholds or highly non-continuous distributions use custom bins.

Human Cold Stress scores are inverted: lower (more extreme) temperatures correspond to higher scores. The global landslide layer uses a fixed categorical mapping from raw susceptibility classes to hazard scores (see table footnote).

Table 4 shows the minimum metric value required for each hazard rating. Values below the score 1 threshold receive a rating of 0.

Table 4: Hazard rating scale: minimum metric value for each score level. Layers marked (P) use percentile-based scoring; others use custom bins.

Hazard Layer	**Unit**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Wildfire	%/yr	0.05	0.14	0.22	0.29	0.38	0.47	0.58	0.75	1.0	1.5
Extreme Heat (P)	°C	22.8	29.9	32.4	34.2	35.2	36.8	38.6	40.2	42.7	46.1
Human Heat Stress (P)	°C	14.9	19.9	21.6	23.3	24.4	25.4	26.8	28.3	29.2	29.7
Human Cold Stress (P)†	°C	14.0	8.1	1.4	−3.3	−10.1	−24.7	−34.2	−42.2	−46.5	−49.9
Drought (P)	score	17	27	37	44	51	60	62	71	83	91
River & Coastal Flood (P)	km^2	0.034	0.073	0.103	0.130	0.158	0.189	0.231	0.295	0.403	0.597
Surface Water Flooding (P)	mm	39.7	45.8	50.4	55.6	63.8	75.3	96.3	117.5	130.9	165.4
Extreme Wind (P)	km/h	68.1	84.3	94.9	102.9	111.1	121.1	132.1	145.0	162.4	199.4
Lightning	hr/yr	1	5	10	20	40	60	80	120	180	300
Severe Thunderstorm (WMAXSHEAR)	m^2/s^2	100	200	300	400	500	600	700	850	1 000	1 200
Large Hail	d/yr	0.0001	0.001	0.002	0.005	0.01	0.02	0.05	0.1	0.2	0.3
Tornado (EF2+)	d/yr	0.001	0.005	0.01	0.02	0.04	0.06	0.08	0.1	0.15	0.2
Earthquake	g	0.01	0.02	0.03	0.05	0.08	0.10	0.20	0.30	0.40	0.50
Land Subsidence	mm/yr	1.5	2	3	4	5	7	10	15	25	50
Landslide (U.S.) (P)	index (0–81)	0	0	1	3	6	12	20	32	46	60
Landslide (Global)‡	class (0–5)	Fixed mapping: raw class $2 \to 4$; $3 \to 7$; $4 \to 8$; $5 \to 10$. Raw class ≤ 1 = no hazard.									
Heating Demand (P)	°C-day	61	344	853	1 800	3 012	4 098	5 214	6 396	8 393	10 700
Cooling Demand (P)	°C-day	18	84	280	658	1 156	1 771	2 349	2 749	3 010	3 267

† Inverted: lower temperatures correspond to higher scores.
‡ Uses a fixed categorical mapping rather than threshold bins.

9 General Limitations and Considerations

1. **Interpretation and Intended Use:** The estimates herein provide indicative insights into the frequency and severity of natural hazards. They are intended to support high-level screening and strategic planning, not to replace site-specific engineering studies. Before applying results to planning or design, users should evaluate relevance in their local context and consult domain experts. Fitness for purpose is application- and assumption-dependent.
2. **Modeling Philosophy:** Methods aim to balance rigor and utility. While higher-resolution or specialized models may be more accurate for any single location, they often require substantial data, time, and cost. The methods used here are peer-reviewed and globally consistent, making them suitable for portfolio-scale screening where credible local data are unavailable or difficult to procure.
3. **Inherent Uncertainties:** All hazard models involve simplifying assumptions (e.g., limited resolution, imperfect observational records, parametric distributional choices). These limitations are especially relevant for localized or event-specific hazards such as convective storms, wildfires, and extreme rainfall. Users should incorporate multiple lines of evidence, including observed trends, accounts of past events, and alternative models.
4. **Relevance Over Time:** The data and methods reflect the best available understanding at publication. As observational records lengthen and methodologies improve, estimates should be revisited and updated accordingly.

10 Data Sources and Licensing

Table 5 lists all principal third-party datasets used in this product, together with their licence terms and the hazard layers in which they appear. Datasets produced by U.S. federal agencies are in the public domain under 17 U.S.C. §105 unless otherwise noted. Creative Commons (CC) licence abbreviations follow standard notation (e.g., CC BY 4.0 = Attribution 4.0 International). Where a dataset is used in multiple hazard layers, all are listed.

Table 5: Principal third-party data sources and their licence terms.

Dataset	Provider	Licence	Hazard Layer(s)
ERA5 reanalysis	ECMWF / Copernicus C3S	CC BY 4.0	Coastal flood, wind, thunderstorms, hail, drought
CHELSA V2.1 climatologies	WSL / EnviDat	CC0 1.0	Heat/cold stress, energy demand, wildfire
GHCNd (daily stations)	NOAA NCEI	Public domain	Heat/cold stress, stormwater flood
GHCNh (hourly stations)	NOAA NCEI	Public domain	Heat/cold stress
WFDE5 (bias-adjusted ERA5)	Copernicus C3S	CC BY 4.0	Riverine flood
Alfieri et al. global riverine flood maps	European Commission JRC	CC BY 4.0	Riverine flood
GTSM-ERA5 (CoDEC)	Deltares / VU Amsterdam	CC BY 4.0	Coastal flood
Deltares global flood maps	Deltares	CC BY 3.0	Coastal flood
NASADEM	NASA JPL	Public domain	Coastal flood
HydroSHEDS / HydroBASINS	WWF	Custom (free use)	Drought
STORM synthetic TC tracks	VU Amsterdam	CC BY 4.0	Extreme wind
ESA-CCI Land Cover	ESA	Custom (free for research)	Extreme wind
Global Wind Atlas	DTU / World Bank	CC BY 4.0	Wildfire, wind
WGLC lightning climatology	ARVE / WWLLN	CC BY-SA 4.0	Thunderstorms
Battaglioli et al. global hail climatology	ERA5 / ECMWF	CC BY 4.0	Large hail
SPC tornado records	NOAA SPC	Public domain	Tornadoes
USFS Burn Probability	USDA Forest Service	CC BY	Wildfire
Copernicus Land Cover 100 m	Copernicus CGLS	CC BY 4.0	Wildfire
USGS landslide susceptibility	USGS	Public domain	Landslides (US)
USGS landslide inventories	USGS	Public domain	Landslides (US)
Palau et al. global landslide model	Univ. Padova	CC BY 4.0	Landslides (global)
Davydzenka et al. global subsidence	Various	CC BY 4.0	Land subsidence
Seismic: USGS models	USGS	Public domain	Seismic hazard
Seismic: GEM regional models	GEM Foundation	CC BY-SA 4.0 / CC BY 4.0 / AGPL v3.0	Seismic hazard

Continued on next page

Table 5 continued

Dataset	Provider	Licence	Hazard Layer(s)
Seismic: ESHM20	EFEHR	CC BY-SA 3.0	Seismic hazard
Seismic: GSHAP	GSHAP	Open	Seismic hazard
Seismic: other regional	Various	Open	Seismic hazard

References

References

Alfieri L, Campo L, Gabellani S, et al (2023) The GIRI global flood hazard model. Contributing/background paper, Coalition for Disaster Resilient Infrastructure (CDRI) and UNEP/GRID, contributing/Background paper to the Global Infrastructure Risk Model & Resilience Index (GIRI)

Allen TI, Griffin JD, Clark DJ, et al (2023) The 2023 national seismic hazard assessment for Australia: Model overview. Tech. Rep. Record 2023/53, Geoscience Australia, Canberra, DOI 10.26186/148969

Battaglioli F, Taszarek M, Groenemeijer P, Púčik T, Rädler A (2026) Contrasting trends in very large hail events and related economic losses across the globe. Nature Geoscience 19:52–58, DOI 10.1038/s41561-025-01868-0

Belair GM, Jones ES, Slaughter SL, Mirus BB (2022) Landslide inventories across the United States (ver. 2.0, june 2022). U.S. Geological Survey Data Release, DOI 10.5066/P9FZUX6N

Bloemendaal N, Haigh ID, de Moel H, et al (2020) Generation of a global synthetic tropical cyclone hazard dataset using STORM. Scientific Data 7:40, DOI 10.1038/s41597-020-0381-2

Bloemendaal N, de Moel H, Martinez AB, et al (2022) A globally consistent local-scale assessment of future tropical cyclone risk. Science Advances 8(17):eabm8438, DOI 10.1126/sciadv.abm8438

van den Brink HW (2019) An effective parametrization of gust profiles during severe wind conditions. Environmental Research Communications 2(1):011,001, DOI 10.1088/2515-7620/ab5777

Brooks HE (2013) Severe thunderstorms and climate change. Atmospheric Research 123:129–138, DOI 10.1016/j.atmosres.2012.04.002

Chan CH, Ma KF, Shyu JBH, et al (2020) Probabilistic seismic hazard assessment for Taiwan: TEM PSHA2020. Earthquake Spectra 36(S1):137–159, DOI 10.1177/8755293020951587

Danciu L, Nandan S, Reyes C, et al (2021) The 2020 update of the European seismic hazard model: Model overview. Tech. Rep. Technical Report 001, v1.0.0, EFEHR, DOI 10.12686/a15

Davydzenka T, Tahmasebi P, Shokri N (2023) Unveiling the global extent of land subsidence: The sinking crisis. Geophysical Research Letters 51:e2023GL104,497, DOI 10.1029/2023GL104497

Deltares (2021) Flood hazard maps. Planetary Computer and Deltares global data (version 11206409-003-zws-0003) [dataset]. https://microsoft.github.io/AIforEarthDataSets/data/deltares-floods.html

Erdik M, Sesetyan K, Demircioglu M, et al (2012) Assessment of seismic hazard in the Middle East and Caucasus: EMME (Earthquake Model of Middle East) project. In: Proceedings of the 15th World Conference on Earthquake Engineering, Lisbon, Portugal

Floors R, Badger J, Troen I, Grogan K, Permien FH (2021) Satellite-based estimation of roughness lengths and displacement heights for wind resource modelling. Wind Energy Science 6(6):1379–1400, DOI 10.5194/wes-6-1379-2021

Gerstenberger MC, Bora S, Bradley BA, et al (2024) The 2022 Aotearoa New Zealand national seismic hazard model: Process, overview, and results. Bulletin of the Seismological Society of America 114(1):7–36, DOI 10.1785/0120230182

Ghasemi H, Cummins PR, Weatherill GA, et al (2020) Seismotectonic model and probabilistic seismic hazard assessment for Papua New Guinea. Bulletin of Earthquake Engineering 18:6571–6605, DOI 10.1007/s10518-020-00966-1

Giardini D, Grünthal G, Shedlock KM, Zhang P (2000) The GSHAP Global Seismic Hazard Map. Seismological Research Letters 71(6):679–691, DOI 10.1785/gssrl.71.6.679

Giardini D, Danciu L, Erdik M, et al (2016) Seismic hazard map of the Middle East. Tech. rep., Global Earthquake Model Foundation, DOI 10.12686/a1

Harris RI, Deaves DM (1981) The structure of strong winds. Construction Industry Research and Information Association (CIRIA), London

Hersbach H, Bell B, Berrisford P, et al (2020) The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society 146:1999–2049, DOI 10.1002/qj.3803

Hoch JM, Probyn I, Marra F, et al (2025) BURGER: A bottom-up regionalization approach for global sub-daily intensity-duration-frequency data. Water Resources Research 61(10):e2024WR039,773, DOI 10.1029/2024WR039773

Irsyam M, Cummins PR, Asrurifak M, et al (2020) Development of the 2017 national seismic hazard maps of Indonesia. Earthquake Spectra 36(S1):112–136, DOI 10.1177/8755293020951206

Kaplan JO, Lau KHK (2021) The WGLC global gridded lightning climatology and time series. Earth System Science Data 13(7):3219–3237, DOI 10.5194/essd-13-3219-2021

Kaplan JO, Lau KHK (2022) World Wide Lightning Location Network (WWLLN) Global Lightning Climatology (WGLC) and time series, 2022 update. Earth System Science Data 14(12):5665–5670, DOI 10.5194/essd-14-5665-2022

Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS 2017), Curran Associates, Inc., pp 4768–4777

Midzi V, Zulu B, Manzunzu B, et al (2020) Macroseismic survey of the M_L 5.5, 2017 Moiyabana, Botswana earthquake and seismic hazard assessment of South Africa. Journal of African Earth Sciences 168:103,845, DOI 10.1016/j.jafrearsci.2020.103845

Mirus BB, Belair GM, Wood NJ, Jones JM, Martinez SN (2024) Parsimonious high-resolution landslide susceptibility modeling at continental scales. AGU Advances 5(5):e2024AV001,214, DOI 10.1029/2024AV001214

Muis S, Verlaan M, Winsemius HC, Aerts JCJH, Ward PJ (2020) A high-resolution global dataset of extreme sea levels, tides, and storm surges, based on the Global Tide and Surge Model (GTSM) forced with ERA5. Journal of Geophysical Research: Oceans 125(4):e2019JC015,513, DOI 10.1029/2019JC015513

Nath SK, Thingbaijam KKS (2012) Probabilistic seismic hazard assessment of India. Seismological Research Letters 83(1):135–149, DOI 10.1785/gssrl.83.1.135

Ornthammarath T, Warnitchai P, Chan CH, et al (2020) Probabilistic seismic hazard assessments for northern Southeast Asia (Indochina): Smooth seismicity approach. Earthquake Spectra 36(S1):69–90, DOI 10.1177/8755293020942528

Pagani M, Garcia-Pelaez J, Gee R, et al (2020) The 2018 version of the Global Earthquake Model: Hazard component. Earthquake Spectra 36(S1):226–251, DOI 10.1177/8755293020931866

Palau RM, Nadim F, Paulsen E, Storrøsten E (2023) A new model for global landslide susceptibility assessment and scenario-based hazard assessment. Contributing paper, Norwegian Geotechnical Institute, contributing to the Global Infrastructure Resilience Index (GIRI) for the Coalition for Disaster Resilient Infrastructure (CDRI) Biennial Global Infrastructure Resilience Report, 2023

Petersen MD, et al (2018) Seismic hazard, risk, and design for South America. U.S. Geological Survey Data Release, DOI 10.5066/F7Wm1BK1

Petersen MD, et al (2023) Data release for the 2023 U.S. 50-state national seismic hazard model – overview. U.S. Geological Survey Data Release, DOI 10.5066/P9GNPCOD

Poggi V, Durrheim R, Tuluka GM, et al (2017) Assessing seismic hazard of the East African Rift: A pilot study from GEM and AfricaArray. Bulletin of Earthquake Engineering 15(11):4499–4529, DOI 10.1007/s10518-017-0152-4

Rodríguez O, Bech J (2020) An evaluation of convective parameters for significant severe weather in northeastern Iberia. Atmospheric Research 240:104,930, DOI 10.1016/j.atmosres.2020.104930

Shumway AM (2019a) Data release for the 2012 American Samoa and neighboring South Pacific Islands seismic hazard model. U.S. Geological Survey Data Release, DOI 10.5066/P9GXLNIB

Shumway AM (2019b) Data release for the 2012 Guam and Northern Mariana Islands seismic hazard model. U.S. Geological Survey Data Release, DOI 10.5066/P9USPEKG

Shumway AM (2020) Data release for the 2010 Haiti seismic hazard model. U.S. Geological Survey Data Release, DOI 10.5066/P9OY8WT2

Shumway AM, et al (2025) Datasets for the 2025 USGS national seismic hazard model for Puerto Rico and the U.S. Virgin Islands. U.S. Geological Survey Data Release, DOI 10.5066/P14ATJLJ

Sokolov V, Zahran HM, Youssef SEH, El-Hadidy M, Alraddadi WW (2017) Probabilistic seismic hazard assessment for Saudi Arabia using spatially smoothed seismicity and analysis of hazard uncertainty. Bulletin of Earthquake Engineering 15(7):2695–2735, DOI 10.1007/s10518-016-0075-5

Taszarek M, Brooks HE, Zrnić DS (2017) Wet microburst environments in Europe. Monthly Weather Review 145(6):2373–2390, DOI 10.1175/MWR-D-16-0336.1

Taszarek M, Brooks HE, Wieckowski W (2019) Severe thunderstorm environments in Europe: Reanalysis-derived proxies and future changes. Weather and Climate Extremes 22:100–110, DOI 10.1016/j.wace.2018.03.003

Taszarek M, Allen JT, Púčik T, Hoogewind KA, Brooks HE (2020) Severe convective storms across Europe and the United States. Part II: ERA5 environments associated with lightning, large hail, severe wind, and tornadoes. Journal of Climate 33(23):10,263–10,286, DOI 10.1175/JCLI-D-20-0346.1

Thøgersen E (2021) Downscaling of near-surface wind speed in complex terrain using machine learning and roughness corrections. PhD thesis, Department of Wind Energy, Technical University of Denmark (DTU), DOI 10.11581/DTU:00000053

Ullah S, Bindi D, Pilz M, et al (2015) Probabilistic seismic hazard assessment for Central Asia. Annals of Geophysics 58(1):S0103, DOI 10.4401/ag-6687

Vafeidis AT, Schuerch M, Wolff C, et al (2019) A new framework for coastal risk assessment. Nature Climate Change 9(11):889–895, DOI 10.1038/s41558-019-0601-7

To learn more about Degree Day's climate risk solutions,
please contact info@degreeday.org or visit www.degreeday.org.

Degree Day

www.degreeday.org