



ThinkHazard!

Identify natural hazards in your project area
and understand how to reduce their impact

• • • • • • •

Methodology report

Updated for ThinkHazard! Version 2

VERSION: 2.0

DATE: September 2017

Version 2.0: Stuart Fraser¹

ORIGINAL AUTHOR(S): John Douglas², Alanna Simpson¹, Stuart Fraser¹, Marijn Kuijper³, Hessel Winsemius³, Andreas Burzel³, Audrey Hohmann², Nicolas Taillefer², Sophie Vermooten³, with technical contributions from Frederic Jacon⁴ and Pierre Giraud⁴.

¹GFDRR, ²BRGM, ³Deltas, ⁴Camptocamp



Summary

This report describes the World Bank Global Facility for Disaster Reduction and Recovery (GFDRR)¹ hazard screening tool, *ThinkHazard!* (thinkhazard.org), which enables users to screen project locations for multiple natural hazards, globally. The simple workflow of the online user interface streamlines communication of hazard level and risk management guidance, to mainstream disaster risk information in development project planning. Further information resources and national contacts are presented to guide users in managing risk.

ThinkHazard! translates technical hazard data describing hazard intensity, frequency and susceptibility in scientific parameters, into four categories of hazard: very low, low, medium and high. The hazard classification is used to communicate hazard to users who are not expert in natural hazards but require hazard information for project planning and disaster risk management purposes.

ThinkHazard! version 2, released in July 2017, provides guidance in three languages (English, French and Spanish) for 11 hazards hydro-meteorological and geological hazards:

1. river ('fluvial') flood
2. urban flood (comprising river and surface, 'pluvial', flood in urban areas only)
3. coastal flood
4. tropical cyclone
5. water scarcity
6. extreme heat
7. wildfire
8. earthquake
9. landslide
10. tsunami
11. volcano

¹ GFDRR: www.gfdrr.org.

Table of contents

<i>Summary</i>	2
<i>Table of contents</i>	3
<i>Figures</i>	4
<i>Tables</i>	5
1 <i>Introduction</i>	6
1.1 <i>Background</i>	6
1.2 <i>Objective of ThinkHazard!</i>	7
1.3 <i>Report outline</i>	8
2 <i>Classifying hazard levels</i>	8
2.1 <i>Input data types</i>	9
2.2 <i>Data storage</i>	10
2.3 <i>Hazard data selection</i>	11
2.4 <i>Classification Process</i>	13
3 <i>Communication of hazard levels</i>	18
4 <i>Risk management guidance</i>	21
4.1 <i>Risk management recommendations</i>	21
4.2 <i>Further resources</i>	21
4.3 <i>Contact Information</i>	22
5 <i>Other tool features</i>	22
5.1 <i>Downloadable hazard reports</i>	22
5.2 <i>User feedback</i>	22
5.3 <i>Languages</i>	23
5.4 <i>Data map</i>	23
5.5 <i>Frequently Asked Questions</i>	23
5.6 <i>Extracting and re-using content</i>	24
5.7 <i>Licenses</i>	24
6 <i>Architecture and technical workflow</i>	24
6.1 <i>Open source architecture</i>	24
6.2 <i>Database structure</i>	25
6.3 <i>Technical workflow – hazard classification</i>	26
7 <i>Administrator role</i>	28
8 <i>Hazard specific classification methods</i>	30
8.1 <i>River Flood and Urban Flood Hazard Levels</i>	30
8.2 <i>Coastal Flood Hazard Levels</i>	33

8.3	Water Scarcity Hazard Levels.....	34
8.4	Tropical Cyclone Hazard Levels	38
8.5	Extreme Heat Hazard Levels (new in version 2)	40
8.6	Wildfire Hazard Levels (new in version 2).....	42
8.7	Earthquake Hazard Levels	44
8.8	Tsunami Hazard Levels	46
8.9	Volcanic Hazard Levels	50
8.10	Landslide Hazard Levels	54
9	References.....	56
10	Appendix 1: Document revisions	58
11	Appendix 2: Hazard coverage	60
12	Appendix 3: Metadata template	62
13	Appendix 4: Data Quality Criteria	65

Figures

Figure 1	The three page levels of thinkhazard.org. From left: homepage location search, location overview of all hazards, single hazard level and risk reduction recommendations	7
Figure 2	An earthquake hazard map for Europe (from the SHARE project). Hazard is shown as expected peak ground acceleration (PGA) with a 10% chance of being exceeding in a 50-year interval (return period of 475 years)	8
Figure 3	Decision-making algorithm for selection of data used in hazard classification	12
Figure 4	Comparison of Think Hazard frequency-based approach, and common intensity-based approach	14
Figure 5	Principles of hazard level aggregation from ADM2 up to ADM0 (country level)	18
Figure 6	Hazard levels communicated via the multi-hazard overview page	19
Figure 7	Hazard levels communicated via the single-hazard and risk reduction guidance page	20
Figure 8	Google form used to obtain user feedback	22
Figure 9	Global data map, showing the geographic level of data used in ThinkHazard, for each hazard and country	23
Figure 10	A Frequently Asked Questions section on the site provides information about most aspects of the tool, in non-technical language	24
Figure 11	General architecture of the tool and back office workflow	26
Figure 12	Screenshot of the administrator view, showing the management of technical recommendations	29
Figure 13	Final classification for river flood at ADM2 level: blue = very low; yellow = low; orange = medium; red = high	32
Figure 14	Final classification for urban flood at ADM2 level: blue = very low; yellow = low; orange = medium; red = high	33

<i>Figure 15 Graph representing the number of countries with a WCI below 1700 m³ per capita per year in relation to the return period</i>	36
<i>Figure 16 Water Crowding Index – Annual water availability per capita by water province</i>	37
<i>Figure 17 Water Crowding Index – Annual water availability per capita by ADM2 Unit</i>	37
<i>Figure 18 Hurricane zoning in the United States, ranging from 60 miles per hour (first damages) up to higher than 90 miles per hour (most severe damages) (Source: http://www.edgetech-us.com/map/MapHurriZone.htm)</i>	39
<i>Figure 19 Preliminary cyclone hazard classification for ADM2 Units using GAR15 global data</i>	40
<hr/>	
<i>Figure 20: Heat hazard classification map at the ADM2-level. The figure visualizes the location with very low (dark green), low (green), medium (orange) and high (red) extreme heat risk.</i>	41
<i>Figure 21: Classification of fire hazard (dark green = very low, light green = low, yellow = medium, red = high) based on variable thresholds of FWI and return period intervals of 2, 10 and 30 years.</i>	43
<i>Figure 22 Example seismic hazard map (source: FP7-SHARE project)</i>	44
<i>Figure 23 Tsunami maximum inundation height data at offshore points (output of Davies et al., 2017), for Sulawesi, Indonesia</i>	47
<i>Figure 24 Tsunami maximum inundation height data at the coast, interpolated from the points offshore Sulawesi, Indonesia</i>	48
<i>Figure 25 Global tsunami hazard levels at ADM1.</i>	50
<i>Figure 26 Global tsunami hazard levels at ADM2.</i>	50
<i>Figure 27 Hazard class at all locations within 100 km of a volcano. Red = high hazard, orange = medium, yellow = low. Very low and unknown hazard are not shown.</i>	53

Tables

<i>Table 1 A summary of damaging intensity parameters, units and thresholds used in the classification of probabilistic data (only one common unit shown for hazard)</i>	15
<i>Table 2 Suggested return periods for each hazard classified from probabilistic data</i>	16
<i>Table 3 Example of rationale justifying the choice of return periods, for river flood hazard ..</i>	31
<i>Table 4 Water scarcity risk category based on return period and its rationale, based on human experience</i>	35
<i>Table 5 Illustration of probability of damages for vulnerable buildings according to the intensity of seismic shaking (EMS98)</i>	45
<i>Table 6 Conversion table for units of earthquake data</i>	46
<i>Table 7 Example of pre-processed rasters for tsunami in Indonesia. Red zones show regions where wave height exceeds both elevation and the threshold on the buffered coastline</i>	49
<i>Table 8 Annual frequency and corresponding raster score for landslide, based on NGI (2013)</i>	55
<i>Table 9 Metadata template</i>	62
<i>Table 10 Data quality rating criteria and rating scheme. These ratings are applied in layer metadata and used in ThinkHazard! to determine which hazard data to use</i>	65

1 Introduction

1.1 Background

The Global Facility for Disaster Reduction and Recovery (GFDRR), established in 2006, is a multi-donor partnership and grant-making financing mechanism. The Facility supports on-the-ground technical assistance to help developing countries integrate disaster risk management (DRM) and climate change adaptation into development strategies, policies and investment programs, including post-disaster recovery and reconstruction.

GFDRR implements most of its activities in countries through the World Bank, in partnership with national, regional, and other international agencies. To support country engagements, GFDRR maintains eight thematic initiatives that provide grant recipients with specialized knowledge and quality assurance in the design and implementation of activities. These initiatives allow GFDRR to collaborate with a broad array of partners, facilitate global engagements and capacity building, and produce innovative knowledge. [GFDRR's eight areas of engagement](#) are: Promoting open access to risk information; Promoting resilient infrastructure; Scaling up the resilience of cities; Strengthening hydromet services and early-warning systems; Deepening financial protection; Building resilience at community level; Deepening engagements in resilience to climate change; and Enabling resilient recovery.

Risk information is a critical input to disaster risk management, building resilience of urban areas, infrastructure and communities. The *Open Access to Risk Information* area of engagement aims to make high-quality risk information available faster and at lower costs, and develop new tools that allow decision-makers and communities to collect, share, and understand risk information.

The *ThinkHazard!* project was initiated in 2015 to facilitate greater access to hazard information and risk management guidance for development sector professionals. The open access thinkhazard.org website enables users to screen potential project locations for the existence of multiple natural hazards, then to obtain guidance on how to manage the risks to their project, and any impacts of their project on the local hazard. Version 1 of *ThinkHazard!* was used over 140,000 times in 200 countries, and has been adopted into World Bank Operations Portal for core use in project planning. This report has been updated to reflect new hazards and revisions in methodology, made in the development of *ThinkHazard!* version 2, released in July 2017.

1.2 Objective of *ThinkHazard!*

Risk information is a critical input to disaster risk management (DRM) and building resilience of urban areas, infrastructure and communities. Interpretation of hazard information to determine potential impact and, in turn, implement robust a risk management strategy requires knowledge of often highly technical data. Additionally, hazard data is generated in many different formats from many organizations. If the data is made available publicly at all, it is via a growing number of online sources and data portals. So, the task to find hazard data sets, decide which data set to use for a given purpose, and then interpret the data can require an increasing amount of time and prior knowledge.

DRM projects have too often been executed without full appreciation of the range of hazards and level of each hazard that exist in a project area. This can result in underestimation of disaster risk and development of a project that is potentially not robust enough to withstand the impacts of one or more hazards that could affect the project in the future. Access to, and appropriate use of hazard and risk information remains a barrier to mainstreaming DRM into development projects.

***ThinkHazard!* is an analytical tool dedicated to improving knowledge and understanding of natural hazards.** It was envisioned that the *ThinkHazard!* primary user would be development sector professionals, who need to gather hazard information while planning projects. However, the benefits of *ThinkHazard!* stretch beyond the development sector, into general education about global distribution of multiple hazards and how to manage them.

To achieve its objectives, the online user interface has a simple structure and workflow, comprising, 1) a location search function, 2) an overview of hazard level for 11 hazards for a selected location, and 3) a hazard-specific screen that presents risk management guidance, relevant contact information and further information in the form of useful websites and reports for that hazard and location.

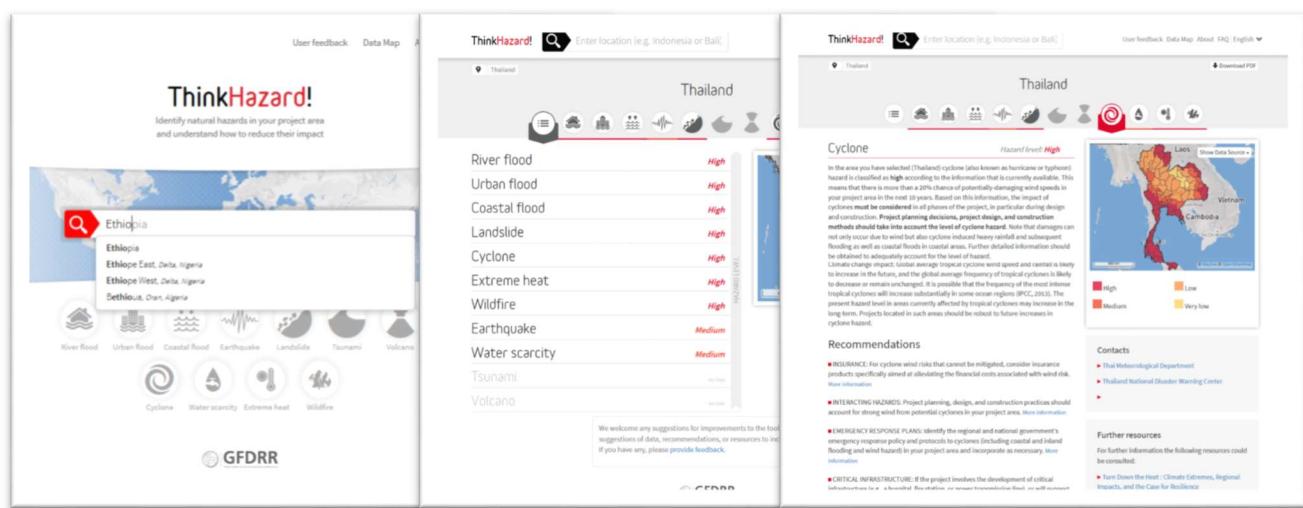


Figure 1 The three page levels of thinkhazard.org. From left: homepage location search, location overview of all hazards, single hazard level and risk reduction recommendations

1.3 Report outline

This report communicates the methods used to communicate technical hazard data as simpler hazard levels (section 2). It then presents: the strategy for communicating multiple hazard levels (section 3); risk management guidance or ‘next steps’ after identifying the hazard level (section 4); other tool features (section 5); software architecture and workflow (section 6); the role of the tool administrator (section 7); and the classification method used for each hazard (section 0).

2 Classifying hazard levels

ThinkHazard! communicates level of hazard for 11 natural hazards, for each Administrative Unit 2 (ADM2) globally, in the classes High, Medium, Low, and Very Low. The hazard levels are used to communicate how aware a user needs to be of each hazard when planning a project at a selected location. Hazard level is aggregated to ADM1 and ADM0 unit levels, providing a view of hazard for users with a less certain project location, or projects spanning large areas in a country.

The four hazard levels are derived from **hazard maps**, which present the spatial distribution of hazard intensity (e.g., flood depth, ground shaking) at a given frequency, or ‘return period’ (e.g., Figure 2). A hazard map is the visualization of hazard at one point of a frequency-severity curve; the distribution of hazard intensity varies differs at each frequency. Such maps contain valuable information on where and how often an event of certain magnitude or intensity might occur, but often require specialist knowledge for interpretation. Often, it can be difficult to identify and access such maps. *ThinkHazard!* aims to overcome these issues by collating specialist hazard maps and communicating hazard level accurately but more simply for hazard non-specialists.

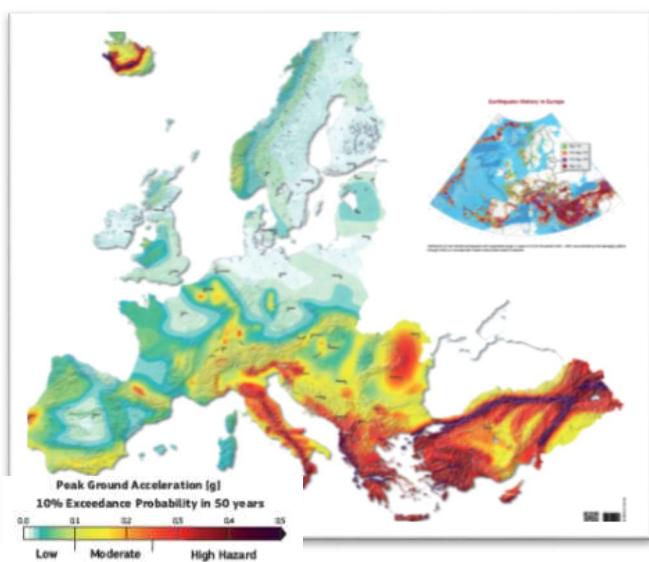


Figure 2 An earthquake hazard map for Europe (from the SHARE project). Hazard is shown as expected peak ground acceleration (PGA) with a 10% chance of being exceeded in a 50-year interval (return period of 475 years)

Hazard levels can be described as:

- **High:** Users should be highly aware of potentially severe damage from this hazard for the project location. Without taking measures to mitigate the hazard and risk, high levels of damage can be expected to occur within the project or human lifetime (and potentially frequently in that timeframe, for hydro-meteorological hazards, e.g., floods, extreme heat).
- **Medium:** Users should be aware of potentially damaging effects of this hazard for the project location. Potentially damaging events can be expected to occur within the project or human lifetime and measures to mitigate the hazard and risk should be considered. For hydro-meteorological hazards, damaging effects could occur frequently in that timeframe.
- **Low:** Potentially damaging events are less likely to occur within the project or human lifetime but are still possible. Measures to mitigate the hazard and risk would be prudent at critical locations. Hazard has been classified based on long-term averages, and there is still potential that damaging events could occur in this timeframe.
- **Very Low:** Available data suggest that potentially damaging effects are unlikely to occur, on average, in the project or human lifetime. Hazard has been classified based on long-term averages, and there is still potential that damaging events could occur in this timeframe.
- **No Data Available:** No dataset covering the chosen location is currently available in *ThinkHazard!*

NOTE: The timeframe considered in classification of each hazard is dependent on historical data available to assess long-term averages, and the timescales over which the hazard causal processes operate. This is elaborated later in this chapter.

Hazard data is available at different geographic scale, spatial extent or domain, and content. It can be provided as probabilistic data, providing estimated hazard severity and frequency, or index data showing susceptibility of an area to a hazard. The type of data typically differs between hazard – earthquake and flood are generally available as probabilistic data but landslide is not. Different organizations also provide different data according to their geographic remit or focus on one hazard or one group of hazards. It has been necessary to develop a consistent framework of classification, but which is implemented in different ways to accommodate different data types. The sections below describe the input data types, data storage, selection of data source, and the classification process.

2.1 Input data types

ThinkHazard! accepts data in **raster (grid) format only**. It classifies **probabilistic and index-based data using different methods**, though the theory remains consistent. Probabilistic data are commonly available for: flood (river, urban and coastal), earthquake, cyclone, extreme heat, wildfire, water scarcity, and tsunami. Probabilistic data are produced for landslide and volcano in local-scale analysis, but at the global scale data are more commonly available as a spatial susceptibility index for landslide. For volcanic hazards there are very limited global

sources of information: a spatial hazard index is available, but otherwise the other global source is a catalogue of previous events.

Probabilistic data are imported to *ThinkHazard!* directly and all classification is conducted in the tool algorithms, described in section 0. In some cases, **pre-processing** is required to convert data into a format accepted by *ThinkHazard!*. Pre-processing comprises spatial analysis and / or data conversion prior to use in *ThinkHazard!*, using external GIS software such as ArcGIS or QGIS. This is described in section 2.4.2.

2.2 Data storage

GeoNode, an open source geospatial content management system, is used to store all hazard data and additional information layers for *ThinkHazard!*, which is set up to retrieve data directly from a defined GeoNode server. The current version of *ThinkHazard!* harvests data from the [GFDRR Innovation Lab GeoNode](http://geonode-gfdrrlab.org) (geonode-gfdrrlab.org).

***ThinkHazard!* administrators curate input data directly in GeoNode** (see Section 7). This includes hazard data upload and metadata curation using the template shown in

Appendix 3: Metadata template. The metadata template is adapted from an ISO standard to contain additional information specifically required for *ThinkHazard!*, including: hazard type, intensity unit, geographic coverage, return period and hazard set ID (all used in hazard classification), and data quality score (used in deciding which data set to use in classification). Documents provided in the further information section of *ThinkHazard!* are also managed in GeoNode.

GeoNode is also used in *ThinkHazard!* to communicate layer and document metadata and provides the route for users to download data. Clicking on a ‘data source’ or ‘further information’ link in *ThinkHazard!* takes the user to GeoNode, from where they can download publicly available data and reports, or navigate to the data owners’ website for access. License information for data is also stored in the layer metadata and communicated via GeoNode.

2.3 Hazard data selection

ThinkHazard! aims to uses the **best available hazard information to derive the hazard level for each administrative unit**. A data selection algorithm determines which layer to use in classifying hazard levels, where more than one hazard data set is available for a hazard and ADM2 Unit (see Figure 3). This algorithm is based on the spatial extent of the data, identified from its metadata; and a data quality rating, which is stored in metadata by *ThinkHazard!* administrators according to criteria in Appendix 4: Data Quality Criteria.

The data quality criteria provide a high-level assessment of ‘Scientific Quality’ (i.e. is the data produced using a peer-reviewed method; is it official government data) and ‘Calculation Method Quality’, which focuses on the fidelity of the hazard model and its components (e.g. resolution of base data, vintage of analysis).

Hazard data may be available with global, regional or national coverage, depending on the scope of the generating project, or intended primary use of the data. *ThinkHazard!* can accept all levels of data, and implements a hierarchical selection to prefer local data (regional and national) where it is available, over global data. Global data is considered to provide a less robust view of hazard for an individual country than local data, because it is generally produced at lower resolution, and uses a greater number of global assumptions (not tailored to a regional or national situation). However, global data can provide a consistent view of hazard across all (or most) countries and is valuable in the absence of data with more local extent, which is expected to be of greater fidelity due to its more focused scope.

To summarize the data selection algorithm, conducted for each ADM2 Unit:

1. If only one dataset is available, that dataset is used by default to classify hazard level;
2. If more than one dataset is available, the dataset with the highest summed data quality score is used to classify hazard level;
3. If more than one dataset is available, with equal summed data quality scores, then local data is preferred over global data to classify hazard level;

NOTE: The data quality scores are assigned for the purposes of *ThinkHazard!* are in no way a judgment on the applicability of the data for other applications.

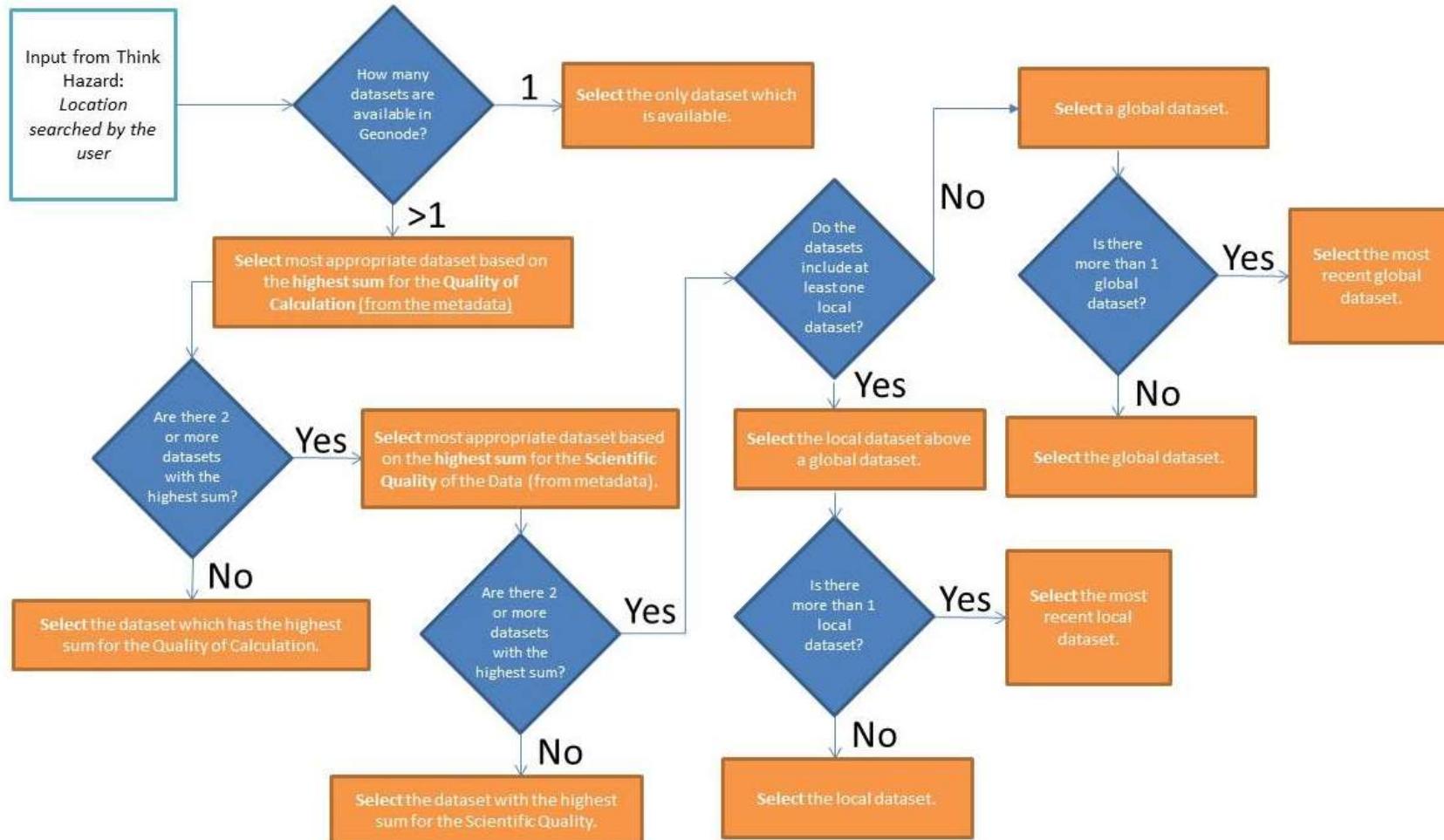


Figure 3 Decision-making algorithm for selection of data used in hazard classification

2.4 Classification Process

The *ThinkHazard!* hazard classification employs a **spatial analysis that intersects raster hazard data, with a global raster layer describing every ADM2 Unit in the world**. The ADM2 raster is derived from the FAO GAUL vector dataset of administrative unit boundaries². Three levels of administrative units are used in *ThinkHazard!*: ADM0 (country), ADM1 (e.g., state in the U.S.) and ADM2 (e.g., county in the U.S.).

Following the above data selection process, to determine which layer should be used for which ADM2 Unit, *ThinkHazard!* imports the selected data from the GeoNode database, classifies the hazard level and stores the hazard level for each hazard and ADM2 Unit in the tool database. This hazard level is communicated in text descriptions on the user interface, and the spatial distribution of hazard is also shown in the map window.

The classification of hazard level is based on the aim to communicate **how aware a user needs to be of each hazard** when planning a project at a selected location, or alternatively, **how frequently a project location may sustain damage from a hazard**. The process relies on defining a level of hazard intensity, above which damage may occur, and the likelihood of that intensity occurring at each location.

2.4.1 Probabilistic data

ThinkHazard! uses frequency and severity information to communicate **how frequently a project location may sustain damage from a hazard**. This is a well-defined process for probabilistic data, which provide both elements required for this assessment (estimates of hazard frequency and severity).

To do this, we **first identify an intensity level for each hazard, above which damage is expected to occur, and then assess how frequently that intensity might be exceeded**. This information is available on frequency-severity curves, which are a product of probabilistic analysis (see Figure 4, right). The more common use of these curves is to define how severe an event would be at a defined frequency (Figure 4, left) to define required building strength, for example. The first formulation is chosen as the key communication point, because *ThinkHazard!* aims to help users prioritize and manage multiple hazards with the greatest chance of causing damage to their interests, rather than identify the exact impact (this should be done in a more focused risk assessment as part of the project).

Frequency of a hazard intensity being exceeded can be defined in terms of average recurrence interval, or **return period**, expressed as ‘1 in 100 years’, or the ‘100-year return period’. Alternatively, this can be expressed as the chance of the intensity value being exceeded on an annual basis: for the 100-year return period hazard this would be 1% chance of exceedance in any given year ($1.0\% = 1/100$); for the 500-year return period this is 0.2% ($0.2\% = 1/500$). This report uses return period as the reference to frequency. Longer return periods correspond to having a smaller chance that the damaging intensity will be exceeded during the reference timeframe lifetime, hence the risk of damage is lower.

² Food and Agriculture Organisation (FAO) Global Administrative Unit Layer (GAUL); <http://www.fao.org/geonetwork/srv/en/metadata.show?id=12691&currTab=simple>

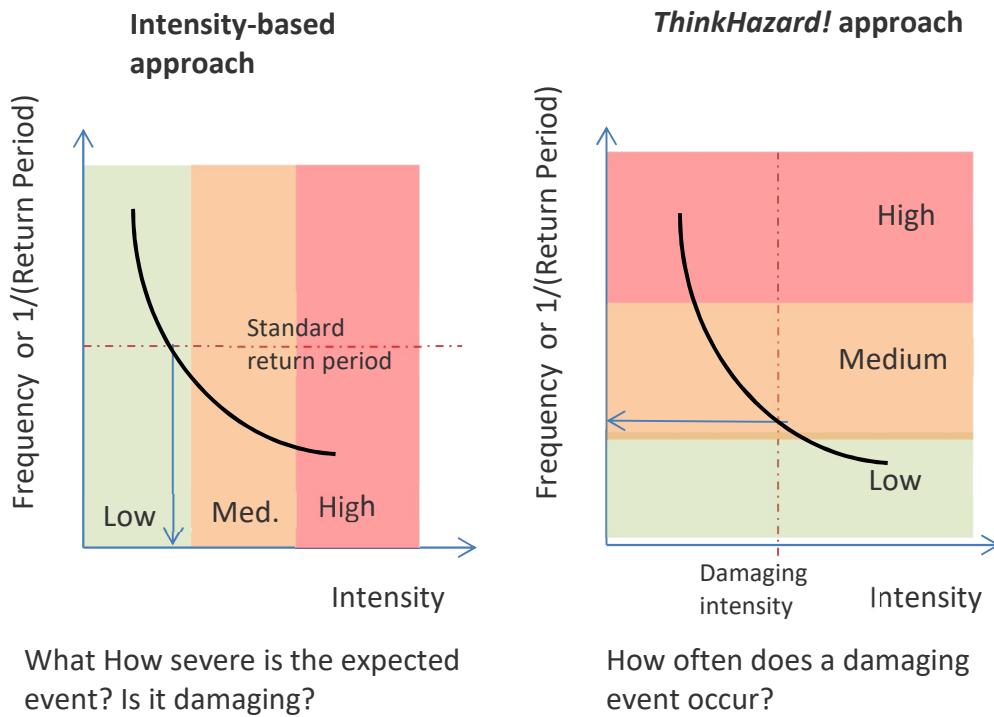


Figure 4 Comparison of Think Hazard frequency-based approach, and common intensity-based approach

2.4.1.1 Classification steps

A step by step procedure is applied to classify hazards based on probabilistic data in *ThinkHazard!*. This procedure is applied for each ADM2 Unit:

1. Hazard = HIGH if hazard intensity exceeds, at any location in the ADM2 Unit, the damaging intensity threshold at the shortest return period (highest frequency threshold);
2. Hazard = MEDIUM if (1) is not satisfied and if hazard intensity values exceed the damaging intensity threshold at the medium return period;
3. Hazard = LOW if (1, 2) are not satisfied and if the hazard intensity values exceed the damaging intensity threshold at the longest return period (lowest frequency threshold);
4. Hazard = VERY LOW if (1, 2, 3) are not satisfied in data values given for the ADM2 Unit.
5. NO DATA is recorded if there are no data values given for the ADM2 Unit.

The ‘damaging intensity threshold’ and ‘frequency threshold’ are described in the next two sections.

2.4.1.2 Damaging intensity threshold

Damaging intensity threshold, the intensity above which damage would be expected to occur, is defined specifically for each hazard (Table 1). In this version of *ThinkHazard!* the **type of project is not specified by the user, and guidance is intended to be project- and sector-agnostic, to be applicable to a wide range of project types**. The defined damage threshold is therefore not tailored specifically to account for damage to any specific group of structures or infrastructure. **Conservative (i.e. low) damage thresholds are used because they are intended to reflect intensity that can cause damage for projects in International Development Association (IDA) countries**, in which investments may be more vulnerable (in relation to construction and/or availability of resources for recovery). Therefore, the thresholds used may be not be realistic in relation to highly-engineered structures. Later sector-specific versions of *ThinkHazard!*, may have damage thresholds tailored to that sector, for example, using construction standards for critical facilities to determine the intensity of event that could be considered damaging.

The three intensity thresholds for a hazard can be identical across the hazard levels, or they can differ to provide the flexibility to further distinguish each hazard level. For example, earthquake shaking reaching intensity VI on EMS98 scale are considered damaging for most projects, so the equivalent ground acceleration value of 0.1 g is defined as the intensity threshold for very low, low and medium hazard. For high hazard, the threshold reflects an intensity value typical of more damaging events that cause damage to structures: 0.2 g. In doing so, the number of high hazard zones is reduced compared to if a threshold of 0.1 g had been used, preserving the significance of high hazard by not distributing ‘high’ widely.

Hazard data for the same hazard may be provided with different intensity units. For example, earthquake ground acceleration can be provided in *gal*, *g* as a decimal, or *g* as a percentage. Flood depth may be provided in *m*, *cm*, or *dm (decimeter)*. The unit for each data layer is defined in its metadata, and *ThinkHazard!* classifies the data according to the correct unit.

Table 1 A summary of damaging intensity parameters, units and thresholds used in the classification of probabilistic data (only one common unit shown for hazard)

Hazard	Intensity parameter	Intensity Unit	Intensity threshold value		
			High	Medium	Low
Earthquake	Acceleration (PGA)	g	0.2	0.1	0.1
Extreme heat	Wet Bulb Globe Temperature (WBGT)	Degrees C	32	28	25
River flood*	Inundation depth	m	0.5	0.5	0.5
Urban flood*	Inundation depth	m	0.5	0.5	0.5
Coastal flood	Inundation depth	m	2	0.5	0.5
Cyclone	Mean wind speed	km/h	80	80	80
Tsunami	Coastal maximum amplitude	m	2	1	0.5
Water scarcity	Water availability	m ³ /capita/yr	<=500	<=1000	<=1700
Wildfire	Canadian Fire Weather Index	FWI	30	20	15

* Although river flood and urban flood data are provided as probabilistic data, it was found in *ThinkHazard!* version 1 that use of a damaging intensity alone overestimates high hazard (it is common for flood data to have a very small number of cells attributed to high hazard, which overestimate the hazard for the whole ADM2 Unit). Therefore, it was decided to use an additional spatial assessment in a pre-processing stage, to classify hazard based on a *minimum area exceeding the damaging threshold*. Further information is given in section 8.1.

2.4.1.3 Frequency thresholds

Frequency thresholds determine which hazard level is assigned to a location, given that the damaging intensity threshold is achieved. These thresholds are assigned on a hazard-specific basis, based on *de facto* standards for a given peril – i.e. the return periods at which data are usually produced. The common return periods reflect the timescales over which a hazard operates. For example, damaging floods occur more frequently than earthquakes due to atmospheric process operating on much shorter timeframes than tectonic processes.

Return periods may vary depending on the source of the data: earthquake data are more commonly produced at the 100, 250, 500, and 1000 years by the insurance industry due to their required outputs, instead of the standard 475, 975 and 2475 years used by researchers and engineers. For flood data, return periods commonly include some but not all of 5, 10, 25, 50, 100, 500, and 1000 years. *ThinkHazard!* has some flexibility to account for these differences in the classification, i.e., a 475 or 500-year return period is used to classify medium hazard. This enables as many datasets as possible to be used. *De facto* return periods do not exist for all hazards. In these cases, the frequency thresholds are defined using expert judgement based on the frequency of impacts. Generally, hazards threatening life due to construction damage (earthquake, cyclone, volcanoes) can be assessed considering, at a minimum, a 50 years lifespan. Hazards threatening people's living conditions more regularly (flood, drought) are often assessed using shorter durations (e.g., 10 years).

ThinkHazard! uses the return periods given in Table 3Error! Reference source not found. to classify hazard based on probabilistic data. A value range indicates the maximum range that can be used in the calculation. The administrator can determine the exact return period used by including (or excluding) certain return periods from data storage and import to *ThinkHazard!*.

Table 2 Suggested return periods for each hazard classified from probabilistic data

Hazard	Return period (years)		
	High	Medium	Low
Earthquake	100 - 250	475 - 500	1000 - 2500
Extreme heat	5	20	100
River flood	10	50	1000
Urban flood	10	50	1000
Coastal flood	10	50	100
Cyclone	50	100	1000
Tsunami	100	500	2500
Water scarcity	5	50	1000

2.4.2 Pre-processed data

Preprocessing of data into a format suitable for *ThinkHazard!* may be needed if:

1. Original hazard data are not available with frequency and severity information as provided by probabilistic data. These are ‘non-probabilistic data’.
 - a. Volcanic hazard data, globally, is not widely communicated probabilistically, so global hazard levels can only be derived based on databases of eruptive histories. Hazard data for individual sites / small areas may be the exception. See section 8.9 for more information.
 - b. Landslide hazard data. Global landslide data is presently available as susceptibility classes or hazard classes, because there is insufficient historical information to understand the frequency of events on such a wide scale. See section 8.10 for more information.
 - c. Note that probabilistic data for the volcanic and landslide hazards are more likely to be available at regional and local scale. Where such analysis is available, we will seek to use the information in *ThinkHazard!*.
2. Available hazard data do not intersect administrative units, preventing use of the general procedure. This is true of some coastal hazard data, e.g., tsunami, which are often provided as data points at some distance offshore because there is high computational expense of simulating flow onto land large areas. Depending on the location of points, some pre-processing is required to transfer the data points to the coastline or onshore. This is may be through spatial analysis to relocate points with the same values, or may involve analytical steps that account for change in values to the new location (e.g., application of tsunami run-up equations). See section 8.7 for more information.
3. Where classification of intensity is insufficient to determine variations in hazard level at the Administrative Unit scale. River flood and urban flood damaging intensity thresholds are found to be exceed at least one grid cell in most Administrative 2 Units, leading to a large proportion of areas with high hazard. This was the case in *ThinkHazard!* version 1, and makes it difficult to define and communicate the truly high hazard areas. To resolve this, hazard classification in *ThinkHazard!* version 2 uses the ‘area of an Admin 2 unit flooded’ to the damaging intensity, which identifies the units with most area flooded to a damaging flood depth. See section 8.1 for more information.

2.4.3 Aggregating hazard levels

Hazard classification is conducted for at the level of ADM2 Unit. That is, *ThinkHazard!* calculates the hazard level for each Admin 2 Unit in the world, by applying the process shown in section 2.4.1.1 to the grid cells within each ADM2 Unit boundary. The hazard level of a higher level ADM1 or ADM0 units is defined as the maximum hazard level in all lower units that it contains. In Figure 5, the Moroccan ADM1 unit named *Center* comprises eight ADM2

units. At the ADM2 level, there are some classified as ‘Low’ hazard (e.g. *Settat*) but the maximum hazard level is ‘High’ (*Azilal* and *Beni Mellal*) so the hazard level of ADM1 unit *Center*, is classified as ‘High’. The hazard levels in the six ADM1 units of Morocco are ranked in this example as either ‘Low’ or ‘High’. The ADM0 unit (Morocco) takes the maximum hazard level: in this case, ‘High’.

The aggregation process means that, in the extreme case, just one ADM2 unit in a country categorized as ‘High’, would result in the user seeing a ‘High’ hazard level for the whole country. This is intentionally conservative, so that *ThinkHazard!* shows the maximum hazard in any ADM unit. This is to safeguard against underestimating the overall hazard, when a large area (e.g., Province or country) is selected. The hazard map on the user interface provides further context, showing the distribution of hazard level across ADM units within the selected unit, enabling the user to see how the hazard level varies across the whole unit.

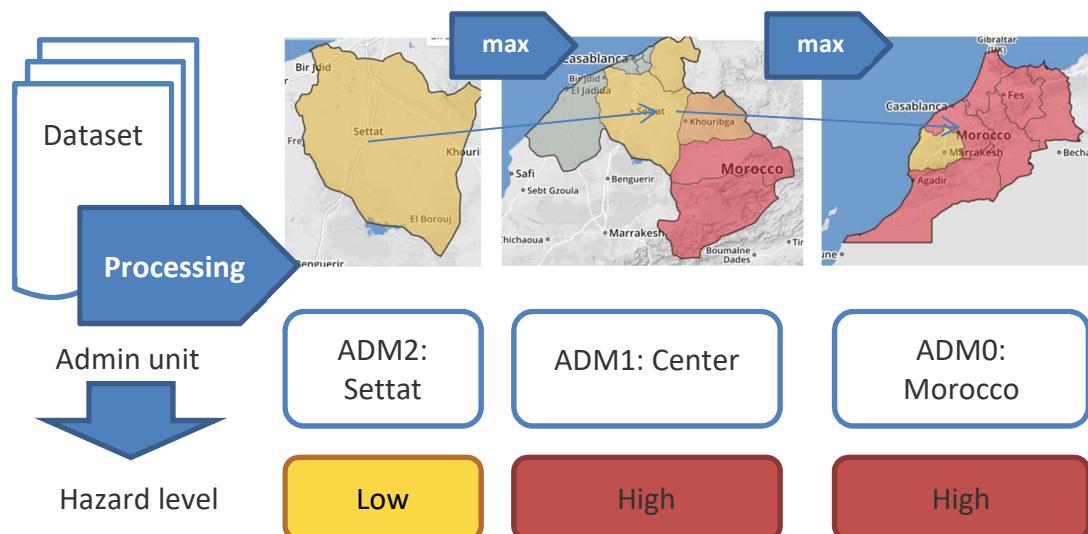


Figure 5 Principles of hazard level aggregation from ADM2 up to ADM0 (country level)

3 Communication of hazard levels

Hazard level is communicated through text description and color. On the location overview page (Figure 6), the horizontally arranged hazard icons are colored according to hazard level (High: red, through orange, to Very Low: yellow). On the overview page, the 11 hazards are listed vertically according to hazard level, with a colored text label.

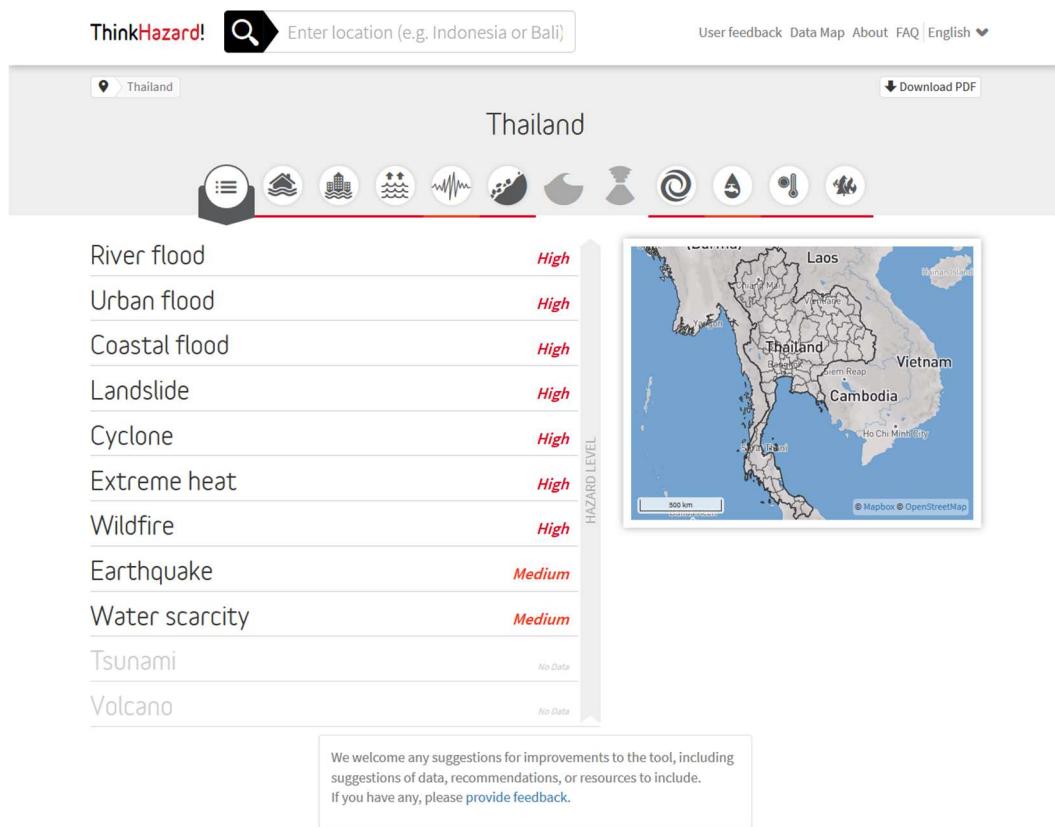


Figure 6 Hazard levels communicated via the multi-hazard overview page

On the hazard-specific page (Figure 7), the colored icons of all hazard are retained, and a colored text label specific to the hazard is shown. **All icons are shown on every page, to always highlight the presence of multiple hazards, and to provide easy navigation to other hazards.**

The spatial distribution of hazard is also shown in the map window. The map window provides limited user navigation. The user can select neighboring units, or zoom in and out to hierarchically-associated units (e.g., from ADM0 to ADM1 to ADM2, and vice versa).

There is a text description of what the hazard means for the user, in terms of the likelihood of damage, and encourages the user to take the hazard into consideration appropriately. For example:

*'In the area you have selected river flood hazard is classified as **high** based on modeled flood information currently available to this tool. This means that potentially damaging and life-threatening river floods are expected to occur at least once in the next 10 years. **Project planning decisions, project design, and construction methods must take into account the level of river flood hazard.** The following is a list of recommendations that could be followed in different phases of the project to help reduce the risk to your project. Please note that these recommendations are generic and not project-specific.'*

A brief statement is made to describe the potential impact of climate change on the hazard. This statement is derived from expert knowledge of the hazard, and IPCC³ reports, and is specific to the hazard and the region. For example: '*Climate change impacts: High confidence in an increase in intense precipitation. The present hazard level is expected to increase in the future due to the effects of climate change. It would be prudent to design projects in this area to be robust to river flood hazard in the long-term.*'

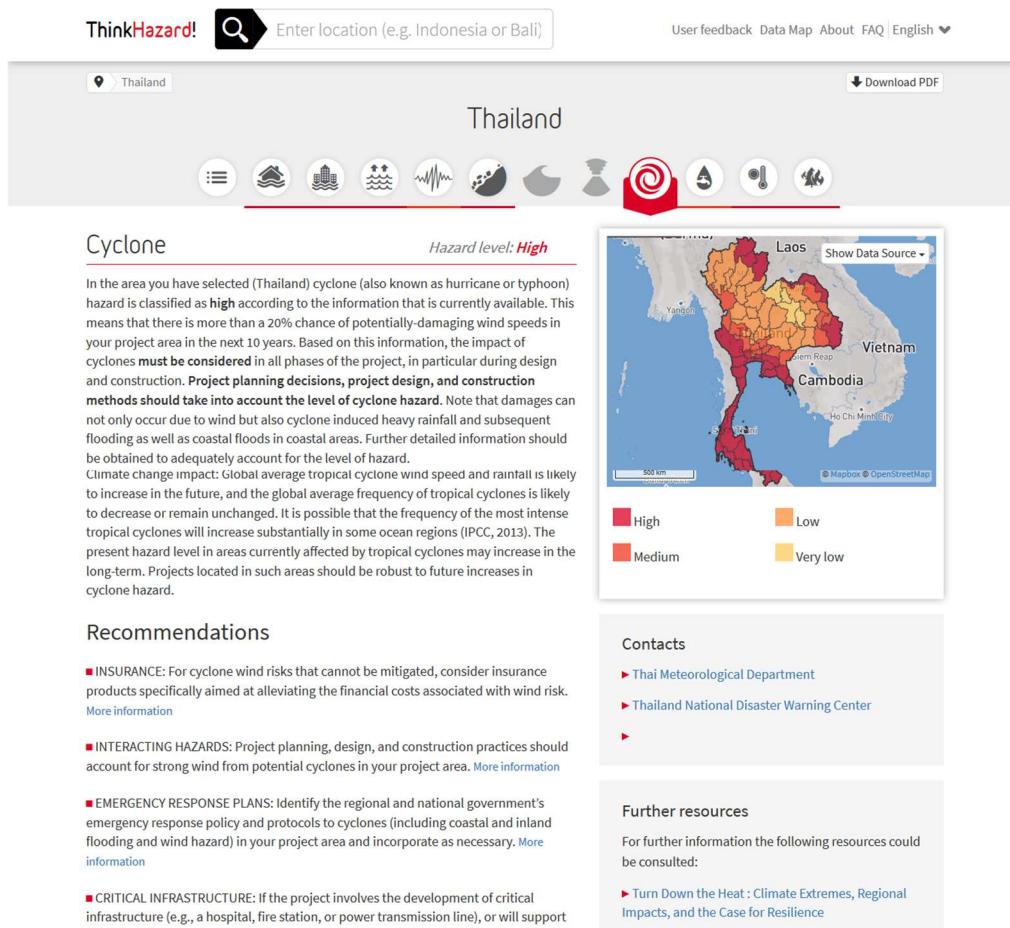


Figure 7 Hazard levels communicated via the single-hazard and risk reduction guidance page

³ Intergovernmental Panel for Climate Change

4 Risk management guidance

One of the most important goals of *ThinkHazard!* is to help users manage disaster risk in their projects, not only to identify the hazard level. This guidance takes the form of:

1. recommendations on how to address the risk from each hazard;
2. links to further resources, such as other data sources, reports on risk management options, or risk analyses relevant to that hazard and/or location; and
3. Contacts information of organizations working on the hazard in that country.

These features are all shown in Figure 7.

4.1 Risk management recommendations

ThinkHazard! displays hazard level for selected locations, not risk level (it does not estimate risk based on exposure and vulnerability to hazard). However, to maximize utility for user, technical recommendations are provided describing actions to be taken to manage risk to the project, based upon the classified hazard level. The recommendations have been developed by external experts in each hazard, consulted for this project. The recommendations are project agnostic (not specific to any project type or sector such as education, agriculture, or sanitation).

Recommendations comprise a short summary statement, which is shown on the user interface, and a more detailed description of risk management actions, which is accessed by clicking on the '*more information*' link at the end of each summary statement. The recommendations have been developed to provide guidance in non-technical language, and to provide practical actions specific to each hazard.

A recommendation may be repeated for more than one hazard category, but higher hazard categories generally have a greater number of recommendations, comprising a greater number of, or the more serious, actions required to manage the highest level of hazard. Each technical recommendation comprises a short summary, which is presented on the user interface, and a more detailed section that can be accessed by clicking the hyperlinked short title. The technical recommendations cover subjects such as seeking professional guidance, considering early warning systems, or not increasing the hazard through the actions of the project. Please see the user interface for examples.

4.2 Further resources

The user interface includes a section providing further resources to the user. This comprises links to documents, analytical tools and websites that provide information relevant to the hazard and selected location. The resources may provide information on previous projects in the area, general advice on managing risk to a hazard, or provide a link to relevant agencies and tools for further analysis.

All resources are stored on the same GeoNode as hazard data layers, and imported to *ThinkHazard!* for display. The resources are filtered by hazard and location, according to their metadata, and are displayed in chronological order in the user interface. A user downloads or accesses a link via the GeoNode record for that document.

4.3 Contact Information

Contact information has been compiled for every country and hazard included in ThinkHazard!. This comprises organization name, website URL, telephone and email address, for up to three organizations per hazard and country. This list was last compiled in early 2017. It will be periodically reviewed and updated as new versions of ThinkHazard! are developed.

5 Other tool features

5.1 Downloadable hazard reports

Hazard level information, map, risk recommendations, further resources and contact information for all hazards and a single location can be downloaded as a .pdf document. The ‘download pdf’ button is available on all overview and hazard pages (see Figure 7, top right of screenshot).

5.2 User feedback

User feedback is a vital component of ongoing improvements and updates to ThinkHazard!. Users can provide feedback on any topic concerning the tool, via a link to a feedback form, available in the header of the user interface.

Feedback is provided via a short series of questions in a Google Form. The administrator notified, and will action any required changes and log requests for new features, responding to the feedback provider as appropriate.

Users can propose additions or amendments to resources or hazard data, and can highlight where hazard levels do not match expectations. If the feedback concerns new data for use in the tool, the administrator will follow up to review the data suitability for ThinkHazard!.

The screenshot shows a Google Form titled "ThinkHazard! User Feedback". The form includes a descriptive text about ThinkHazard! being under development and welcoming suggestions. It features a section for "Type of feedback *", which includes a list of options such as "I have a suggestion for including new hazard data", "The hazard levels do not appear to be correct", and "Amendment or suggestion for recommendations". There is also a "Name *" field and a "Type of Organization" field at the bottom. The form is set against a background image of a world map.

Figure 8 Google form used to obtain user feedback

5.3 Languages

ThinkHazard! is currently available in three languages: **English, French and Spanish**. All components of the tool are available in these languages, except for the feedback form (English only). The language toggle button is in the header of every page on the user interface. Translation has been conducted by World Bank translators, using Transifex.com.

5.4 Data map

A global map of data coverage is provided via a link in the header bar. This shows, for each hazard, the availability of different levels of hazard data. Clicking on a country provides a list of each available data set.

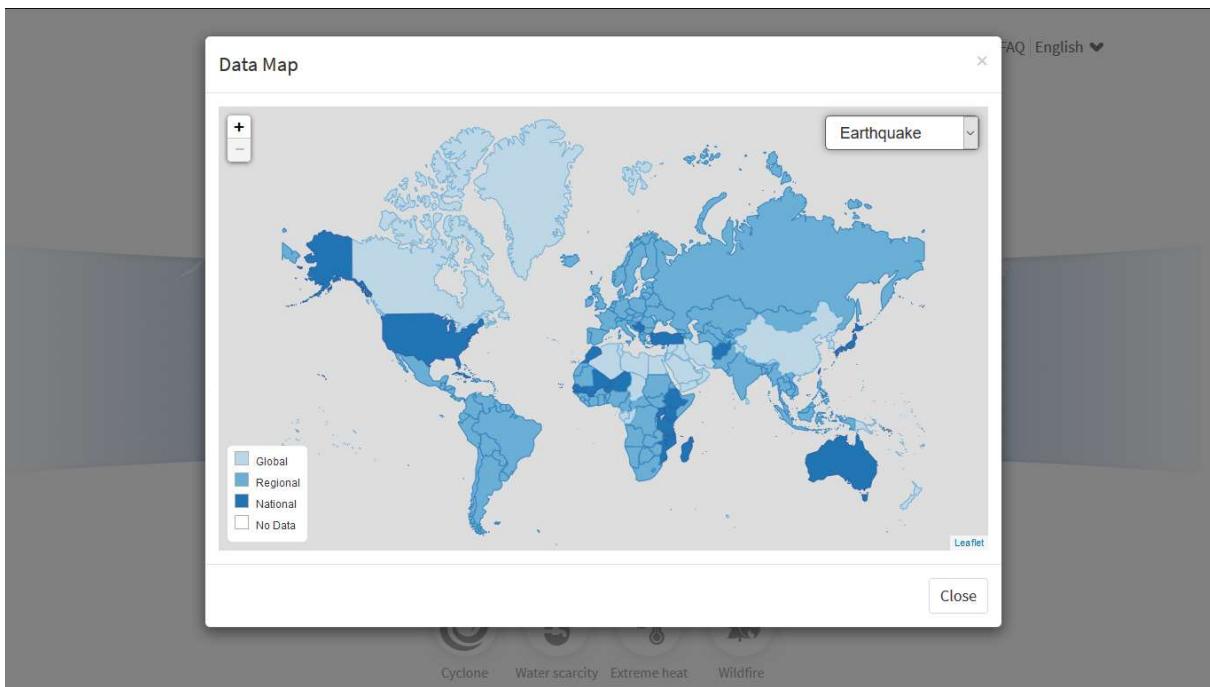


Figure 9 Global data map, showing the geographic level of data used in ThinkHazard, for each hazard and country

5.5 Frequently Asked Questions

A page of non-technical FAQs is available to users via a link on the header bar.

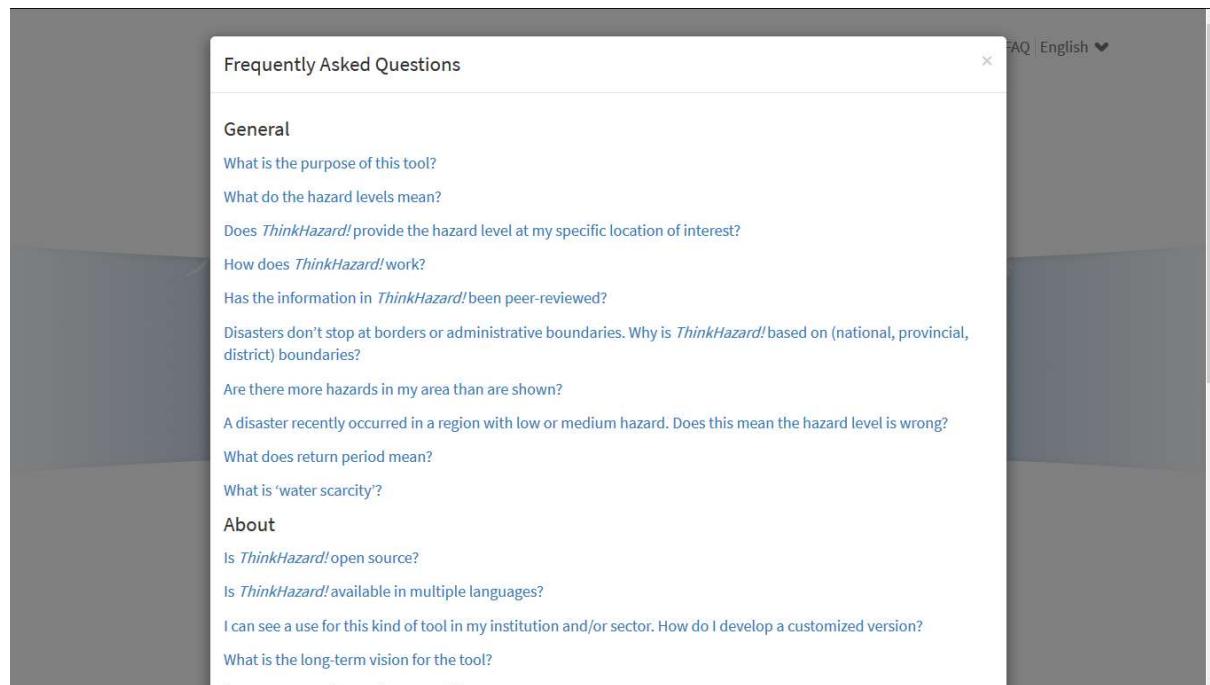


Figure 10 A Frequently Asked Questions section on the site provides information about most aspects of the tool, in non-technical language

5.6 Extracting and re-using content

The *ThinkHazard!* API enables extraction of hazard levels, recommendations, further information and contacts, into code for re-use in other tools. The API instructions are at [thinkhazard/API.md at master · GFDRR/thinkhazard](https://github.com/GFDRR/thinkhazard).

5.7 Licenses

ThinkHazard! is available under the [GNU General Public Licence, Version 3, 29 June 2007](#). Text content is licensed under CC-BY-SA. Classified hazard levels are licensed under CC-BY. Original hazard data are licensed under their original terms, which are contained in the associated layer metadata.

6 Architecture and technical workflow

6.1 Open source architecture

ThinkHazard! uses open-source code, available at <https://github.com/GFDRR/thinkhazard>. The code can be replicated and amended to develop versions specific to an organization or sector.

Potential amendments include including different hazards, or showing a subset of the present hazards. Recommendations may be tailored to a specific use (e.g., with a focus on managing risks in the agricultural sector). Additionally, the underlying method may be adjusted, for

example by changing the classification of hazard levels; if the intended focus of the tool related to minimum building standards, it may be appropriate to revise the intensity threshold at which damage can be expected to occur. It is also possible to link the user interface to a different set of layers and resources on a different GeoNode.

6.2 Database structure

ThinkHazard! uses two dedicated databases, populated with data harvested from a GeoNode. The *processing database* is used for data administration/preview, and the *public database* is used for visualization of information in the user interface.

Tasks performed using the *processing database* include: 1) harvesting and processing of hazard data from GeoNode; 2) modification of recommendations (content, order, hazard level link) via the *administrator interface*. A *preview interface* is available, for administrators to verify any updates to information (i.e., correct data and recommendations) before publication to the *public database*.

The *public database* is a copy of the *processing database* at a given time. To maintain performance of the user interface, data shown in the user interface are retrieved from the pre-calculated hazard levels stored in this database, rather than using on-the-fly data retrieval and processing. The *public database* is dropped and replaced by a new version each time the administrator decides to publish the data. At the same time, an archive of the database is stored in case of any future requests of data before the latest version. This archive is accessible to administrators only. The current database version number is presented on the About page (accessible via link in the user interface header, and on the printed .pdf report).

An update to the *public database* is required when:

1. Changes have been made to recommendations, contact information, or climate change statements
2. New data is available on the GeoNode, or changes have been made to existing data layers (including change of data content, or metadata).

Updates can only be made by administrators, using [prescribed commands](#).

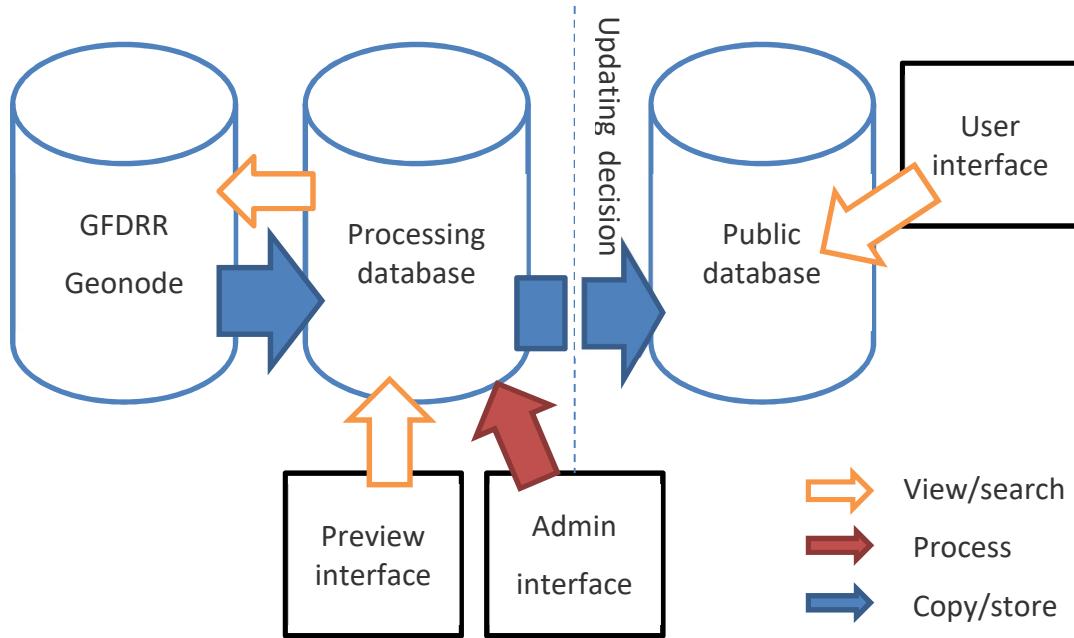


Figure 11 General architecture of the tool and back office workflow

6.3 Technical workflow – hazard classification

Hazard levels are not classified ‘on-the-fly’ (i.e., as a user requests them) because the processing time required would slow down the user search. Instead, classified hazard level is stored for each hazard against each ADM2, ADM1, and ADM0 Unit in the database. The user interface pulls data from that database upon a user requesting data in the user interface. This enables very fast operation of the user interface, but does require administrators to update the tool database when new data are available.

The technical workflow (contained in the *ThinkHazard!* processing code) identifies all hazard layers on GeoNode that are identified in metadata as being for use in *ThinkHazard!*. These layers are downloaded to the processing database, before being assessed for ‘completeness’. Completeness includes: 1) a valid hazard set must contain a data layer for each return period frequency threshold that is used in the classification; 2) data layers in a hazard set must all have matching spatial extent, origin, and resolution. The following hazard classification is then conducted:

1. All complete and previously unprocessed hazard sets are identified. These are datasets that have changed since the last execution of this process. Datasets that have already been processed are already in the production database, and are not reanalyzed.
2. For each hazard set:
 - a. All ADM2 units that intersect the data set bounding box are selected.

- b. The hazard classification is conducted one ADM2 unit at a time. Calculations are made on individual ADM2 units to avoid memory problems.
- c. For each selected ADM2 unit:
 - i. An extent corresponding to the current ADM2 (vector polygon) is read from the raster hazard dataset matching the return period used for the currently analyzed hazard level.
 - ii. If there are hazard intensity values in the selected extent, the ADM2 unit vector is converted to a raster with resolution to match the hazard raster. If there are no values, a value of 'No Data' is stored in the processing database for this ADM2 unit.
 - iii. This raster includes all pixels that intersect fully or partially, the ADM2 unit polygon (so that hazard pixels along the ADM2 unit boundary are accounted for). This raster is applied as a mask on the hazard dataset raster.
 - iv. If the data set is a pre-processed layer:
 1. The maximum hazard intensity value in the masked area of the data set is selected.
 2. This value is compared against the hazard intensity thresholds.
 3. The corresponding hazard level is stored for the ADM2 unit, in the processing database.
 - v. If the data set is not pre-processed (i.e. is probabilistic), the following process is performed with the thresholds corresponding to each hazard level (in the order of High, Medium, and Low):
 1. The maximum hazard intensity value in the masked area of the data set is selected
 2. This value is compared against the hazard intensity threshold for the High hazard level, on the return period data layer corresponding to the High Hazard frequency threshold.
 3. If the intensity threshold is exceeded, High hazard level is stored for the ADM2 unit, in the processing database.
 4. If the High hazard threshold is not exceeded, the process is performed on the Medium hazard threshold and data. If Medium hazard threshold is exceeded, Medium hazard level is stored for the ADM2 unit.
 5. If the Medium hazard threshold is not exceeded, the process is performed on the Low hazard threshold and data. If Low hazard threshold is exceeded, Low hazard level is stored for the ADM2 unit.
 6. If none of the three hazard levels are exceeded and there are intensity values, the value 'Very Low' is stored.
3. A decision is made on which dataset to present in ThinkHazard! (i.e. which values to store in the public database), if there has been more than one dataset available to classify hazard levels.
 - a. Outputs of the above process are prioritized according to:

- i. Data quality,
 - ii. Geographic status (local or global)
 - iii. Date of last update
- b. The dataset with highest priority is identified, and the hazard level for each hazard and ADM2 unit are passed through to the public database and take the first one for each ADM2 entity
 - c. Aggregation of ADM2 hazard level to ADM1 and ADM0 is performed:
 - i. Aggregation to ADM1 is conducted by identifying the maximum hazard levels in all its ADM2 'child' units
 - ii. Aggregation to ADM0 is conducted by identifying the maximum hazard levels in all its ADM1 'child' units
 - iii. For each ADM Unit (2, 1, 0) and hazard, hazard level is stored in the database

7 Administrator role

The *ThinkHazard!* administrator is responsible for identifying data and resources to include in *ThinkHazard!*, and curating the data in such a way that it conforms to *ThinkHazard!* requirements. This includes data content, format, and metadata, including units and data quality ratings. Hazard data and further resource curation is conducted via the GeoNode.

Administration also includes periodic review and update of any text content, using the administrator interface of *ThinkHazard!* (Figure 12): technical recommendations, climate change statements, and contact information using the online *Admin Interface* (a password-protected area of thinkhazard.org), or the data map and FAQs via static files in GitHub. The administrator is responsible for updates to translations, which have been conducted by professional translators and are supplied back to thinkhazard.org via [Transifex](#). The administrator is also responsible for monitoring website traffic (via Google Analytics) and User Feedback, which is submitted via Google Forms, with automatic notifications delivered to selected administrators.

A separate document, '*Administrator Workflow*' has been developed as a guide to these processes.

The screenshot shows a web-based application interface for managing technical recommendations. At the top, there is a navigation bar with links: Hazard sets, Hazardsets / admin div, Hazard categories, Technical recommendations (which is the active tab), Climate Change Recommendations, and Contacts. To the right of the navigation bar is a 'Preview' button with a small icon. Below the navigation bar, the main content area has a title 'Technical Recommendations' and a 'New Recommendation' button. The content is organized into several sections, each with a title, a detailed description, and 'Edit - Delete' links. The sections include:

- GOVERNMENT EXPERTISE:** Contact the governmental organizations responsible for management of landslides in the project country (e.g. ministry of environment, national geological survey and/or local authorities) to obtain more detailed information on areas previously affected by landslides and areas considered to be highly susceptible.
LS - MED, LS - VLO, LS - LOW, LS - HIG
[Edit](#) - [Delete](#)
- REGULATIONS:** Check with local authorities for any local regulations concerning wildfire hazard. Ensure that the project conforms to 1) any applicable wildfire land use planning regulations; 2) any applicable building regulations 3) any existing plans for warning and evacuation; and 4) any national laws, regulations and rules.
WF - VLO, WF - LOW, WF - MED, WF - HIG
[Edit](#) - [Delete](#)
- SEEK INFORMATION:** Conduct a risk assessment to further evaluate if your project is vulnerable to water scarcity/drought and whether a more detailed assessment and/or intervention should be considered.
DG - VLO, DG - LOW, DG - MED, DG - HIG
[Edit](#) - [Delete](#)
- EARTHQUAKE HISTORY AND HAZARD:** Get information about major earthquakes and secondary hazards (fires, landslides, liquefaction, tsunami in coastal areas) that have affected the project area in the past and the effects these caused. Community memory and historical accounts of earthquakes can provide useful information to supplement scientific studies. Contact the governmental organisations (e.g. Ministry of Environment and Geological Survey/ Ministry of Earth Sciences) responsible for earthquake risk management in the project country to obtain more detailed information on the potential earthquake hazard. Remember that low hazard does not mean no hazard.
EQ - LOW
[Edit](#) - [Delete](#)
- EARLY WARNING ACTION:** Ensure the project can act on volcanic early warnings. This may involve having a plan in place to mobilize in the event of a warning being received.
VA - HIG, VA - LOW, VA - MED
[Edit](#) - [Delete](#)
- INSURANCE:** For cyclone wind risks that cannot be mitigated, consider insurance products specifically aimed at alleviating the financial costs associated with wind risk.
CY - LOW
[Edit](#) - [Delete](#)
- INTERACTING HAZARDS:** Project planning, design, and construction practices should account for coastal flood and storm surge from cyclones and other weather events occurring in your project area.
CF - LOW
[Edit](#) - [Delete](#)

At the bottom of the content area, there is a note: **LOCATION ACCURACY:** The high level information available in this tool may indicate the presence of a flood hazard in your project area. Before committing significant resources to address a hazard, it is important to verify the location and severity of the hazard using more detailed methods.

Figure 12 Screenshot of the administrator view, showing the management of technical recommendations

8 Hazard specific classification methods

This section describes in more detail, the hazard classification process for each hazard. This section does not explicitly describe the data sources used for all hazards. The latest data sources can be viewed on the data map at thinkhazard.org and on geonode-gfdrrlab.org.

8.1 River Flood and Urban Flood Hazard Levels

Classification Summary:

The methodology for urban and river flood hazard classification was updated for ThinkHazard! version 2.

River flood and urban flood hazard are classified using a threshold of 'area flooded to damaging intensity threshold'.

The damaging intensity threshold is 0.5 m for both types of flood. The area threshold is 1% of the ADM unit for river flood, and 4% of the ADM unit for urban flood (due to the additional surface flooding included in urban flood, comparably larger areas are flooded)

Hazard is classified using frequency thresholds of 10, 50 and 1000 for high, medium and low hazard, respectively.

The full river and urban flood hazard classification methodology, developed by SSBN Ltd., is provided at the public [ThinkHazard! Methodology Google Drive folder](#).

8.1.1 Intensity

Flood hazard intensity is most commonly expressed as the water depth at a specific location, indicating severity of a flood. Flow velocity may also be given in high resolution local data but required a high level of detailed modeling. *Think Hazard!* is a global application, so water depth is used as the damaging intensity threshold. In determining a suitable intensity threshold, the European joint research project RiskMap⁴ recommends four water depth classes: <0.5 m; 0.5-1.0 m; 1.0-2.0 m; and >2.0 m. The values represent typical thresholds for which significant changes in damages on buildings occur:

- 0.5 m: Flood mitigation by sandbags and other preliminary measures are no longer possible. This is a typical height of tables and light switches.
- 2.0 m: The first floor and its interior are completely flooded.

A 0.5 m threshold is used; this provide a conservative damage threshold in line with the scope of *ThinkHazard!*. In version 1, global data used a threshold of 1.0 m due to concerns about uncertainties in the low-resolution elevation data leading to under- and over-estimation of depth at individual grid cells. Research by SSBN shows that at the ADM2 scale to which data are aggregated in *ThinkHazard!*, any errors seen in individual grid cells are largely cancelled out across the whole ADM2 Unit.

In development of version 1 there was over-estimation of hazard when using a 0.5 m or a 1.0 m threshold. This overestimation was due to the presence of at least one grid cell exceeding these depths in most ADM2 Units, which resulted in a majority of ADM2 Units being classified

⁴ Meyer et al. (2011)

as high hazard. After aggregating maximum hazard level to ADM1 and ADM0, most of the world was classified as high flood hazard. The updated methodology, considering the area flooded to a given depth threshold, overcomes the overestimation of hazard when basing the method solely on the presence of at least one grid cell exceeding the intensity threshold.

8.1.2 Frequency

The selected frequency thresholds for flood hazard are related to likelihood of experiencing flood in a human lifetime. This can be described as shown for the German Federal Office for Citizen Protection and Disaster Support (BBK, 2010; Table 3).

An alternative approach to using ranges of return periods, would be to interpolate between maps for different return periods to obtain a map for the required period. This requires spatial interpolation of hazard intensity, and uncertainties in the linearity of event boundaries and intensity distributions. Therefore, it was proposed to allow the algorithm to select a dataset within a range of return periods (100-1000 years). In version 2, the methodology explicitly uses the 1000-year return period.

The following frequency classes are used in *ThinkHazard!* version 2:

- **high:** 10-year return period
- **medium:** 50-year return period
- **low:** 1000-year return period

Table 3 Example of rationale justifying the choice of return periods, for river flood hazard

Name	Return period	Rationale
High	10 years	<p>In many cases a River Flood event of once per year or once every five years already causes considerable damage.</p> <p>NOTE: This return period may not be available in all flood hazard assessments. Therefore, it is proposed to allow <i>ThinkHazard!</i> to use 25 years as an alternative.</p>
Medium	50 years	<p>An event that would, on average, be expected to occur once or twice in a lifetime.</p>
Low	10,000 years	<p>An event most people will not experience and will only be remembered by previous generations.</p> <p>NOTE: Often the 10,000 year return period will not be available. It is proposed to use the highest available return period (and longer than 50 years) for the 'Low' hazard class. This can typically be between 100 and 1,000 years.</p>
Very low	Intensity not exceeded at the 'low' return period used.	No floods expected based on current climate, current models and data. However, some uncertainty remains.

8.1.3 Other notes on methodology

Areas of permanent water are masked in the method applied, so river channels and lakes are not considered in the calculation – these water bodies would contribute large areas that

exceed the intensity threshold, but due to their permanent nature cannot be considered as periodically flooded. The 3 arc-second (~90m) G3WBM mask (Yamazaki et al., 2015) based on multi-temporal Landsat images is used by SSBN to mask their 90m global flood product.

Urban areas were identified for the definition of urban flood. This was using a combination of two remotely sensed datasets indicating urban activity: the Global Urban Footprint (GUF)⁵ and the 2013 NOAA DMSP ‘stable lights’ (NTLD)⁶ datasets.

8.1.4 Results of classification

The maps below indicate the ADM2 classification based on **SSBN Ltd** 90 m global flood hazard maps, which is **used for global coverage for river flood and urban flood hazard in *ThinkHazard!* version 2.**

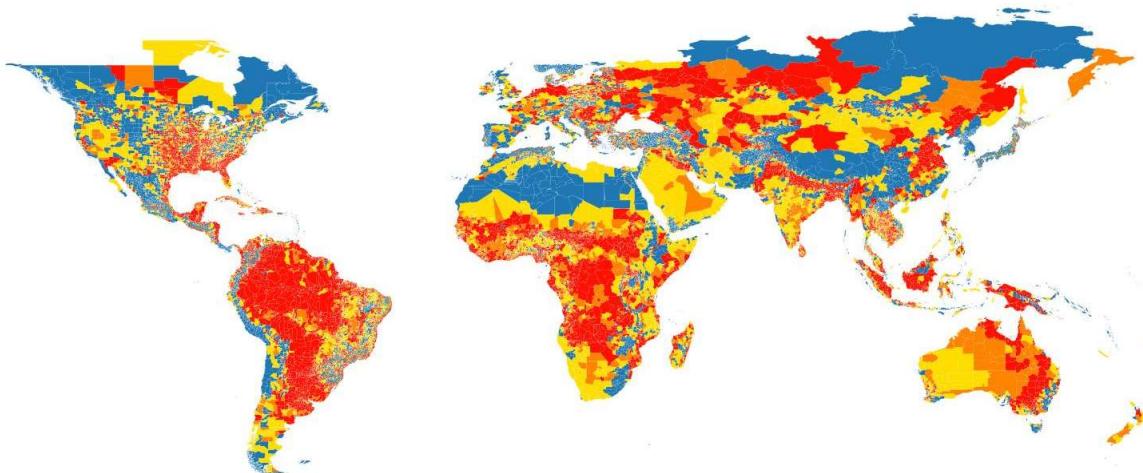


Figure 13 Final classification for river flood at ADM2 level: blue = very low; yellow = low; orange = medium; red = high

⁵ http://www.dlr.de/eoc/en/desktopdefault.aspx/tabcid-9628/16557_read-40454/

⁶ <https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>

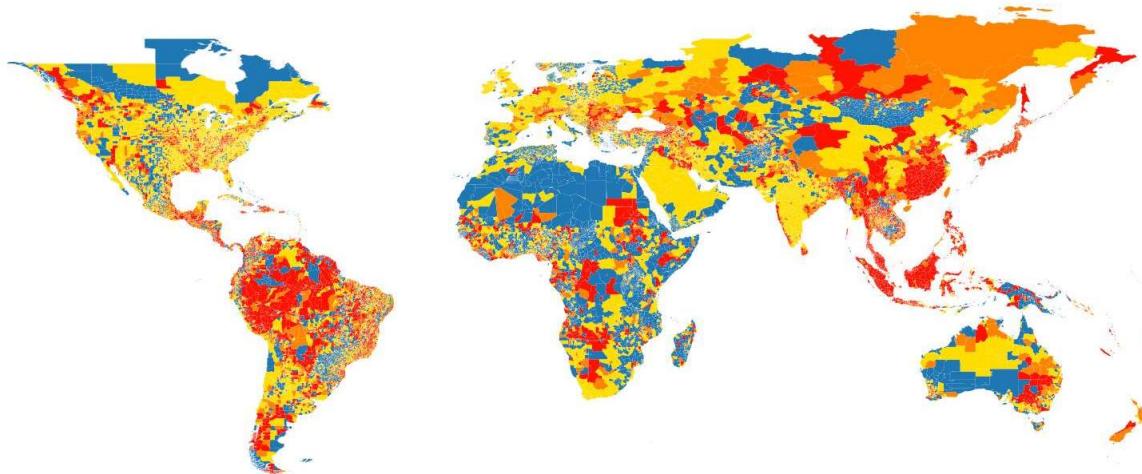


Figure 14 Final classification for urban flood at ADM2 level: blue = very low; yellow = low; orange = medium; red = high

8.2 Coastal Flood Hazard Levels

Classification Summary:

Coastal flood is classified using onshore flood depth data, provided as frequency-severity data in raster format. The methodology follows that described in section 2.4.

The damaging intensity thresholds are 2 m for high hazard, and 0.5 m for low and medium hazard. Hazard is classified using frequency thresholds of 10, 50 and 100 years for high, medium and low hazard, respectively.

Coastal flooding is caused by increased elevation of coastal waters above usual sea levels commonly caused by a combination of phenomena: Astronomical tide (natural cyclic sea level variation); storm surge during a storm or cyclone (raised water elevation due to low atmospheric pressure and accumulation of water due to wind); and wave set-up resulting from the energy transferred from offshore waves to the water column at coast. Storm surge hazard maps do not necessarily include the effect of tide and wave set-up, therefore may not encompass the full potential hazard.

The extent of inundated area and the depth of floodwater depend on intensity of the event and local topography. Extensive studies and simulation are needed to determine inundation parameters from wave data, since it requires an elevation model comprising both onshore topography and sea floor bathymetry. The ideal input to *Think Hazard!* is onshore inundation depth maps or inundation depth at the coastline to process the hazard categorization procedure. The former has been provided with global coverage by **Muis et al. (2016)**, who have conducted this type of analysis for coastal flooding; this dataset **provides the global coverage for coastal flood hazard in *ThinkHazard!* version 2**.

8.2.1 Intensity

Damaging intensity threshold is based on a similar rationale as for river flood. The values representing typical thresholds for which significant changes in damages on buildings occur are:

- For low and medium thresholds, 0.5 m: Flood mitigation by sandbags and other preliminary measures are no longer possible. This is a typical height of tables and light switches.
- For high threshold, 2.0 m: The first floor and its interior are completely flooded. More threat to life occurs with this threshold.

8.2.2 Frequency

The selected frequency thresholds for flood hazard are related to likelihood of experiencing flood in a human lifetime. The following frequency classes are used in *ThinkHazard!* version 2 (unchanged from version 1):

- **high:** 10-year return period
- **medium:** 50-year return period
- **low:** 100-year return period

8.3 Water Scarcity Hazard Levels

Classification Summary:

Hazard is classified using a Water Stress Index, which reflects the availability of water per person per year – a measure of water stress based on hydrological drought and water use.

Water scarcity is the only hazard in *ThinkHazard!* that uses an ‘inverse damaging intensity threshold’. All other hazards are classified based on intensity value being *exceeded*. Water scarcity becomes more serious as the water availability *decreases*, so a lower water availability per capita per year translates to higher hazard.

Hazard is classified as high, when water availability is $<500 \text{ m}^3\text{capita/yr}$ at the 5-year return period. Hazard is classified medium when water availability is $<1000 \text{ m}^3\text{capita/yr}$ at the 50-year return period. Finally, hazard is classified low when water availability is $<1700 \text{ m}^3\text{capita/yr}$ at the 100-year return period.

Discussion is ongoing over how to best represent drought in *ThinkHazard!* – whether meteorological, hydrological, or agricultural drought. From version 1, it was chosen to represent drought in terms of water scarcity, using a **global dataset of Water Crowding Index (WCI) (Veldkamp et al., 2015)** based on water availability per capita (Falkenmark et al., 1989). This index is based on estimates of water requirements in the household, agricultural, industrial and energy sectors, and the needs of the environment.

Water availability data are available as grid rasters and are summarized per water province (a combination of catchments and administrative areas), as well as for several return periods (2, 5, 10, 25, 50, 100, 250, 500 and 1000 years). Water availability is determined over these timeframes from an ensemble of five global circulation models (CMIP5), covering the historic period 1975-2004. This was combined with the population density in 2010 to determine the water availability per capita or Water Crowding Index (WCI) for the current situation.

Assumptions include: 1) This data includes surface water (rivers and lakes) and not soil moisture or groundwater; 2) Aggregation of effects to water provinces includes the effects of flow within a water province and the assumption that it is possible to mitigate drought by

distributing water within a province (but not between provinces). Real water stress risks may be lower than represented by these data due to these assumptions.

The rationale for the categorization choice is given in Table 4. Based on the return periods delimiting each category, the “best case” chance (1/best case Return Period) could be obtained.

8.3.1 Intensity

It was estimated that 1700 m³ of renewable water resources per capita per year was a threshold, below which a country would experience water stress. For *ThinkHazard!* the threshold of <500 m³/capita/yr is used for high hazard, as it represents absolute water stress. Thresholds of <1000 m³/capita/yr and <1700 m³/capita/yr are used for medium and low hazard level, as these represent severe water stress and moderate water stress, respectively.

8.3.2 Frequency

For each category of water scarcity (<500, <1000, and <1700 m³/capita/yr) and return period of original data, the average WCI in each ADM2 Unit was calculated. A trend line and corresponding (power) equation was fitted to this data, to obtain an equation with which the average return period of water stress could be calculated. The average return period and ‘best-case chance’ (i.e. longest possible return period) are related much like average and standard deviation. Both indicate how often water scarcity will occur in a certain water scarcity category, but the former represents the average chance of water scarcity occurring. The other represents the chance of a drought event occurring in the “best case” scenario where an area experiences water scarcity much less frequently than the average country of the same classified hazard.

The following frequency classes are used in *ThinkHazard!* version 2 (unchanged from version1):

- **high:** 5-year return period
- **medium:** 50-year return period
- **low:** 1000-year return period

Table 4 Water scarcity risk category based on return period and its rationale, based on human experience

Name of category	Best Case Return Period	Average Return Period	Best-case Chance	Average Chance	Rationale
High	5 years	2 years	20%	50%	This category represents areas where water scarcity is common but do not occur every year. Many countries in this category experience water stress every other year.
Medium	50 years	25 years	2%	4%	The 50 years return period was selected to represent water scarcity that occur once or twice in a lifetime. It corresponds to a large increase in number of water provinces which fall in this category.

Low	1000 years	250 years	0.1%	0.4%	Water provinces and countries in this category experience droughts less than once in a human life time, but they may occur occasionally.
Very low	>1000 years	1×10^{24} years	0%	0%	In this category no water stress is expected based on longest return period under the current climate, current models and data. However, some uncertainty remains.

For each water province and available return period it was then calculated whether Falkenmark's threshold of 1700 m^3 per capita per year was met. In other words, for each water province and return period it was determined if water stress would occur. Figure 15 shows the number of countries with a WCI below 1700 m^3 of renewable water resources per capita per year for several return periods. This graph is used as basis for the categorization.

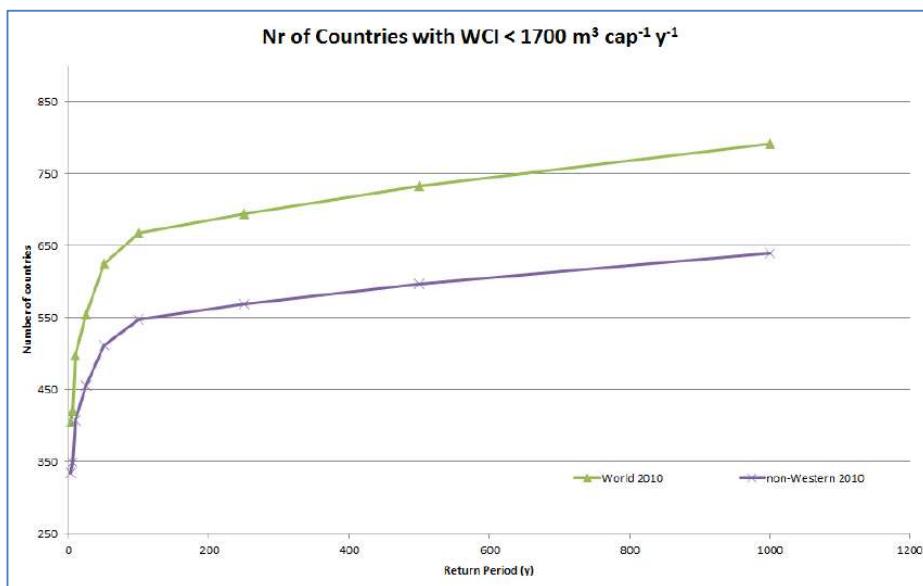


Figure 15 Graph representing the number of countries with a WCI below 1700 m^3 per capita per year in relation to the return period

8.3.3 Results of classification

The resulting water scarcity map per water province and per ADM2 Unit is shown in the figures below. The analysis was made on water province level first, because it is assumed that within a water province (a combination of watershed and countries borders) distributing of available water will occur towards water scarce areas when needed.

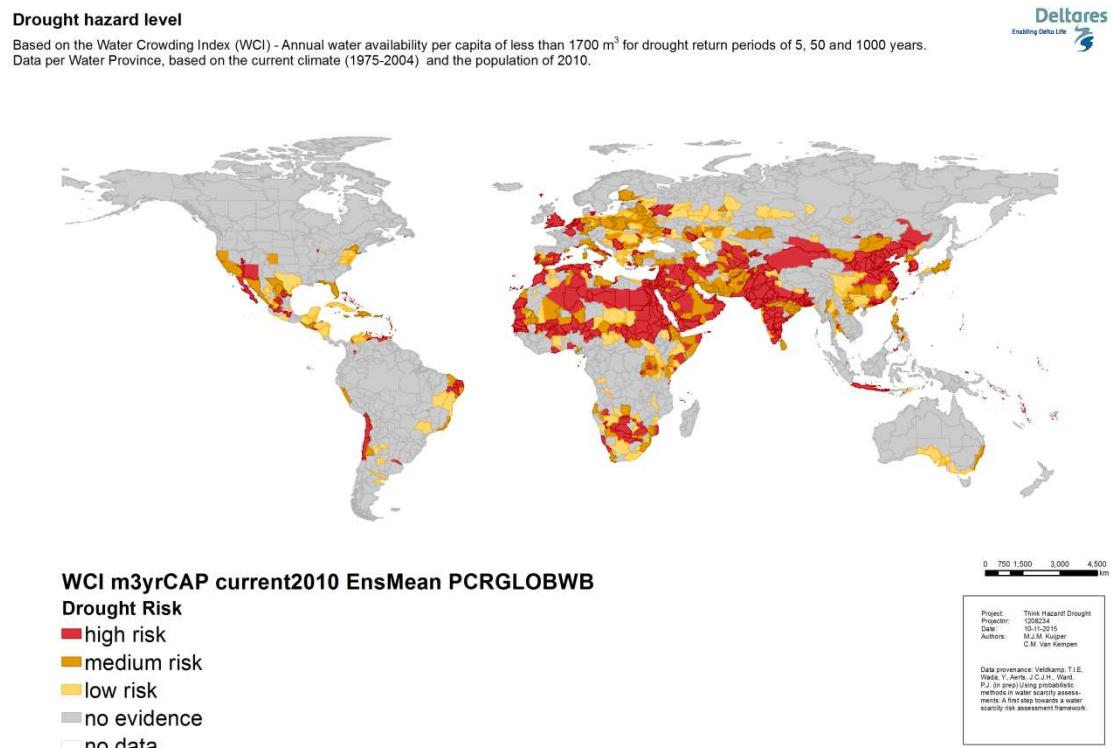


Figure 16 Water Crowding Index – Annual water availability per capita by water province

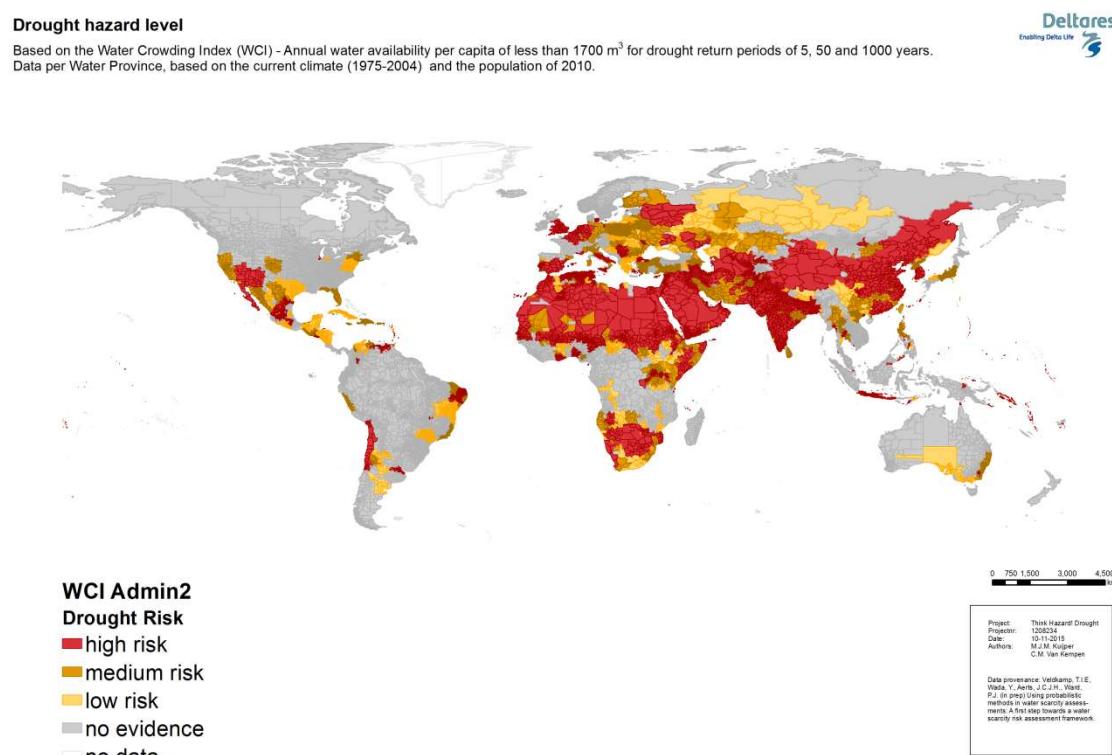


Figure 17 Water Crowding Index – Annual water availability per capita by ADM2 Unit

8.4 Tropical Cyclone Hazard Levels

Classification Summary:

Tropical cyclone is classified using wind speed, provided as frequency-severity data in raster format. The methodology follows that described in section 2.4.

The damaging intensity threshold is 80 kilometers per hour for all hazard levels, using frequency thresholds of 50, 100 and 1000 years for high, medium and low hazard, respectively.

Cyclone data used for hazard classification is the UNISDR Global Assessment Report 2015 (GAR15)⁷. Based on the metadata, the “tropical cyclonic strong wind and storm surge model uses information from 2594 historical tropical cyclones, topography, terrain roughness, and bathymetry”. Topography was taken from the Shuttle Radar Topography Mission (SRTM) of NASA, which provides terrain elevation grids at a 90 meters resolution.

The dataset covers the return periods 50, 100, 250, 500 and 1000 years, containing the peak wind velocity in kilometers per hour (km/h). No return period more frequent than 50 years is available from GAR15. Note that this data does not include extra-tropical cyclone winds, for example those that affect Western Europe. Also note that this data does not include the impact of cyclone-induced storm surges. Storm surge is available in another GAR15 dataset, but is represented in *ThinkHazard!* as one component of a coastal flood dataset (section 8.2).

8.4.1 Intensity

The intensity of cyclones is described by the wind speed, e.g., Figure 18. Based on literature review, an intensity threshold of 80 km/h is applied, which corresponds to the 50-60 miles per hour (MPH) hurricane warning threshold applied by NOAA (U.S. National Oceanic and Atmospheric Administration). The intensity threshold also corresponds to the Beaufort scale 9, described as “strong/severe gale – [at which] first damages occur”.

⁷ <http://www.preventionweb.net/english/hyogo/gar/2015/en/home/index.html>

Hurricane Winds

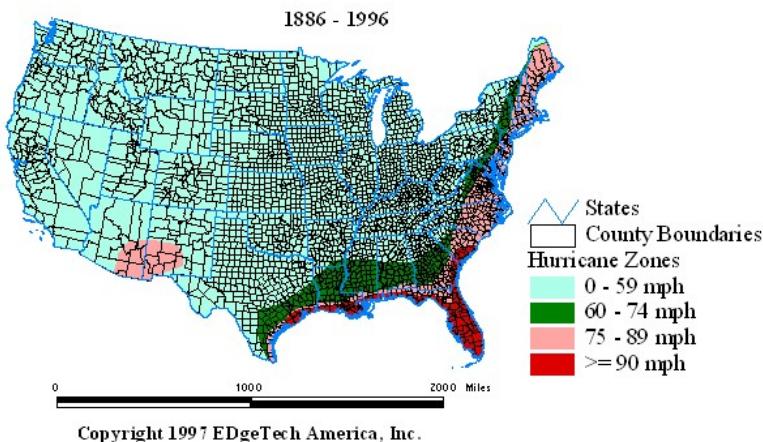


Figure 18 Hurricane zoning in the United States, ranging from 60 miles per hour (first damages) up to higher than 90 miles per hour (most severe damages) (Source: <http://www.edgetech-us.com/map/MapHurriZone.htm>)

8.4.2 Frequency

For cyclones, there are no community standards with respect to the frequency classes or return periods. Therefore, it is recommended to use the frequency classification as applied for river flood, based on human experience. However, as there is no global dataset with a return period more frequent than 50 years, the classification is as follows:

The following frequency classes are used in *ThinkHazard!* version 2 (unchanged from version1):

- **high:** 10-year return period
- **medium:** 50-year return period
- **low:** 1000-year return period

8.4.3 Results of classification

Figure 19 shows the results of cyclone hazard classification based on GAR15 data.

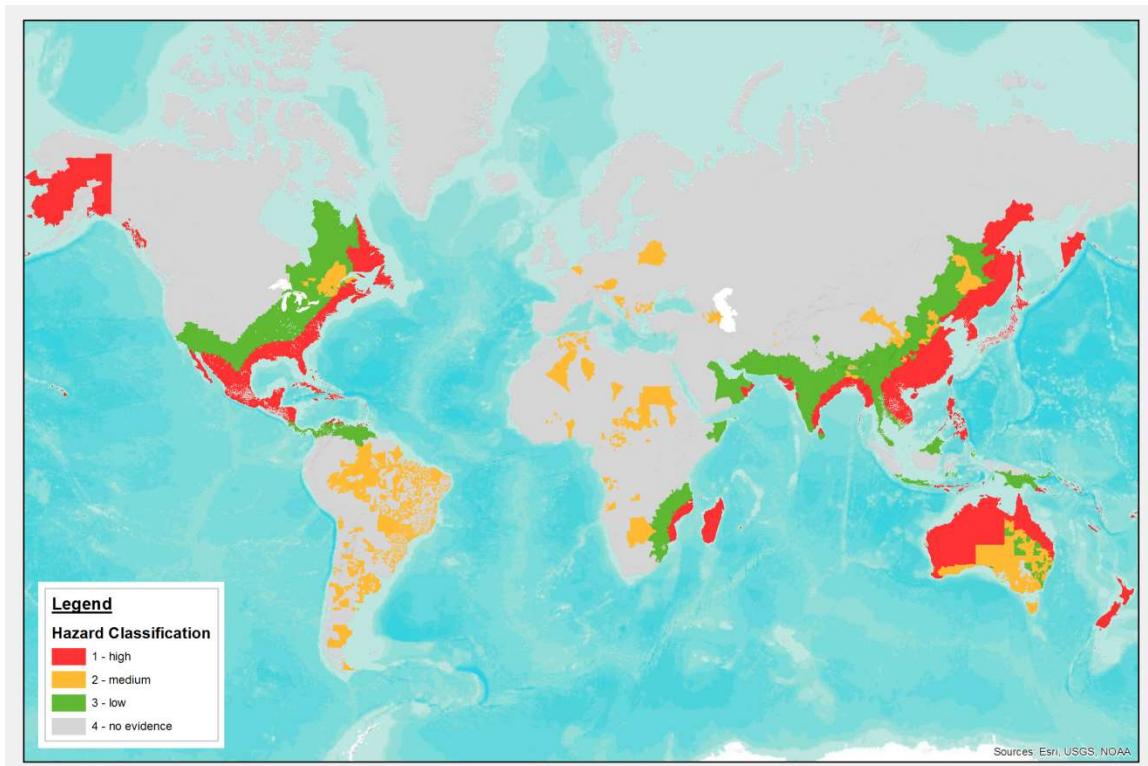


Figure 19 Preliminary cyclone hazard classification for ADM2 Units using GAR15 global data

8.5 Extreme Heat Hazard Levels (new in version 2)

Classification Summary:

Extreme heat hazard classification is based on daily maximum Wet Bulb Globe Temperature, provided as frequency-severity data in raster format. The methodology follows that described in section 2.4.

A specific temperature threshold is defined for each hazard level, at the 5, 20, and 100-year return periods, as described below. **The full river and urban flood hazard classification methodology, developed by VITO, is provided at the public [ThinkHazard! Methodology Google Drive folder](#).**

Extreme Heat hazard is classified based on an existing and widely accepted heat stress indicator, the Wet Bulb Globe Temperature (WBGT, in °C) – more specifically the daily maximum WBGT. The WBGT has an obvious relevance for human health, but it is relevant in all kinds of projects and sectors, including infrastructure related, as heat stress affects personnel and stakeholders, and therefore the design of buildings and infrastructure. In general, the WBGT is a relevant enough proxy to quantify the strain on physical infrastructure (energy, water, transport), such as increased demands for water and electricity, which may also affect decisions related to infrastructure.

8.5.1 Intensity

Heat stress studies in the scientific literature that make use of the WBGT apply thresholds of 28°C and 32°C to categorise heat stress risk. The damaging intensity thresholds are applied

following this definition of slight/low (<28°C), moderate/high (28-32°C) and severe/very high (>32°C) heat stress:

- high: >32°C
- medium: 28-32°C
- low: >25°C
- very low: <25°C

8.5.2 Frequency

There are no standard return periods used in research or engineering design concerning extreme heat events. However, most scientific studies use a 20-year return period in analyzing extreme heat events. This return period is included and because of the consistency with existing literature, results for this return period can be easily compared and verified with previous studies. A short return period (5 years) reflects more frequent extreme heat events, and the longest return period that can be generated based on the 30 years of available daily maximum WBGT data is 100-year return period. For longer return periods, the uncertainties in projected intensity become too large due to the inherent uncertainties in the statistical processing of the input data.

The following frequency classes are used in *ThinkHazard!* version 2:

- high: 5-year return period
- medium: 20-year return period
- low: 100-year return period

8.5.3 Results of classification

The classification of hazard at ADM2 Units based on a global probabilistic WBGT extreme heat dataset developed by VITO specifically for *ThinkHazard!*, is as follows:

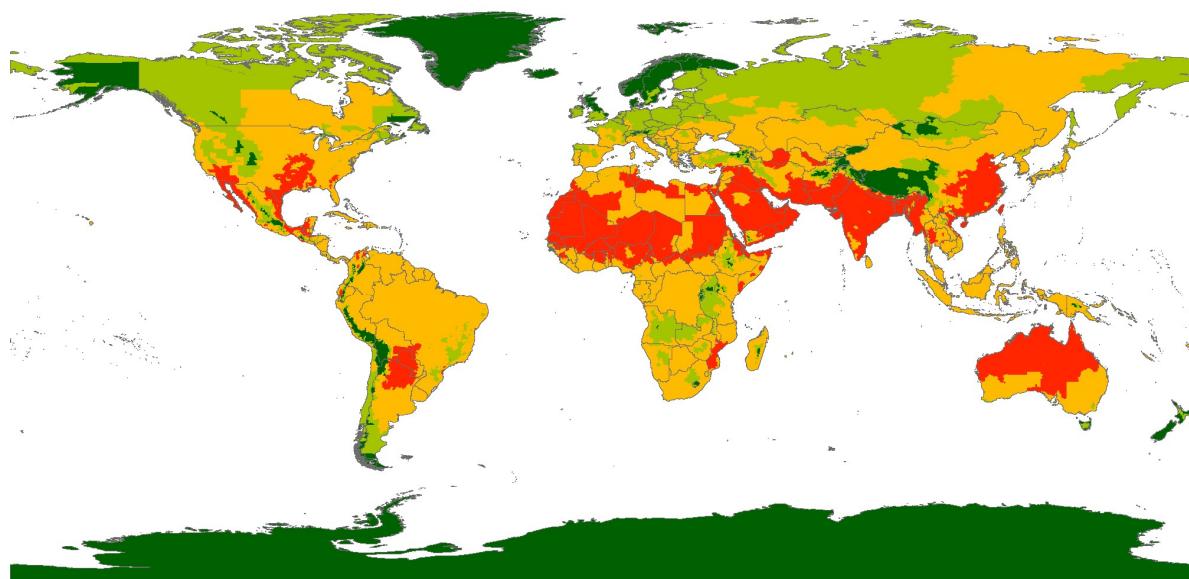


Figure 20: Heat hazard classification map at the ADM2-level. The figure visualizes the location with very low (dark green), low (green), medium (orange) and high (red) extreme heat risk.

8.6 Wildfire Hazard Levels (new in version 2)

Classification Summary:

Wildfire hazard classification is based on a probabilistic dataset generated from the daily Canadian Fire Weather Index, provided as frequency-severity data in raster format. The methodology follows that described in section 2.4.

A specific temperature threshold is defined for each hazard level, at the 5, 20, and 100-year return periods, as described below. **The full river and urban flood hazard classification methodology, developed by CSIRO, is provided at the public [ThinkHazard! Methodology Google Drive folder](#).**

The Canadian Fire Weather Index (FWI) is the most widely used fire danger rating system in the world. It integrates different weather factors that influence the likelihood a vegetation fire will ignite and spread. The index is based on instantaneous measurements (or prediction) of temperature, relative humidity and wind speed. In addition, longer term precipitation history is used to assess the dryness of fuels available for fire spread. It is important to note that the FWI does not consider the amount of fuel present or the likelihood of an ignition source being present. The FWI is most commonly used as the basis for fire hazard warning systems over large areas.

8.6.1 Intensity

Although the Canadian Fire Weather Index (FWI) has been widely adopted to rate levels of hazard, in practice the way it is applied to inform fire management varies greatly, according to regional differences in fire risk. For example, the numerical value of FWI considered high in Europe, would generally be considered benign in fire-prone areas like Australia.

The general, classification of hazard levels first requires identification of what constitutes extreme in the local context. Sub-classes of hazard below extreme are then defined based on perceived characteristics of risk to life, property and the difficulty of suppression. The definitions specified below are considered as a descriptive term appropriate to ThinkHazard! for each hazard level:

- High: wildfire is difficult to control and often spread over large areas;
- Medium: significant probability of fire but controllable in certain circumstances;
- Low: wildfire spread is possible but does not typically cover large areas and is readily suppressed; and
- Very Low: little to no fire spread possible.

Damaging intensity thresholds of FWI value have been constrained using the range of values used in different regions of the world to define fire severity. These vary substantially but extreme fire is consistently indicated by FWI ≥ 30 (see VITO methodology for the regional comparison). Testing of various FWI thresholds was conducted, and based on defining a suitable distribution of hazard in terms of 1) comparison to global distribution of previous fire occurrence, and 2) a meaningful distribution of hazard levels (particularly to distinguish areas of low and medium hazard) in *ThinkHazard!*, it was decided to apply damaging intensity thresholds as follows:

- high: FWI > 30

- medium: FWI >20
- low: FWI >15

8.6.2 Frequency

A fire event requires appropriate fire weather to support its spread but such weather conditions do not guarantee that a fire occurs. This fact influences the appropriate fire weather RPI for a given level of hazard. Given this, and evaluation of the spatial distribution of resulting global hazard the following frequency classes are used in ThinkHazard! version 2 (qualified with descriptors):

- high: 2-year return period – likely frequent extreme fire weather;
- medium: 10-year return period – extreme fire weather likely within the design life of the project;
- low: 30-year return period, extreme fire weather possible within the design life of the project.

8.6.3 Results of classification

The classification of hazard at ADM2 Units based on a global probabilistic FWI dataset developed by CSIRO specifically for *ThinkHazard!*, is as follows:

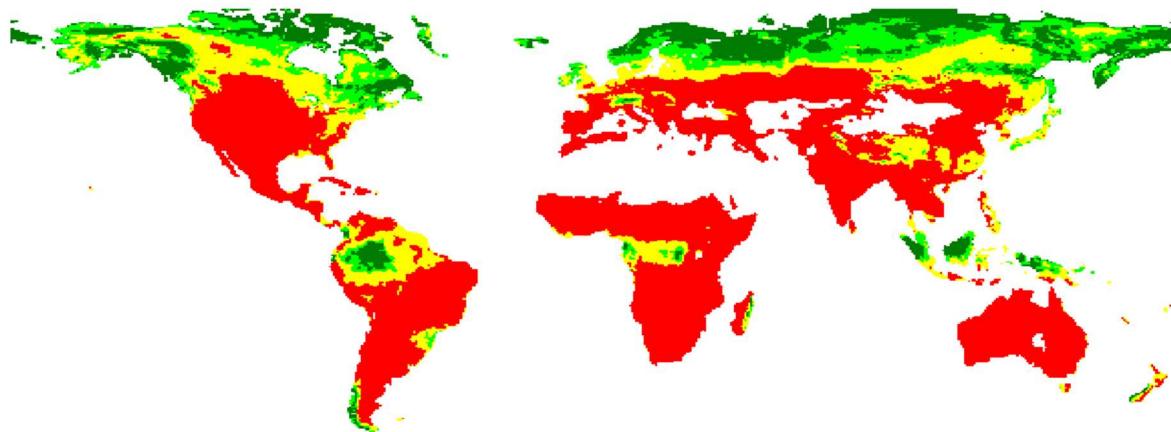


Figure 21: Classification of fire hazard (dark green = very low, light green = low, yellow = medium, red = high)
based on variable thresholds of FWI and return period intervals of 2, 10 and 30 years.

8.7 Earthquake Hazard Levels

Classification Summary:

Earthquake hazard is classified using peak ground acceleration (PGA, representing ground shaking), provided as frequency-severity data in raster format. The methodology follows that described in section 2.4.

The damaging intensity threshold is 0.2 g for high hazard and 0.1 g for medium and low hazard, using frequency thresholds of 100-250 years, 475-500 years, and 1000-2500 years for high, medium and low hazard, respectively.

Earthquake hazard maps generally consist of a grid of the expected peak ground acceleration (PGA) with a 10% chance of being exceeded in a 50-year interval, which translates to a return period of 475 years. Fifty years is often considered as a standard lifespan for infrastructures. For example, the following map was recently published by the SHARE European project (Figure 22, <http://www.efehr.org:8080/jetspeed/portal/hazard.psml>). The 475-year return period is a typical choice for seismic design codes for normal buildings whereas longer return periods are used as the basis of critical infrastructure such as bridges or dams (nuclear installations use even longer return periods, e.g. 10 000 years).

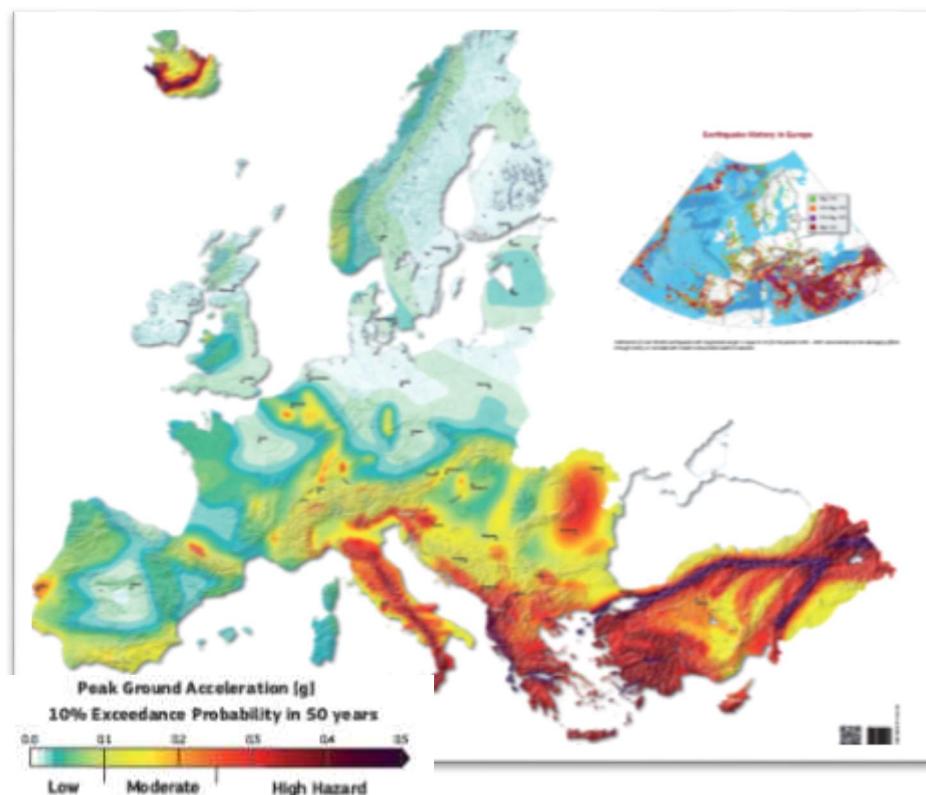


Figure 22 Example seismic hazard map (source: FP7-SHARE project)

ThinkHazard! uses UNISDR GAR15 data to provide global coverage for earthquake hazard, but across the world relies on many regional and local datasets available in the GeoNode to provide higher resolution hazard data with fewer global assumptions. Where available, the higher resolution datasets are utilized in the hazard classification.

8.7.1 Intensity

The severity of earthquake impact is commonly measured according the effects of the shaking on humans and structures. For example, the European Macroseismic Scale (EMS-98) ranks earthquakes on a 12-degree scale from “not felt” to “completely devastating”, each **intensity** degree denoting how strongly an earthquake affects a specific place. Intensity is subjective, determined according to post-disaster survey, from the observed effects of the earthquake.

The European Macroseismic Scale⁸ (EMS98), ranks seismic shaking according to effect on people, buildings and the environment. An event of EMS VI is considered to cause ‘slightly damaging’ effects to structures (Table 5), such as fine cracks in plaster, and can be felt by most people. Shaking with intensity VII result in stronger effects: people are frightened, cracks appear in buildings, and chimney start collapsing. According to widely acknowledged correlations⁹ between intensity and PGA, intensity VI corresponds approximately to a 0.1 g and intensity VII to 0.2 g. Some earthquake hazard data are available only in EMS or Modified Mercalli Intensity (MMI). In such cases, data are converted to PGA value using conversaion equations before storage on GeoNode and import to *ThinkHazard!*.

Table 5 Illustration of probability of damages for vulnerable buildings according to the intensity of seismic shaking (EMS98)

Degree of damage for vulnerable buildings				
Intensity VI	Many buildings	Some buildings		
Intensity VII			Many buildings	Some buildings

Hazard dataset provide seismic hazard different for PGA. To enable those datasets to be included in *ThinkHazard!*, several units are accepted in the processing algorithm, which can read PGA in terms of a decimal or percentage value of Gravity (*g*), or PGA in terms of SI units (e.g., gal or cm/s²) (see Table 6).

⁸ G. Grünthal, A Levret, European macroseismic scale 1998, Cahier du centre européen de géodynamique et de séismologie, vol 19, 2001.

⁹ Atkinson and Sonley, 2000

Table 6 Conversion table for units of earthquake data

Parameter – Unit	Thresholds	
	High	Medium and Low
PGA-g	0.2	0.1
PGA-g-per	20	10
PGA-gal	196.133	98.066
PGA-cm/s ²	196.133	98.066
PGA-m/s ²	1.961	0.981

8.7.2 Frequency

The earthquake field has standard frequencies for presenting earthquake hazard, for which maps are available from many projects. The standard used by research and engineers is to present 475, 975 and 2475 years. More commonly the 100, 250, 500, 1000, and 2500 years are used by the insurance industry. *ThinkHazard!* leverages these standards in setting the return periods to maximize the data that can be incorporated. *Preferred* values for return periods are:

- High: 1 in 100 years (10% chance the value is exceeded in 10 years, 50% in 50 years).
- Medium: 1 in 475 years (2% in 10 years, 10% in 50 years).
- Low: 1 in 2475 years (0.4% in 10 years, 2% in 50 years).

If 2475-year return period data are not available, 975-year is an acceptable value (5% chance of exceedance in 50 years) for the longest return period, since it still corresponds to a low probability of exceedance on the lifespan of common projects.

If 100-year return period data are not available, 250 years is an acceptable value (4% chance of exceedance in 10 years or 20% in 50 years) for the shortest return period, since it still corresponds to a high probability of exceedance on common projects lifespan. For medium return period, data can also be fund for 500-year return period. The 250 and 500-year return periods are often available in datasets produced for the financial sector.

8.8 Tsunami Hazard Levels

Classification Summary:

Tsunami hazard data at the scales appropriate for use in *ThinkHazard!* is generally presented as maximum wave amplitude or height, usually at 100 m water depth, provided as frequency-severity data in raster format. The methodology follows that described in section 2.4.

The damaging intensity threshold is 2 m for high hazard, 1 m for medium hazard, and 0.5 m for low hazard, using frequency thresholds of 100 years, 500 years, and 2500 years for high, medium and low hazard, respectively.

It is decided that a threshold of 0.5 m should be used for low hazard, as it represents a height at which land and buildings are flooded, and which allows for some uncertainty in offshore values of tsunami translated to potential onshore impacts. A value of 1.0 m for medium hazard is proposed to distinguish areas of known hazard e.g., Eastern Mediterranean Sea from areas of lower hazard (e.g., Western Africa). High hazard is classified using a higher threshold of 2.0 m, which corresponds to the depth at which building damage ratio increases significantly, based on post-disaster surveys in Japan, 2011 (MLIT, 2012).

GFDRR commissioned analysis by the Norwegian Geotechnical Institute, Geoscience Australia, and INGV Italy (under the umbrella of the *Global Tsunami Model*) to make data developed by Davies et al. (2017) openly available at several return periods. The dataset provides run-up values calculated at offshore hazard points, which have been projected to the shoreline by simple interpolation (described in the NGI report, in [ThinkHazard! Methodology Google Drive folder](#)). These coastal values have been rasterized at 0.01-degree resolution (c. 1km at the equator) for import to GeoNode and Thinkhazard!.

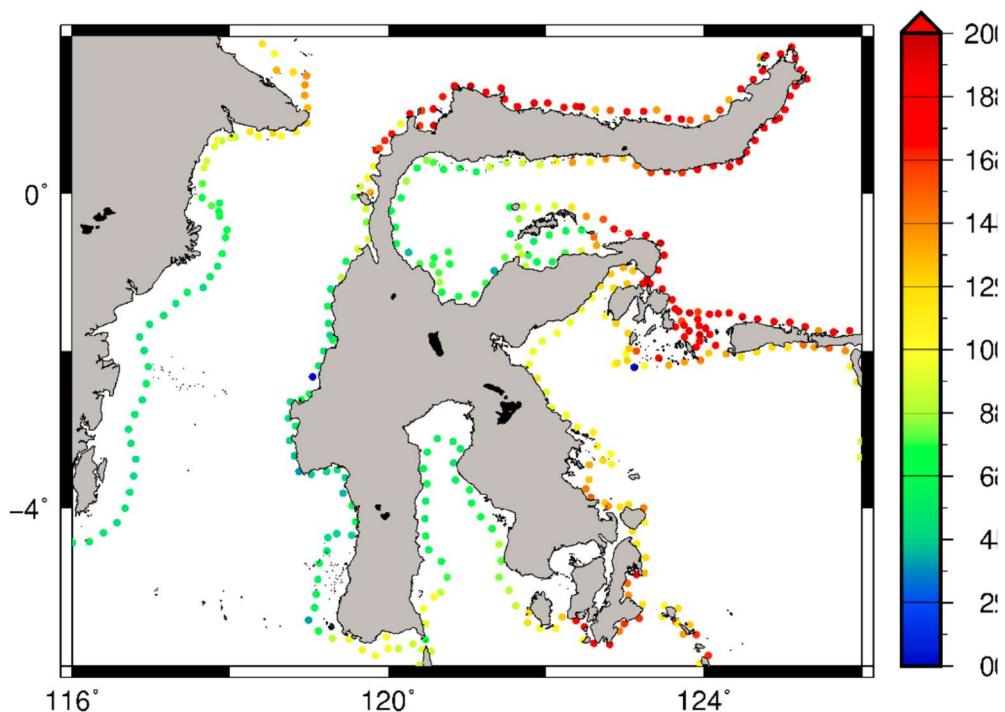


Figure 23 Tsunami maximum inundation height data at offshore points (output of Davies et al., 2017), for Sulawesi, Indonesia

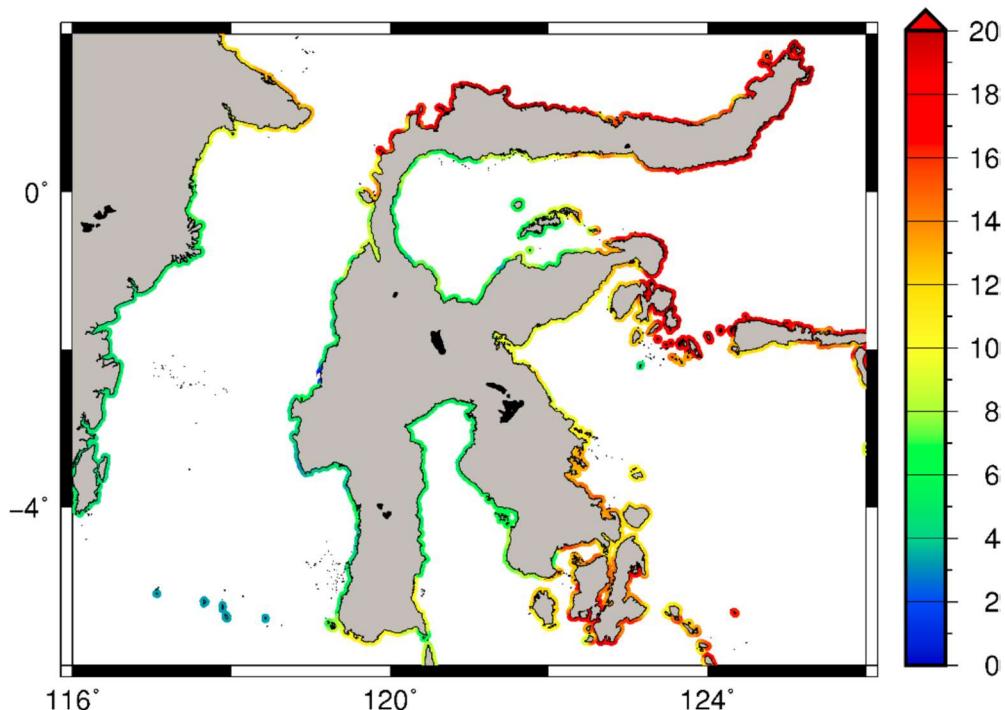


Figure 24 Tsunami maximum inundation height data at the coast, interpolated from the points offshore Sulawesi, Indonesia

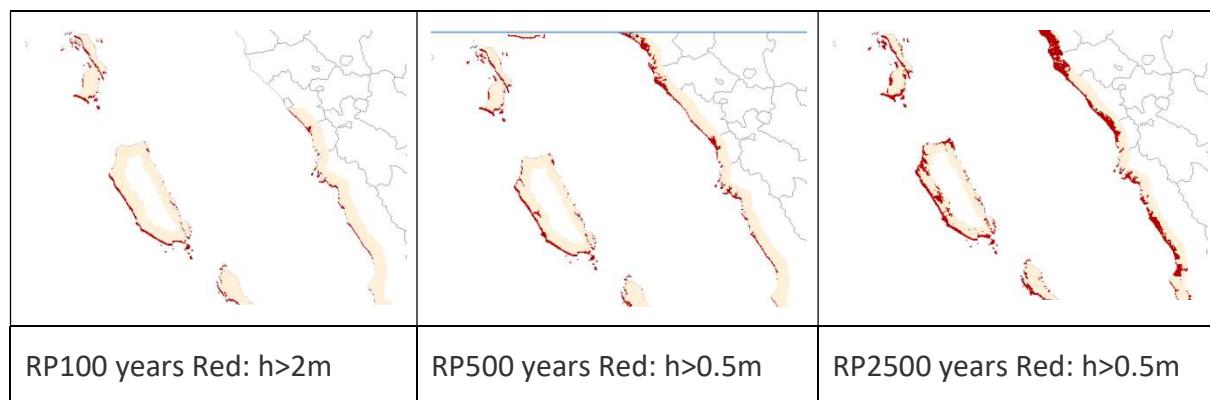
An alternative pre-processing approach can be used to better define potential areas of onshore inundation, which identify low-lying areas where tsunami may inundate far inland. This can lead to non-coastal ADM2 units being assigned a hazard level, which would otherwise not be identified in the global data. This approach relies on SRTM elevation data, which are not available for the entire domain of global tsunami data, therefore use of **this approach is limited to national or regional data**.

Due to the associated computational requirements and aggregated hazard levels presented by ThinkHazard!, it is not appropriate to conduct full inundation modelling to define onshore inundation height and extent. It is instead proposed to apply a simple ‘bathtub’ method which compares wave height from each point, to ground elevation within 10 km of the coastline. This distance is expected to incorporate the maximum inundation of any tsunami. This processing procedure is as follows:

- a. Extract the pixels covering the coast zone (to 10km inland) from SRTM30 tiles covering the tsunami data extent.
- b. Produce a buffer polygon around each tsunami data point, to a distance sufficient to intersect the first 10 km inland from the coastline. Assign the point data value to the corresponding polygon.
- c. Rasterize the polygon layer, aligned with SRTM raster cells, with maximum value of the buffered polygons assigned to the raster value.
- d. Perform a raster calculation; Inundation depth is calculated as tsunami height minus SRTM elevation. Where inundation depth is negative, it should be set to zero (the ground is not inundated).
- e. This process is repeated for each return period layer to produce one ‘bathtub’ inundation map per return period for import to ThinkHazard! (Table 7).

An alternative approach to this, is to use an attenuation-based approach, which applies a degree of attenuation to the tsunami data points, to better represent the potential height inland. The attenuation relationship used is derived from post-tsunami surveys of inundation height decreasing with distance inland (Leonard et al. 2008). Another alternative is to use tsunami run-up equations to estimate tsunami run-up inland. Both approaches would be best suited to local data, and retain a significant amount of uncertainty in the onshore heights produced. These approaches and their impact on hazard level may be explored in future.

Table 7 Example of pre-processed rasters for tsunami in Indonesia. Red zones show regions where wave height exceeds both elevation and the threshold on the buffered coastline



8.8.1 Intensity

Tsunami hazard is the potential for damage by massive volumes of water flowing onshore, with high velocity and depth, as tsunami waves. Damage occurs due to the force of waves, debris contained in the waves, floatation and scour of structures, and deposition of sediment/rocks. Tsunami waves are generated primarily by submarine earthquakes displacing the water column above, but also landslides or volcanic eruptions, which can displace large volumes of water if they occur underwater or at the coast. Small tsunamis can have very localized effects. The largest tsunamis, generated along subduction zones, can travel across the whole ocean affecting coastlines of multiple countries and inundating long distances and up to several kilometers inland, depending on topography.

Tsunami hazard data is commonly probabilistic and provides a view of hazard (generally wave height, rather than velocity) for multiple return periods. Data may be available as onshore inundation maps (typical of local analyses, e.g., for a city or province) or, more commonly for large areas such as national or global analysis, point data at locations offshore. Due to the computational overheads of simulating the nearshore and onshore flow of tsunami, data points are often spaced several kilometers (or tens of km) apart and are located offshore, e.g., at the 100m isobath (water depth). On the other hand, availability of inundation maps is likely to be limited, since models have very high technical and computational requirements.

8.8.2 Frequency

Tsunami hazard data do not have *de facto* standard frequencies. However, given tsunamis' primarily tectonic causes, the timescales used are commonly in the range of 100 to 2500-year return periods are used. Analysis commissioned for ThinkHazard! has enabled global tsunami data at 10, 50, 100, 200, 500, 1000, and 2500-year return periods to be openly available.

Return periods of 100, 500, and 2500-year are used to define high, medium and low hazard, respectively.

8.8.3 Results of classification

The classification of global data results in the following distribution of hazard levels:

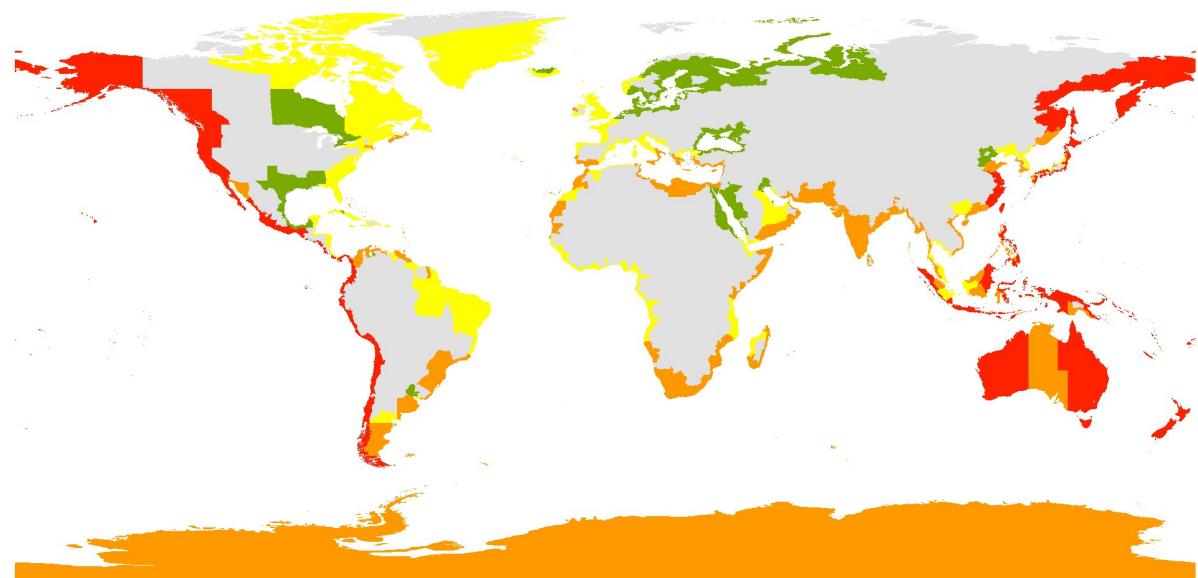


Figure 25 Global tsunami hazard levels at ADM1.

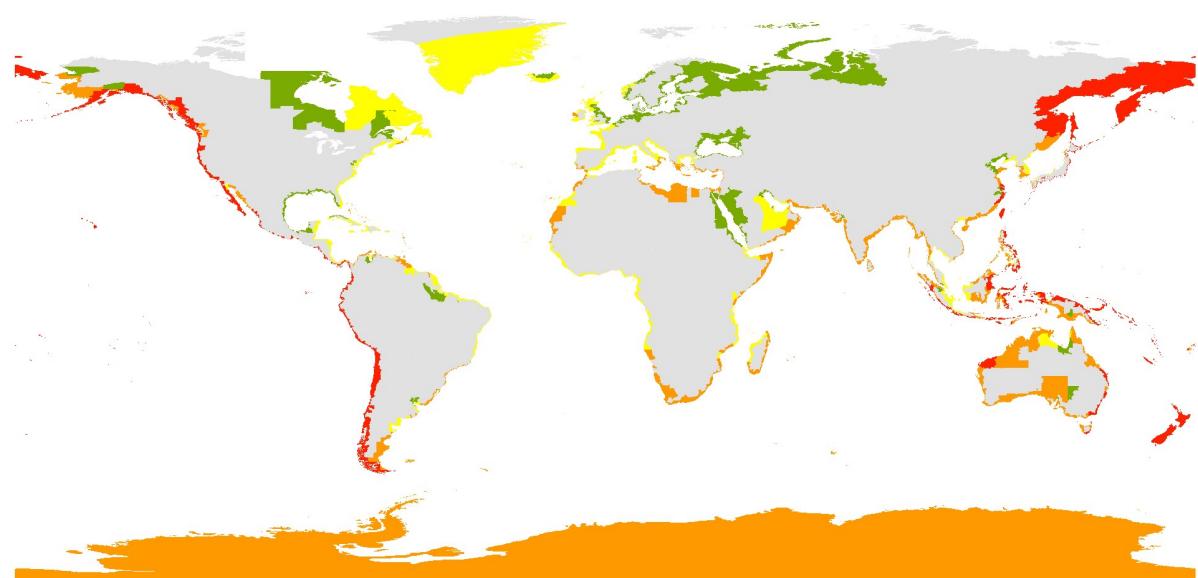


Figure 26 Global tsunami hazard levels at ADM2.

8.9 Volcanic Hazard Levels

Classification Summary:

Hazard levels are defined outside of *ThinkHazard!*, using a combination of historical eruption records: maximum recorded eruption size (Volcanic explosivity index) and last known eruption date).

A volcano is a place where magma comes to the surface, during volcanic activity. Volcanoes present potential threats to people and property, due to:

- Proximal hazards, affecting the area within 100 km around a volcano
 - Lava flows are very hot materials that can destroy the built environment and can release harmful gases.
 - Pyroclastic flows are a mixture of extremely hot rocks, sediment and gases that moves downhill, destroying most things it encounters and depositing large quantities of debris.
 - Lahars are flows of water, rock and sediment that destroy most things they contact, and deposit large quantities of debris. They can occur without an eruption during wet weather when rainfall mobilizes loose volcanic material.
- Distal hazards, affecting larger distances
 - Ash (or tephra) consists of small fragments projected in the air during eruption, which drop over large area. They cause health problems, disrupt services and agriculture, and the weight of ash (especially when wet) can cause roofs to collapse.
 - Gases released into the atmosphere.

Eruptions can be relatively rare events, and volcanoes can remain quiet for several hundreds of years between eruptions. For most volcanoes globally, there is little information to assess how frequently eruptions might occur, and how big they might be in an eruption. For some well-studied volcanoes this information is available through assessment of previous deposits. Ashfall modelling is gradually providing more data that shows potential depth of ash at multiple return periods, though a lack of eruption data introduces significant uncertainties to the outputs.

Volcanic hazard levels are defined outside of *ThinkHazard!*, using a combination of historical eruption records: maximum recorded eruption size (Volcanic explosivity index) and last known eruption date). Several global eruption databases record the coordinate location, dates, type and magnitude of past eruptions. The *Smithsonian Global Volcanic Program (GVP)*¹⁰ and *LaMEVE*¹¹ databases contain a volcanic eruption index (VEI)¹² which quantifies the eruptive magnitude of past events. These are the best global information from which we can classify volcanic hazard, however, they contain significant uncertainty. Volcanoes with no recorded eruption, or that has never erupted in living memory may pose a hazard. The most widely used measure of volcanic intensity is VEI. Volcanoes can display variable VEI in different eruptions, and VEI varies over time during the same eruption. The maximum eruption VEI per eruption are provided in the GVP and LaMEVE eruption databases. The maximum VEI at each volcano forms the basis of the intensity thresholds used in preprocessing. Not all volcanoes in these

¹⁰ <http://volcano.si.edu/>

¹¹ <http://www.bgs.ac.uk/vogripa/view/controller.cfc?method=lameve>

¹² VEI – Volcanic Explosivity Index – is measure of how powerful an eruption is. The scale is logarithmic, and ranges from non-explosive eruptions (0) to the largest eruption in history assigned VEI 8.

databases have an associated VEI value. Where this is the case, date of last eruption is used (section 7.2.3.1). Where VEI is available, VEI 3 is used as a threshold to define a non-explosive volcano (those with a $\text{VEI} < 3$), are rated as ‘Low’ hazard. Volcanoes with maximum $\text{VEI} > 5$ are rated High, and those with a maximum recorded $\text{VEI} 3\text{-}5$ are ‘Medium’.

The above is done in combination with consideration of the eruptive history (dates of previous eruptions). The most available, but basic, measure of eruption frequency is date of last eruption. Many volcanoes do not have a complete history of all eruptions because the average time between eruptions is can be several hundred or thousand years (these tend to be the volcanoes with more powerful eruptions) so it is difficult to reliably and consistently compare frequency across all volcanoes in the GVP database. Conversely, some volcanoes erupt with surprising regularity – monthly or even daily – with much smaller eruption and little impact. Last eruption date is only a very approximate guide to frequency of eruption, and further work should be done in future to better define volcanic hazard level for *ThinkHazard!*.

Considering the above, last known eruption data is used as a guide to eruption potential only if there is no VEI data for a volcano. If the last known eruption was over 10000 years ago, it is considered Low hazard; if the last eruption occurred within the last 2000 years ago, it assigned High hazard. Anything with a last know eruption in the intervening period is Medium.

The procedure is applied as follows:

- For each volcano of the GVM database, location, date of last known eruption and the maximum VEI index is extracted from the database
- Hazard level is associated to the VEI index value (when available):
 - If $\text{VEI} \geq 5$, then hazard level is **high**,
 - If $5 > \text{VEI} \geq 3$, then hazard level is **medium**,
 - If $\text{VEI} < 3$, then hazard level is **low**,
- If the VEI index value is not available, hazard level is associated according to the date of last known eruption:
 - If it was recorded an eruption in the last 2000 years (CE), then hazard level is **high**,
 - If it was recorded an eruption in the Holocene (last 10000 years), then hazard level is **medium**,
 - If it was recorded an eruption in more ancient times, then hazard level is **low**,
 - If no eruption of the volcano has been reported, then hazard level is **low**,

To account for the fact that damage from a volcano does not occur only at the vent, but several tens of kilometers around the vent, the hazard level of each volcano is applied to a circular area around the volcano coordinate location.

The maximum extent of proximal hazards is approximately 100 kilometers from a volcanic vent. This distance does not account for topographic influences that constrain the flow of lahars and lava. The resulting raster map of hazard levels provides a crude assessment of proximal volcanic hazard (excluding impacts of ash and gas), see Figure 27. This map is uploaded to the *ThinkHazard!* database. The tool then associates the hazard level to administrative units following the normal procedure: intersection with administrative polygons and maximum of hazard level on a given unit.

8.9.1 Results of classification

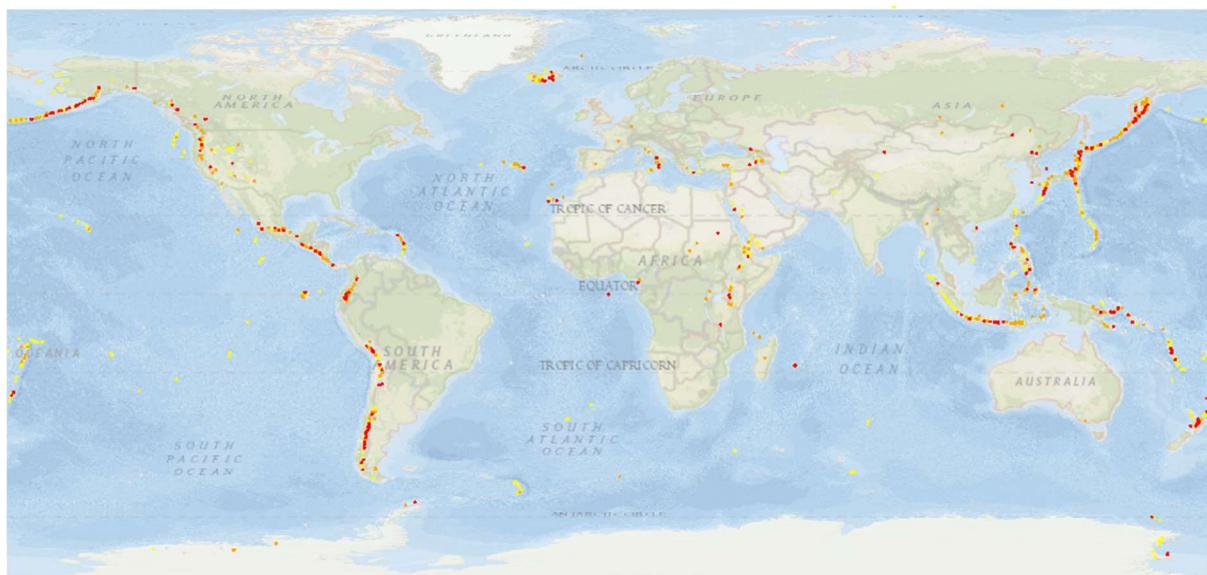


Figure 27 Hazard class at all locations within 100 km of a volcano. Red = high hazard, orange = medium, yellow = low. Very low and unknown hazard are not shown.

8.9.1 Future probabilistic classification

When available, probabilistically simulated ashfall data can be imported as maps of ash depth at multiple return periods, enabling hazard levels to be defined as described in section 2.4. However, such analysis are available for only a few regions, and are not typically represented at the multiple return periods required for *ThinkHazard!*. Work may have to be commissioned to achieve good coverage of volcanic ash hazard levels in *ThinkHazard!*.

Probabilistic volcanic data are available as raster containing ash depth (mm) at each grid cell, or as maps of isopachs indicating the limits of ash depth distribution, for several return periods. Recommended ash depth thresholds for hazard levels are: 0.5 mm, 10 mm, and 50 mm. These values are selected based on differential impact. Ash thicknesses of 0.5 mm can impact transportation by reducing visibility and obscuring road markings. At 10 mm, minor damage to buildings and infrastructure may occur through ash infiltration requiring extensive clean-up and this thickness of ash may cause agricultural productivity loss (<50%) and health implications. At 50 mm, major agricultural productivity loss (>50%) and damage to buildings and infrastructure (i.e., potential roof collapse), ash infiltration to buildings and health implications would be expected (GVM, 2016).

There is no industry standard frequency for presenting ash fall modeling. Frequency of eruption varies significantly for different volcanoes, but generally ash impacts are relevant presented at return periods greater than 100 years. A decision on recommended return period for probabilistic ash data would have to be taken on receipt of global ash modeling data, but it is expected that suitable return periods would include 100, 500, 1000, and 10000 years.

8.10 Landslide Hazard Levels

Classification Summary:

Landslide susceptibility or hazard index is reclassified into four hazard levels. Presently, this is done differently for each hazard data source (see below). Investigations are underway to make this more consistent across multiple hazard sources.

Large-scale landslide hazard maps generally present landslide frequency per grid cell, derived from GIS-based analysis of terrain conditions to produce a landslide susceptibility map, and the coincidence of landslide triggers. Landslide is a locally variable hazard, but the factors determining landslide susceptibility are well defined (e.g., Nadim et al., 2006). Terrain factors include slope, land or vegetation cover, soil type and geological conditions, and soil moisture. This is often presented as a ‘susceptibility index’, or when combined with triggering factors, (precipitation and seismic activity) this may be presented as a ‘hazard index’. In either case, hazard is generally classified into discrete categories describing the frequency which landslide may occur there. This is consistent with availability of regional and national landslide data too. Probabilistic landslide data are rare, and only available on a local basis, sometimes including analysis of landslide runout to establish potential impact. It is envisaged that *ThinkHazard!* Will rely on using national, regional, or global landslide hazard maps, which present frequency of landslides, rather than such high-resolution local datasets.

Two of the most prominent global landslide datasets are *NASA SEDAC*¹³ host the Global Landslide Hazard Distribution, v1 (CHRR, CIESIN and NGI, 2005), which presents landslide and avalanche hazard using an index of 6-10 (below 5 is considered negligible hazard). International Centre for Geohazards/NGI later produced a global landslide frequency map for the *UNISDR GAR2013*¹⁴ (NGI, 2013), based on a similar method using global terrain and trigger data. This dataset presents expected annual probability per grid cell, and percentage of the grid cell of a potentially destructive landslide event.

The available landslide hazard maps can be categorized into hazard level based on the frequency of occurrence, e.g., as used in the GAR13 data (table below), or based on the mapped hazard index, as supplied in SEDAC data. Given the variation in original data units and categories, it is necessary to apply a different processing method to each, while preserving a consistent conversion of landslide frequency into *ThinkHazard!* hazard levels. The categorization procedure applied to each of the available global datasets is described below:

NGI GAR13 data are supplied as two raster datasets, each indicating the annual frequency of landslide due to seismic trigger and Rainfall (precipitation) trigger. The two rasters were combined for the purposes of *ThinkHazard!*, with the maximum annual frequency per grid cell assigned to the resulting combined trigger raster. The resulting raster was classified into four hazard levels defined by NGI (2013) (Table 8).

¹³ <http://sedac.ciesin.columbia.edu/data/set/ndh-landslide-hazard-distribution>

¹⁴ <http://preview.grid.unep.ch/index.php?preview=data&events=landslides&evcat=1&lang=eng>

Table 8 Annual frequency and corresponding raster score for landslide, based on NGI (2013)

Hazard class	Annual frequency per km ² (1x10-4)	Annual frequency per km ²	Raster score (Ann. Freq. * 1,000,000)
Negligible	<1.8	<0.00018	<180
Low	1.8-3.2	0.00018-0.00032	180-320
Medium	3.2-7.5	0.00032-0.00075	320-750
High	>7.5	>0.00075	>750

SEDAC data is available as a single raster dataset, giving landslide hazard as a decile value of relative global landslide hazard per grid cell. The range 5-9, was adjusted to 6-10 for consistency with other datasets. The original data layer is reclassified in GIS, to produce a raster containing hazard level, as follows (partly based on the classification description given in Dilley et al. (2005): 6 = negligible (translated to very low for *ThinkHazard!*); 7 = Low; 8 = Medium; 9-10 = High).

After assigning a consistent integer value (1-4) to the processed hazard levels, the layers are imported to *ThinkHazard!*, where hazard level is mapped onto ADM2 units.

8.10.1 Future classification

Implementation of new dataset produced by NASA in 2017 (Stanley and Kirschbaum, 2017) is under investigation. This is a global landslide susceptibility map was created by combining information from four principal sources of information for five explanatory variables: slope, distance to fault, geological classification, presence of roads, and forest loss. Susceptibility is classified into five categories: very low, low, moderate, high, and very high. The classes are defined were divided at susceptibility values of 0.11, 0.49, 0.671, and 0.75. Each category was twice as large as the next highest, e.g., the high category contains approximately twice the number of pixels as the very high category (3% of the global area). It has been currently proposed to combine the very high and high categories into ‘high’, and transfer the remaining categories as originally defined.

Investigation is required into how this relates to landslide frequency, and in consistency with the already implemented layers. Additionally, initial testing shows a tendency towards all ADM0 units showing high hazard, which poses a problem for communicating the truly high hazard areas. Investigation will also consider whether an area threshold (similar to that applied to river and urban flood) could be used to avoid this.

9 References

Atkinson and Sonley (2000): Empirical Relationships between Modified Mercalli Intensity and Response Spectra, Bulletin of the Seismological Society of America, 90, 2, pp. 537–544, April 2000

Davies, G., Griffin, J., Løvholt, F., Glimsdal, S., Harbitz C., Thio, H.K., Lorito, S., Basili, R., Selva, J., Geist, E., and Baptista, M.A. (2017), A global probabilistic tsunami hazard assessment from earthquake sources, Geological Society, London, Special Publications, 456, doi:10.1144/SP456.5

Federal Office for Citizen Protection and Disaster Support (German: Bundesamt für Bevölkerungsschutz und Katastrophenhilfe, BBK) (2010): Methode für die Risikoanalyse im Bevölkerungsschutz. Wissenschaftsforum Series No. 8. Available online: http://www.bbk.bund.de/SharedDocs/Downloads/BBK/DE/Publikationen/Wissenschaftsforum/Bd8_Methode-Risikoanalyse-BS.pdf?blob=publicationFile

HFA: <http://www.unisdr.org/we/coordinate/hfa>. The HFA is a 10-year plan to make the world safer from natural hazards. It was endorsed by the UN General Assembly in the Resolution A/RES/60/195 following the 2005 World Disaster Reduction Conference.

Center for Hazards and Risk Research - CHRR - Columbia University, Center for International Earth Science Information Network - CIESIN - Columbia University, and Norwegian Geotechnical Institute - NGI. 2005. Global Landslide Hazard Distribution. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). <http://dx.doi.org/10.7927/H4P848VZ>

Dilley, M., Chen, R. S., Deichmann, U., Lerner-Lam, A. L., & Arnold, M. (2005). Natural Disaster Hotspots A Global Risk Analysis. Washington DC.

Falkenmark M, Lundqvist J & Widstrand C (1989): Macro-scale water scarcity requires micro-scale approaches: aspects of vulnerability in semi-arid development. Natural Resources Forum, 13, 258–267

Global Volcano Model, British Geological Survey and University of Bristol. Development of National Disaster Risk Profiles for Sub-Saharan Africa - Volcanic Risk Hazard and exposure assessment for selected volcanoes – Final report. 2016. 56 p.

Meyer et al. (2011): CRUE Final Report. RISK MAP: Improving Flood Risk Maps as a Means to Foster Public Participation and Raising Flood Risk Awareness: Toward Flood Resilient Communities. Available online at <http://risk-map.org>

Muis, S., Verlaan, M., Winsemius, H.C., Aerts, J.C.J.H., Ward, P.J., 2016. A global reanalysis of storm surge and extreme sea levels. Nat. Commun. 7, 1–11. doi:[10.1038/ncomms11969](https://doi.org/10.1038/ncomms11969)

Nadim, F., Kjekstad, O., Peduzzi, P., Herold, C. and Jaedicke, C. (2006). Global landslide and avalanche hotspots. Landslides, Vol. 3, No. 2, 159-174. <http://dx.doi.org/10.1007/s10346-006-0036-1>

NGI (2013). Landslide Hazard and Risk Assessment in El Salvador – UNISDR Global Assessment Report 2013 - GAR13. 20120052-01-R

Leonard, G.S.; Power, W.; Lukovic, B.; Smith, W.; Johnston, D.; Downes, G., 2008. Tsunami evacuation zones for Wellington and Horizons regions defined by a GIS-calculated attenuation rule, GNS Science Report 2008/30. 22p.

Lindenschmidt et al. (2006): Risk assessment and mapping of extreme floods in non-dyked communities along Elbe and Mulde. *Adv. Geosci.*, 9, 15–23.

Merz and Thielen (2004): Flood risk analysis: Concepts and challenges, *Osterreichische Wasser- und Abfallwirtschaft*, 56(3–4), 27–34.

Stanley, T. and Kirschbaum, D. B. (2017): A heuristic approach to global landslide susceptibility mapping. *Natural Hazards* 87:145–164. doi:10.1007/s11069-017-2757-y

Veldkamp, T.I.E., S. Eisner, Y. Wada, J.C.J.H. Aerts, and P.J. Ward, 2015: Sensitivity of water scarcity events to ENSO driven climate variability at the global scale. *Hydrol. Earth Syst. Sci.*, 19, 4081-4098, doi:10.5194/hess-19-4081-2015.

Winsemius, H.C., L. P. H. Van Beek, B. Jongman, P. J. Ward, and A. Bouwman (2013): A framework for global river flood risk assessments. *Hydrol. Earth Syst. Sci.*, 17, 1871–1892, 2013. www.hydrol-earth-syst-sci.net/17/1871/2013/ doi:10.5194/hess-17-1871-2013

Woessner et al (2013) – FP7- SHARE – Report D5.5: Final seismic hazard assessment including de-aggregation, Swiss Seismological service.

Yamazaki, D., Trig, M.A., Ikeshima, D., 2015. Development of a global ~ 90 m water body map using multi-temporal Landsat images. *Remote Sensing of Environment*, 171, 337-351, doi: 10.1016/j.rse.2015.10.014

10 Appendix 1: Document revisions

Version	Date	Author	Object
2.0	22.09.2017	Stuart Fraser	Addition of new hazards and features included in release of version 2
1.8	28.04.2016	Nicolas Taillefer	Final version for release of version 1
1.7	23.03.2016	Andreas Burzel, Marijn Kuijper, Sophie Vermooten, Stuart Fraser	Methodology cyclones and water scarcity added, general review
1.6	18.03.2016	Nicolas Taillefer	Update on database workflow (input from C2C)
1.5	08.03.2016	Stuart Fraser, Nicolas Taillefer, Andreas Burzel, Marijn Kuijper	Update of the document, and addition of peril specific sections content (EQ, VA, FL)
1.4	28.01.2016	Stuart Fraser, Marijn Kuijper, Nicolas Taillefer	Update of the document, and addition of peril specific sections (template)
1.3	18.12.2015	Nicolas Taillefer, Stuart Fraser	Review, reordering; Revised data quality assessment
This is the version released with Beta1 version of the tool for review by GFDRR (12/2015)			
1.2	14.12.2015	Nicolas Taillefer, Nicolas Chauvin	New template, precision for RP selection
1.1	10.11.2015	Stuart Fraser	Review, reordering
1.0	02.11.2015	Nicolas Taillefer	Update to fit with the Beta version of the tool. Upgrade of the decision tree
0.10	02.10.2015	Nicolas Taillefer, Audrey Hohmann	Review of the global categorization methodology. Chapter 4
0.9	12.08.2015	Stuart Fraser, Marijn Kuijper and Andreas Burzel	River Flood and wind classification, revision of the decision tree, quality rating and meta data template.
0.8	16.06.2015	Nicolas Taillefer, Audrey Hohmann	Data format, methodology for Volcanic and Coastal Flood
0.7	10.06.2015	Marijn Kuijper	Revision of the decision tree and quality rating
0.6	09.04.2015	John Douglas	Took account of outstanding comments, tidied up text, added background to project, deleted some superfluous detail
0.5	01.04.2015	Stuart Fraser	
0.4	18.03.2015	Marijn Kuijper, Hessel Winsemius and Andreas Burzel	
0.3	18.03.2015	Stuart Fraser, John Douglas	
0.2	03.03.2015	Audrey Hohmann, John Douglas	
0.1	25.02.2015	Marijn Kuijper, Hessel Winsemius and Andreas Burzel	

0.0

29.01.2015

John Douglas, Alanna Simpson and
Stuart Fraser

draft

11 Appendix 2: Hazard coverage

The following hazards are presently available in *ThinkHazard!*¹⁵:

1. **River Flood:** Overflow of a body of water (river, lake) that submerges land otherwise not normally inundated. Also referred to as fluvial flood.
2. **Urban Flood:** Flood data in urban areas only, from both fluvial and pluvial (surface flooding from intense rainfall) sources.
3. **Coastal Flood:** Inundation of land from coastal waters, due to high tidal levels, or storm surge. Storm surge is a temporary rise in sea level as water is pushed toward the shore by the force of winds associated with a tropical or extra-tropical cyclone.
4. **Tsunami:** A series of multiple ocean waves generated by submarine earth movements, earthquakes, volcanic eruptions or landslides.
5. **Water Scarcity:** Water Scarcity originates from a deficiency of precipitation over an extended period, usually a season or more. This deficiency results in a water shortage for some activity, group, or environmental sector. Different from other hazards in that it develops slowly, sometimes over years, and its onset can be masked by several factors. Water Scarcity can be devastating: water supplies dry up, crops fail to grow, animals die and malnutrition and ill health become widespread. Different types of drought can be distinguished (e.g. Wilhite, 2006): meteorological, hydrological, agricultural and socio-economic droughts. In *ThinkHazard!* drought hazard refers to hydrological drought, a shortage of river runoff, in relation to the population density.
6. **Cyclone:** A non-frontal storm system characterized by a low-pressure center, spiral rain bands and strong winds. Usually it originates over tropical or subtropical waters and rotates clockwise in the southern hemisphere and counter-clockwise in the northern hemisphere.
7. **Extreme Heat:** Extreme heat, and heatwave, refer to an extended period of high atmosphere-related heat stress, relative to usual conditions.
8. **Wildfire:** Wildfire, or bushfire refers to fire in occurring in an area of vegetation. Wildfire can spread rapidly in the right conditions and cause destruction of landscapes and the built environment at the urban-wildland interface. Wildfire require high temperatures, combustible vegetation and an ignition source to develop and spread. Ignition may be natural (e.g., lightning) or human-influenced such as arson or accidental ignition.
9. **Earthquake:** Shaking, trembling or displacement of the earth surface due to seismic waves or other phenomena of volcanic or tectonic origin.

¹⁵ Definitions are based on www.preventionweb.net/unisdr.

10. **Volcanic:** Volcanoes are vents in the surface of the Earth through which magma and associated gases erupt, and the resulting structures that are produced by the erupted material. Volcanic hazard comprises proximal hazards such as ballistics, lava, lahars and debris flows, in addition to the more distal effects of volcanic ash fall.
11. **Landslide:** The movement of a mass of rock, debris, or earth down a slope. These encompass events such as rock falls, topples, slides, spreads, and flows, such as debris flows commonly referred to as mudflows or mudslides. Landslides can be initiated by disturbance and change of a slope due to rainfall, earthquakes, volcanic activity, changes in groundwater, man-made construction activities, or any combination of these factors.

12 Appendix 3: Metadata template

Metadata describe data layer content and provides some of the information displayed to the user in the front end of *ThinkHazard!*. Metadata is stored in the GFDRR Innovation Labs GeoNode. The following template is proposed by GFDRR. It contains all information required for further processing of the data by the tool. It is based on ISO 19115:2003 and the World Bank Metadata Standards Quick Guide¹⁶.

Table 9 Metadata template

Field Name	Field Description
Owner*	Default is the person who loaded the dataset, this should be changed to show the data owner/creator – GeoNode Username
Title*	Name by which the dataset is known and should be cited. At a minimum, the name should indicate where, what, and when.
Creation Date*	Reference date for the cited dataset
Data Update Date	Automatically updated every time the layer is replaced.
Date	Reference date for the cited dataset
Date Type*	Event used for reference date. Drop down options are: -creation (identifies when the dataset was brought into existence) -publication (identifies when the dataset was issued) -revision (identifies when the dataset was examined and improved or amended)
Edition	Version of the cited resource
Abstract*	Brief narrative summary of the content of the resource(s)
Purpose	Summary of why this resource was developed?
Maintenance Frequency	Frequency with which the data will be updated by the creator and thus updated on GeoNode
Regions*	Choose the country or region from the menu associated with the dataset
Restrictions*	Constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitation or warning on using the resource or metadata. Drop down options are: - copyright (exclusive right to the publication, production, or sale of the rights to a literary, dramatic, musical, or artistic work, or to the use of a commercial print or label, granted by law for a specified period of time to an author, composer, artist, distributor) - intellectual Property Rights (rights to financially benefit from and control distribution of non-tangible property that is a result of creativity) - license (formal permission to do something other -Restrictions (limitation not listed)

¹⁶ <https://collaboration.worldbank.org/docs/DOC-3008>

- patent (government has granted exclusive right to make, sell, use or license an invention or discovery)
- patent Pending (produced or sold information awaiting a patent)
- restricted (withheld from general circulation or disclosure)
- trademark (a name, symbol, or other device identifying a produce, officially registered and legally restricted to the use of the owner or manufacturer)

License*	License of the dataset <ul style="list-style-type: none"> - NextView - Not Specified - Open Data Commons Open Database License/OSM - Public Domain - Public Domain/USG - Varied/Derived - Varied/Original
Language*	Language used within the dataset
Hazard Type*	<ul style="list-style-type: none"> - Earthquake - Water scarcity - River Flood - Tsunami - Coastal Flood - Strong Wind - Volcanic Ash - Landslide
Hazard Set ID*	This ID will link the associated layers for each hazard dataset so that the analytical framework knows to reference these layers in the same query.
Glide Number	Not used for <i>ThinkHazard!</i> , only for Labs Use; ID associated with hazard event; only to keep historic layers;
Intensity Unit*	The units of intensity specified in the hazard layer (e.g. meters, feet, PGA, m/s, MMI etc.)
Return Period*	The return period of the layer in years

Calculation Method Quality*	This will be a number from 0-10 and will be provided by the Administrator
Scientific Quality*	This will be a number from 0-2 and will be provided by the Administrator
Spatial Representation Type	Method used to present geographic information in the dataset
Temporal Extent Start	If the dataset is temporal data then enter a start date
Temporal Extent End	If the dataset is temporal data then enter an end date
Supplemental Information	Enter any other descriptive information about the dataset
Distribution URL	Information about online sources from which the dataset, specification or community profile name and extended metadata elements can be found. Automatically adds the URL to the items location on GeoNode.
Distribution Description	Detailed text description of what the above online resource is/does.
Data Quality Statement	General explanation of the data producer's knowledge about the lineage of a dataset
Point of Contact- Individual Name*	Name of the responsible person – surname and given name
Featured Check Box	Leave this box unchecked
Is Published Check Box*	Check this box to publish an item to the public GeoNode and to be discoverable by <i>ThinkHazard!</i>
Thumbnail URL	Automatically populated by GeoNode
Download URL	Automatically populated by GeoNode
Keywords*	A space or comma separated list of keywords.
Point of Contact *	Name of the responsible GFDRR person – GeoNode Username
Metadata Author*	Name of the GFDRR person who authored the metadata – GeoNode Username
Topic Category*	Specify the main ISO category through which your map of data could be classified. e.g., boundaries, climatology, geoscientific Information

*Required

13 Appendix 4: Data Quality Criteria

Table 10 Data quality rating criteria and rating scheme. These ratings are applied in layer metadata and used in *ThinkHazard!* to determine which hazard data to use

Criteria	Rating (points)	Comments / Considerations
1. Quality of the calculation method	1-10 (based on sum of values below)	to be added to the metadata
1A Model set-up	Highest level of detailed modeling approach (e.g. 1D-2D for floods, locally-specific earthquake ground motion prediction equations): 2 Detailed modeling approach (e.g. 2D for floods, regionally-appropriate earthquake ground motion prediction equations): 1 Generic modeling approach (e.g. 1D with level/volume spreading for floods, global earthquake ground motion prediction equations): 0	Would require knowledge of model method, but can be obtained.
1B Elevation data (terrain / bare ground only, or elevation of additional features included)	No model: 0 General (Elevation model): 1 Precise (Terrain model): 2	Provide examples for each hazard, if relevant: FL: elevation/terrain, CF, TS, CY: elevation/terrain, EQ: topography/soil
1C Data resolution and components	Explicit modeling of required components, high-resolution data (e.g. explicitly modeled flood protection, local-scale geology/landcover data): 2 Statistical approximation of components, or low-resolution data (e.g. protection level applied statistically, global or continental scale geology/landcover data): 1 Major model components excluded and low-resolution data (e.g. no flood protection, global data): 0	Would require knowledge of model method, but can be obtained.
1D Vintage of data (year in which the data was generated)	2015 to present: 2 2010-2015: 1 Older than 2010: 0	Yes, would be given with data
1E “Official” government data	Official governmental data: 2 Not-official governmental data: 0	Yes, would be given in source (as owner / publisher)
2. Scientific quality of the data	0-2 (based on sum of values below)	to be added to the metadata
2A Peer review	Peer reviewed publication available about dataset: 2 No peer reviewed publication available: 0	Requires knowledge of research or project / input from owners.

2A1	If not peer reviewed: clear description available and acceptable method	Clear description available and acceptable method to determine boundary conditions and model input (e.g. what is the model input and corresponding frequencies): 1 No description available: 0	Requires knowledge of research or project / input from owners.
2A2	If not peer reviewed: description available of processes	Description available of processes modeled to quantify the hazard: 1 No description available: 0	Requires knowledge of research or project / input from owners.