ThinkHazard!

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

Methodology report

VERSION: 1.8
DATE: 28/04/2016

AUTHOR(S): John Douglas, Alanna Simpson, Stuart Fraser, Marijn Kuijper, Hessel Winsemius, Andreas Burzel, Audrey Hohmann, Nicolas Taillefer, Sophie Vermooten

Review	Sylvain Grellet	
Approval	Approver's name (GFDRR Lab team member)	*Signature*

Summary

This report presents the method used in *ThinkHazard!,* to transform technical hazard data layers provided by the Global Facility for Disaster Reduction and Recovery (GFDRR)[1] into four categories of hazard (from high to very low), which are presented to *ThinkHazard!* users who are not experts in natural hazards. The method is based on an intensity threshold, beyond which a hazard could cause damage or disruption to a development project, combined with frequency of that threshold being exceeded (represented by hazard maps for various return periods, or frequencies of exceedance). The level of hazard category assigned to a given location is defined by whether the intensity at that location at each of the different return periods exceeds or does not exceed the intensity threshold. This report presents the decision tree that is used to decide which data source is used for a given location, based on the assessed quality of the data source. In addition, the report presents a description and tests of this method for earthquake and river flood data.

This document was continually modified and extended as the project progressed; leading to this final version.

Document revisions

Version	Date	Author	Object
0.0	29.01.2015	John Douglas, Alanna Simpson and Stuart Fraser	draft
0.1	25.02.2015	Marijn Kuijper, Hessel Winsemius and Andreas Burzel	
0.2	03.03.2015	Audrey Hohmann, John Douglas	
0.3	18.03.2015	Stuart Fraser, John Douglas	
0.4	18.03.2015	Marijn Kuijper, Hessel Winsemius and Andreas Burzel	
0.5	01.04.2015	Stuart Fraser	
0.6	09.04.2015	John Douglas	Took account of outstanding comments, tidied up text, added background to project, deleted some superfluous detail
0.7	10.06.2015	Marijn Kuijper	Revision of the decision tree and quality rating
0.8	16.06.2015	Nicolas Taillefer, Audrey Hohmann	Data format, methodology for Volcanic and Coastal Flood
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.

[1] GFDRR: www.gfdrr.org.

0.10	02.10.2015	Nicolas Taillefer, Audrey Hohmann	Review of the global categorization methodology. Chapter 4
1.0	02.11.2015	Nicolas Taillefer	Update to fit with the Beta version of the tool. Upgrade of the decision tree
1.1	10.11.2015	Stuart Fraser	Review, reordering
1.2	14.12.2015	Nicolas Taillefer, Nicolas Chauvin	New template, precision for RP selection
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)**
1.4	28.01.2016	Stuart Fraser, Marijn Kuijper, Nicolas Taillefer	Update of the document, and addition of peril specific sections (template)
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)
1.6	18.03.2016	Nicolas Taillefer	Update on database workflow (input from C2C)
1.7	23.03.2016	Andreas Burzel, Marijn Kuijper, Sophie Vermooten, Stuart Fraser	Methodology cyclones and water scarcity added, general review
1.8	28.04.2016	Nicolas Taillefer	Final version

Table of contents

Table of illustrations

1 Introduction

1.1 Background

The Global Facility for Disaster Reduction and Recovery (GFDRR)[2] is a global partnership, managed by the World Bank and funded by 25 donor partners, to help high-risk, low-capacity developing countries better understand and reduce their vulnerabilities to natural hazards and to adapt to climate change in the spirit of the Hyogo Framework for Action (HFA)[3].

GFDRR focuses on the five following pillars of action:

- Pillar 1: Risk Identification
- Pillar 2: Risk Reduction
- Pillar 3: Preparedness
- Pillar 4: Financial Protection
- Pillar 5: Resilient Recovery

[2] GFDRR: www.gfdrr.org.
[3] 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.

The Risk Identification pillar aims to improve access to information about disaster and climate risks in vulnerable countries and to support a greater capacity to create, manage and use this information. This pillar then provides valuable and often critical inputs to the other pillars to reduce and manage disaster risk. The *ThinkHazard!* project was initiated in 2015 in order to design and setup an analytical tool for non-DRM (Disaster Risk Management) specialists to understand the level of natural hazards in development project areas.

1.2 Main objective of *ThinkHazard!*

Interpretation of hazard information requires knowledge based on heterogeneous data, which are often highly technical and difficult to find and to exploit for non-specialists of the DRM domain (referred to below as 'the user' or 'users'). Rapid and simple access to high added-value hazard information is a key barrier for various stakeholders of the DRM domain. Thus, availability, accessibility and processing of these information sources are currently crucial issues.

The main aim of the *ThinkHazard!* project is to develop an analytical tool dedicated to facilitating improved knowledge and understanding of natural hazards among non-DRM specialists. It aims to provide users with a view of natural hazards in project areas, and relevant guidance on appropriately handling the threats posed by those hazards. Thus, the intuitive web-based tool that has been developed, enables users to:

i) rapidly undertake simple and preliminary hazard screening of a project, receiving a list, map, and short report of hazard levels for multiple hazards
ii) access recommendations on how to account for the hazards present, in their project planning, and
iii) access additional information related to the hazards and locations searched (including previous studies, related projects).

Furthermore, all aspects of this tool will be open and transparent, to provide users with enough information to understand its operational principles. This tool will also be easy to update in order to integrate new hazard datasets and analysis processes.

ThinkHazard! version 1 can be accessed at:

www.thinkhazard.org

1.3 Aim and structure of the document

This document describes the *ThinkHazard!* method for categorization of an area (national, to county or region) into four hazard levels (from High to Very Low) for multiple natural hazards. These hazard levels are related to how often potentially damaging events occur in a users' location of interest. It also describes the basis on which recommendations and additional information are provided to the user in *ThinkHazard!*.

Section 2 describes the input of hazard data to *ThinkHazard!*, including the assessment of source data quality prior to upload to the tool, and the model decision tree that determines which data is to be used for each administrative area. This section also discusses the storage of data within the GFDRR Geonode structure (See section 4.3.1), external to *ThinkHazard!*.

Section 3 presents the hazard categorization procedure, including the development of a process which uses both intensity and frequency to determine hazard level. This procedure is applied in an internal process to data that are available in a suitable format (a set of gridded data layers). The same section introduces preprocessing of data, which are additional hazard-specific processing steps to include useful hazard data, whose format does not permit application of the main hazard categorization procedure.

Section 4 presents the development process behind recommendations and additional information that are provided to users, to guide them on incorporating hazard information to reduce disaster risk in project design.

Section 5 presents hazard categorization procedures that are specific to each hazard. This section comprises greater detail on preprocessing of some data types.

2 Hazard data (input)

Data may come from numerous organizational and academic sources. They are generally of three formats:

- A set of GIS raster (grid) layers in which each grid cell represents the value of a mapped intensity parameter. One raster layer represents the intensity at a specific associated with multiple probabilities of exceedance.
- A set of GIS vector (point, polygon or line) layers that contain values of a parameter characterizing the intensity of a hazard, at a given return period. Layers of this type may be processed into raster format for use in the hazard categorization procedure.
- A single GIS vector or raster layer providing non-probabilistic hazard data (e.g., a susceptibility index). This type of data must be preprocessed before import to *ThinkHazard!*, into a raster layer containing mapped values that directly represent a hazard level. *ThinkHazard!* will directly map the hazard levels provided in the layer.

The first type is exemplified by Figure 2-1, in which earthquake 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). Such maps are produced for different return periods, for use in projects that require estimated probability of loss or impact. For earthquakes, return periods of 475, 975 and 2475 years are common. A typical choice for seismic design codes for normal buildings is 475 years; longer return periods are used as the design basis for critical infrastructure such as hospitals, bridges or dams. This is because longer return periods correspond to a smaller chance that the design level will be surpassed during the lifetime of the structure, hence there is a lower risk of damage to buildings designed to those levels. A hazard map is the visualization of a 'slice' at a certain return period (or probability of exceedance) through a hazard curve, which is a graphical representation of hazard calculated for all points in an area. Figure 2-2 explains the procedure for the generation of earthquake hazard maps.

Such maps contain valuable information about earthquake hazard, but rely on specialist technical knowledge for interpretation, making them unsuitable for communicating hazard to non-specialists in DRM. The hazard categorization process described in Section 2 translates the technical data (e.g., PGA) contained in these maps, into accurate but simpler hazard information. It translates this information into the degree of awareness that the user should have of the potential impacts of a given hazard on their project.

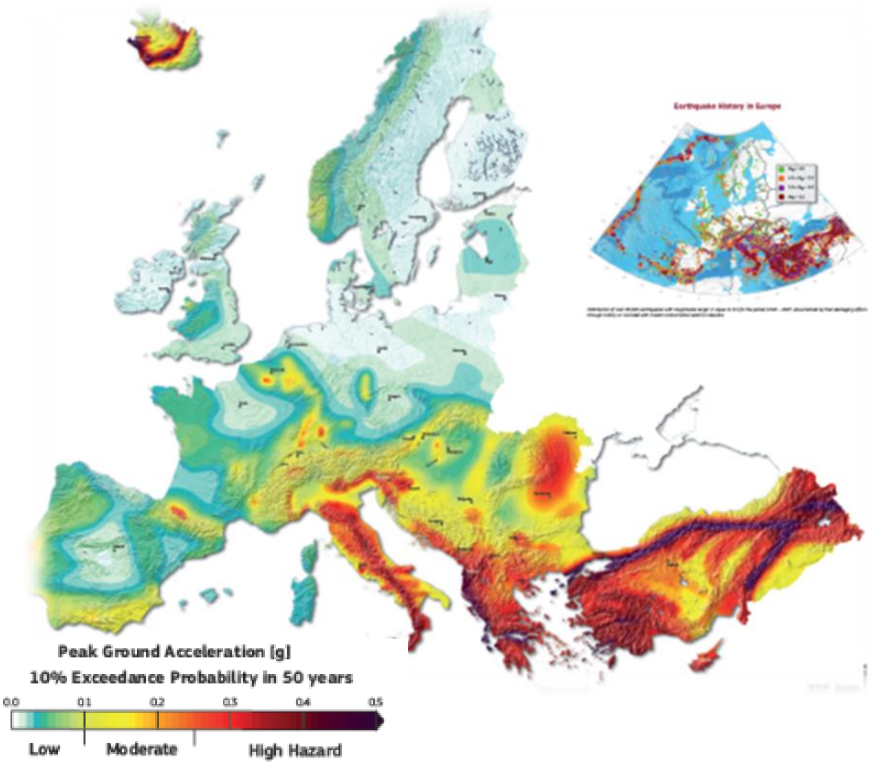


Figure 2-1: Example seismic hazard map, PGA for Europe at 10% exceedance probability in 50 years (source: SHARE[4]).

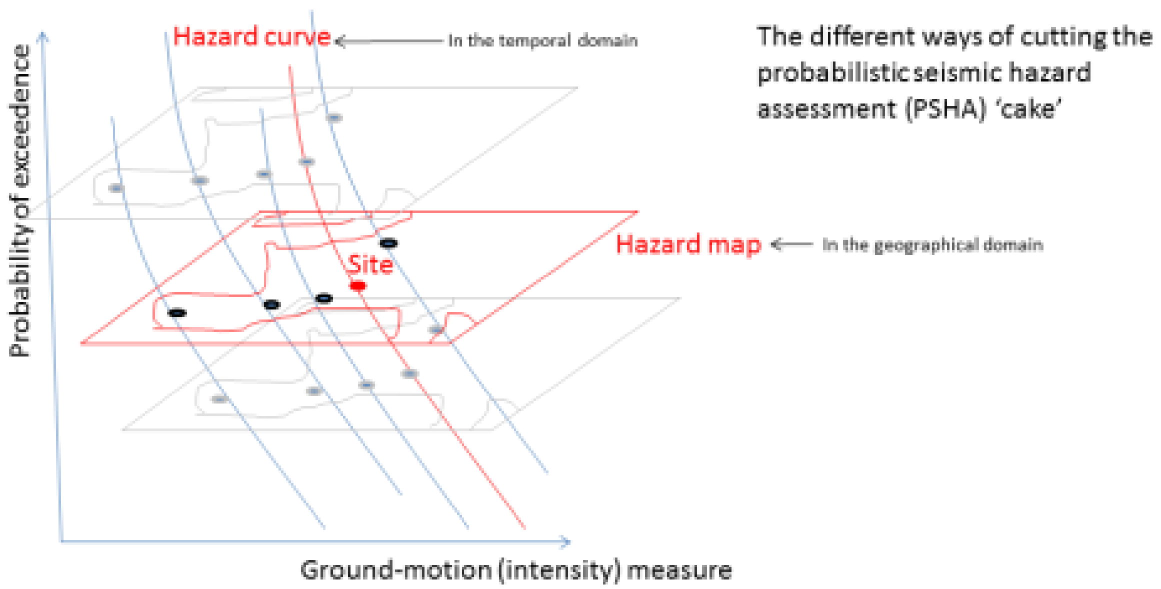


Figure 2-2: Modified illustration of how seismic hazard maps relate to hazard curves (source: J. Douglas within SHARE).

[4] http://www.efehr.org:8080/jetspeed/portal/hazard.psml

2.1 Hazard coverage

The following hazards are implemented in *ThinkHazard! version 1*[5]:

1. **River Flood** (FL): Overflow of a body of water (river, lake) that submerges land otherwise not normally inundated.

2. **Earthquake** (EQ): Shaking, trembling or displacement of the earth surface due to seismic waves or other phenomena of volcanic or tectonic origin.

3. **Water Scarcity** (DG): Water Scarcity originates from a deficiency of precipitation over an extended period of time, 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 a number of 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 de 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.

4. **Cyclone** (CY): 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.

5. **Coastal Flood** (CF): 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.

6. **Tsunami** (TS): A series of multiple ocean waves generated by submarine earth movements, earthquakes, volcanic eruptions or landslides.

7. **Volcanic** (VO): 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.

8. **Landslide** (LS): 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.

---

[5] Definitions are based on www.preventionweb.org of UNISDR. Abbreviations are those used in the *Think Hazard!*.

2.2 Global and local data

ThinkHazard! is flexible in the extent of data it uses. Data are considered 'global' if they cover the entire earth, and 'local' if their coverage is limited to a given number of administrative divisions. 'Local' data includes data that cover a region, i.e., several countries. Table 2-1 outlines the type of data available for version 1.

Table 2-1 : Type of datasets available for version 1 of the tool.

Hazard	Mapped parameter values (raster)	Point or polygon values (vector): partial preprocessing	Proximal data: full preprocessing
River Flood	Global and local		
Earthquake	Global and local		
Water Scarcity		Global (polygon)	
Cyclone	Global and local		
Coastal Flood	Global and local		
Tsunami		Local (point)	
Volcanic			Global
Landslide	Global		

The coverage of a local datasets is taken into account into the processing following three steps, aiming at limiting hazard categorization to administrative units where there is actually data:

1. Screening of the bounding box of the raster and selection of the administrative units that intersects the bounding box,
2. If one administrative unit is only covered by no-data pixels, it is excluded from the selection,
3. Mapping procedure (see 5.1.3) is applied on each administrative unit of the selection.

2.3 Tool administration

In order to protect original data, control the processing and the quality of results, and provide the user with a fluid user experience, a specific workflow involving various databases has been set up (Figure 2-3).

Main datasource is a database managed by GFDRR Innovation Lab for general purposes.

Suitable data for *ThinkHazard!* processes are identified and copied to a processing database, where specific data processing (hazard levels, mapping, conversion to raster), and management of added value information (recommendation, further information) take place. Finally, when the administrator decides to make new data or recommendations available for the user, the processing database is copied and dropped into a public database.

Consequently the public database is an image of the processing one at a given date, and is searchable by the user through the user interface.

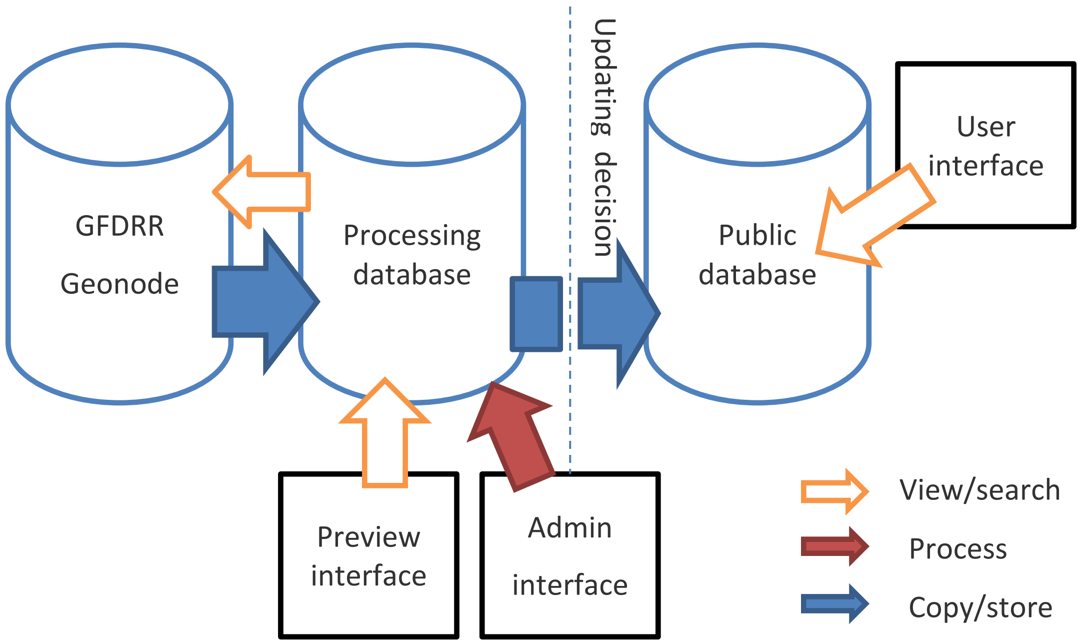


Figure 2-3: General architecture of the tool and back office workflow.

2.3.1 Geonode storage

All hazard data and additional information layers are stored on a Geonode instance – an Open Source Geospatial Content Management System. For *ThinkHazard! Version 1*, data sources are provided by the GFDRR Innovation Labs Geonode instance. The Geonode enables *ThinkHazard!* administrators to store and curate data with a uniform standard for styling and metadata. Metadata are completed according to the template shown in section 2.3.4, to provide information regarding data sources, and enable selected description of data within *ThinkHazard!*. *ThinkHazard!* uses several of these metadata information fields to implement the model decision tree, and accesses the hosted data values to apply the hazard classification procedure. Users of the tool are able to access the geonode page of the underlying dataset or the data provider's web page, from hyperlinked text in the user interface, giving the name of the data owner.

2.3.2 Databases and workflow of data

ThinkHazard! works with two different databases dedicated to the tool, populated with data harvested from the GFDRR Geonode instance:

- The processing database for data administration/preview (aka processing database),
- The public database for the end-users (aka public database).

The harvesting of hazard data from Geonode, as well as actions like downloading, completing and processing of hazard sets, happens on the processing database.

The modifications made on recommendations (text, order, etc...) via the administrator interface also happen on the processing database.

Administrators can verify that the processing database displays the right information (ie. correct data and recommendations) in the preview interface, within *ThinkHazard!* administration user interface. This preview interface takes any modification in the processing database into account.

The public database is a copy of the processing database at a given date. It's dropped and replaced by a new version each time the administrator decides to publish the data. At that time an archive of the database would be stored for future needs (a complaint from a user for example). This archive is accessible to administrators only; users would need to request archived results from the administrator if required. The version number is presented on the About page.

In is recommended that an update of the public database is done in the following cases:

- Changes concerning recommendations: modifications have been made on recommendations, and the modifications are validated,
- Changes in the data processing: new data has been processed or processing has been re-launched after quality ratings have been changed, and the processing results sanity has been checked.

2.3.3 Administrator role

The *ThinkHazard!* administrator is responsible for identifying data and resources that should be included in *ThinkHazard!*, and curating the data in such a way that it conforms to *ThinkHazard!* requirements in terms of metadata, including data quality ratings. Data curation requires the use of a Geonode. The administrator should follow the separate document, 'Administrator Workflow' in order to achieve this.

2.3.4 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 (Table 2-2). 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 Guide[6].

Table 2-2: 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.

[6] https://collaboration.worldbank.org/docs/DOC-3008

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) - 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 - 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*	- 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 *ThinkHazard!*, 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 *ThinkHazard!*
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

2.4 Hazard data selection

The premise of the tool is to use the best available hazard information to derive the hazard level for each administrative unit. An internal algorithm, or decision tree, based the global or local nature of the data and the data quality rating identifies the most appropriate information for each unit.

2.4.1 Quality rating

Data quality rating should be based on information available with the original dataset, and any necessary dicsussions with data providers. Most criteria would have to be obtained by direct communication with data owners to obtain detail on the methodology and publication associated with the work, as some information is unlikely to be available in standard metadata. The data quality rating is applied by the administrator before upload of data to the GFDRR Geonode, and the data quality score is recorded in the metadata, using the following process:

- As part of the *ThinkHazard!* Administrator workflow, the criteria in Table 2-3 are assigned a score reflecting the features of the hazard data.
- There are two categories for considering data quality, each with a maximum range of scores: 'quality of the calculation methods' (0-10 points) and 'scientific quality' (0-2 points). Scores for the criteria in each category are summed to provide the category score.
- The scores are entered by an Administrator into the metadata fields: "calculation_method_quality" and "scientific_quality". The full breakdown of scoring is stored by the administrator, in a separate spreadsheet.
- *ThinkHazard!* algorithm reads the above metadata fields to obtain the two calculated scores for use in the decision tree.

2.4.2 Decision tree

For many locations several data sets will be available for instance one global dataset and one or more local datasets. *ThinkHazard!* will, therefore, need a decision tree to automatically determine which data to use. The decision will be made based on the available metadata, and is illustrated in the schema presented in Figure 2-4.

Table 2-3 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.

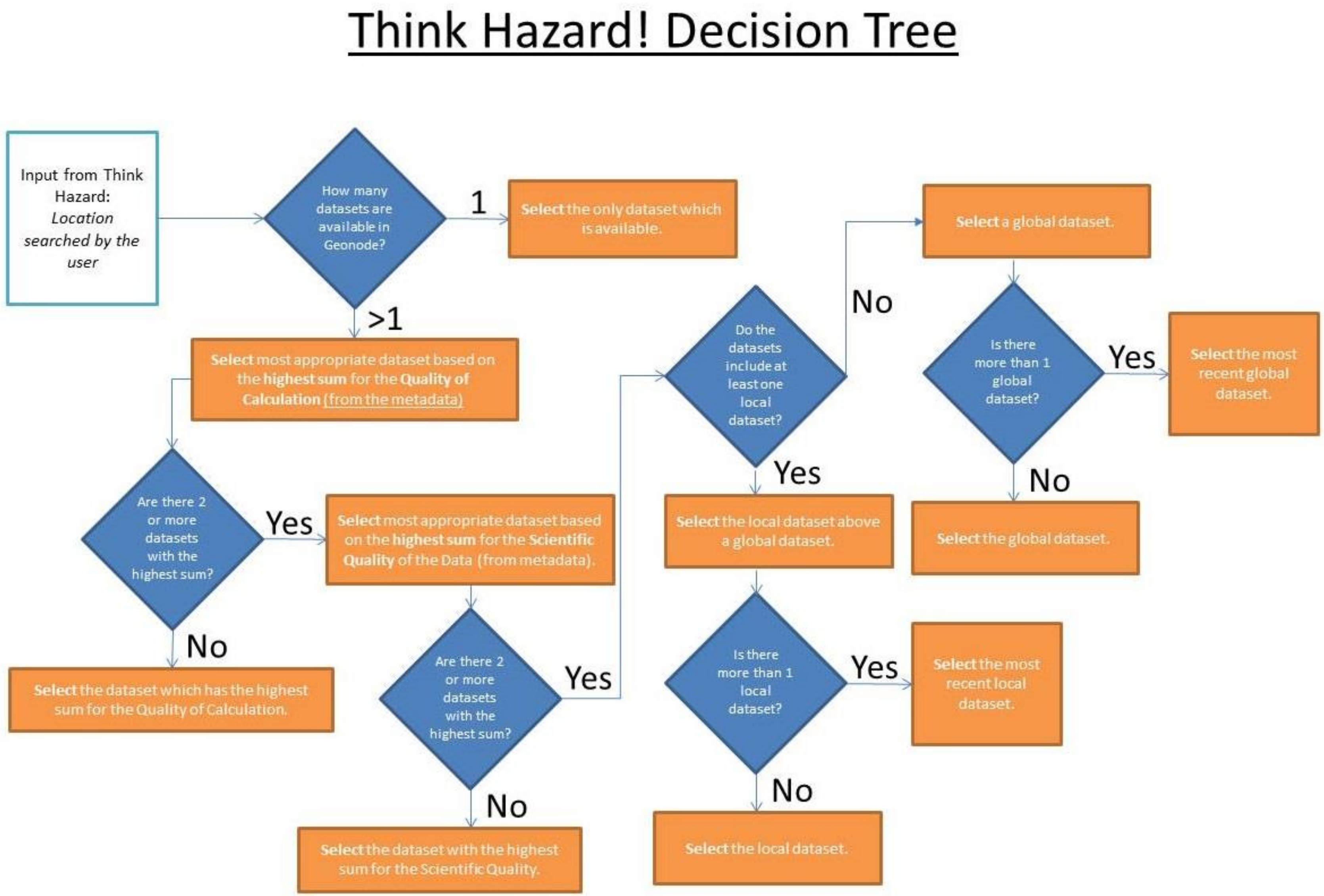


Figure 2-4 : Decision tree for the selection of data to be returned by *ThinkHazard!*

3 Hazard levels

The *ThinkHazard!* approach to produce valuable and easily understandable information for the user consists of three main steps:

1. Hazard categorization
   Translates scientific characterization of hazard (i.e. hazard data) into easily understandable hazard classes (hazard level) referring to the degree of awareness one should have when selecting or preparing a project in a given geographical area.

2. Mapping of results
   Enables the user to visualize hazard levels on a map, aggregated to administrative units 0 (national level), 1, and 2.

3. Provision of recommendations
   Recommendations associated with hazard level are provided to guide a user in reducing risk from multiple hazards to the planned project. This includes the provision of additional information associated with hazard level and location, to guide a user to sources of relevant websites, previous studies or other information.

This section describes the hazard categorization and mapping. Details on the development of Disaster Risk Reduction (DRR) guidance follow in Section 5.

3.1 Hazard categorization

The *ThinkHazard!* objective is to produce a simple tool for use by non-DRM experts and planners of development projects. Users enter a location of interest only; no further information is entered about project type. The hazard category (level) is the key parameter describing the potential hazard impact of the project located in a given area, and determines which hazard-specific recommendations are provided to the user. The tool transforms mapped hazard data, generally contained in a raster file containing intensity values on a grid basis (e.g. *.tif*), into a hazard level for each administrative unit.

To make the results of the tool easier to understand, it was decided to label hazard categories according to expressions that would normally relate to the hazard's intensity: 'high', 'medium' and 'low', and 'very low' rather than labels related to the frequency of occurrence. The tool's focus is on raising awareness of hazards, rather than providing scientific results, and the labels assigned, e.g. 'high', refer to **how aware a user needs to be of that peril for their project**. Effectively, the tool is using frequency and intensity information to provide a view to non-specialists of **how likely it is that a given hazard can achieve damaging intensity in the area** (i.e. its frequency) rather than what size such a hazard would be if it occurs (i.e. its intensity), see Figure 3-1.

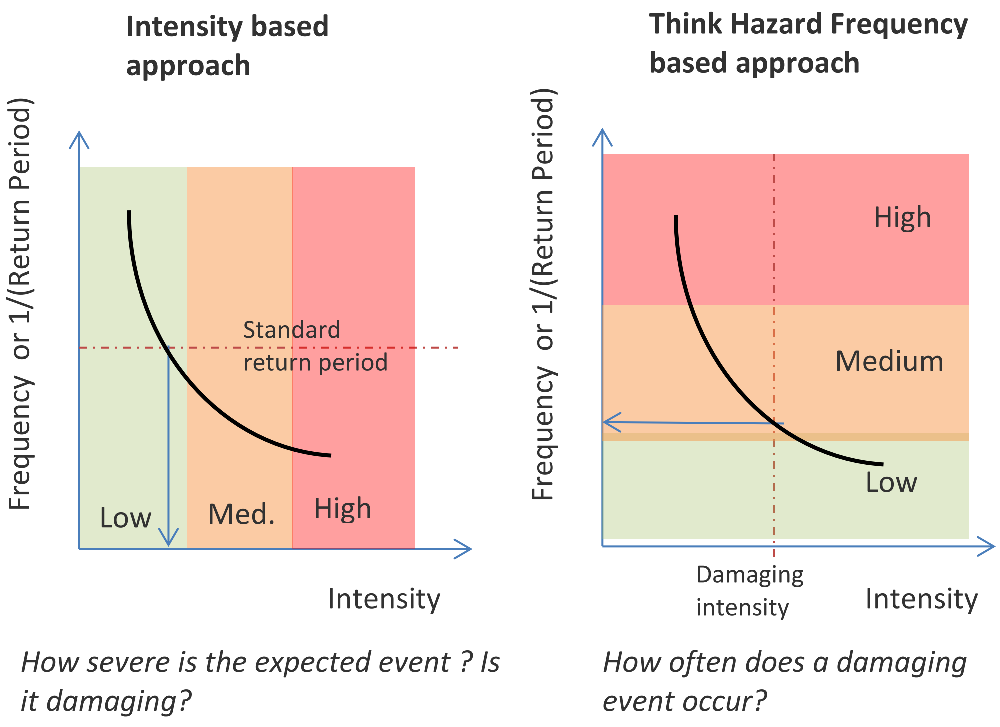


Figure 3-1: Comparison of Think Hazard frequency-based approach, and common intensity-based approach.

To determine what constitutes '**damaging intensity'**, the tool relies on definition of a damage threshold in terms of the hazard parameter. In *ThinkHazard! version 1*, the type of project is not specified by the user, it is difficult to determine hazard category according to intensity with a relevant meaning for very different project types (e.g. agricultural or infrastructure). In version 1, the defined damage threshold is generic, i.e., not specific to any particular project type. In future, sector-specific versions of the tool 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.

A frequency-based approach means that intensity thresholds are calibrated to characterize significant hazards, and the return periods considered are those for which hazard maps exists and hence are often the *de facto* standards for a given peril (e.g. 475, 975 and 2475 years for earthquakes and 5, 10, 25, 50, 100, 1000 and 10000 years for floods). Finally, adopting a frequency-based approach for the categorization can allow comparison between hazards, although only if the same return periods are considered for each peril. Return period are selected among common-practice values for each field (widely available dataset exists). Hazard mainly threatening life of people due to damages to construction (Earthquake, cyclone, volcanoes) can be ranked considering a 50 years lifespan, whereas hazard threatening people's living conditions (flood, drought) are often ranked with regards to lower duration (10 years) more typical from memory and economic impacts.

A step by step procedure is applied to represent potentially damaging hazards, after an intensity threshold has been selected:

1. If **hazard intensity values exceed the intensity threshold for the shortest return period**, hazard category is defined as **high**.

2. Then, if the hazard intensity values do not exceed the intensity threshold for the lower return period, the medium return period is considered. In case the **hazard intensity values exceed the intensity threshold at this medium return period**, hazard category is defined as **medium**.

3. If the hazard intensity values still do not exceed the intensity threshold for the medium return period, the longest return period is considered. In case the **hazard intensity values exceed the intensity threshold at this longest return period**, hazard category is defined as **low**.

4. If the **hazard intensity values do not exceed the intensity threshold for the longest return period**, hazard category is defined as **very low**.

The procedure can be illustrated easily in the case of earthquake. It is assumed that intensity data are available as hazard maps containing PGA values for three return periods (100, 475 and 2475 years). For a given location, if the 100 year return period hazard map shows that PGA value is > 0.1 *g* (as an example intensity), it means that earthquake shaking exceeding 0.1 *g* statistically occur once every 100 years, or with a 1% chance of occurring in one year.

According to the European Macroseismic Scale[7] (EMS98), ranks seismic shaking according to their effect on people, buildings and the environment. An event reaching the intensity degree VI on this scale is considered to cause 'slightly damaging' effects to structures, 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 correlations[8] between intensity and PGA, intensity VI corresponds approximately to a 0.1 *g* and intensity VII to 0.2 *g*.

If intensity VI is chosen to establish the intensity threshold of significant earthquake shakings, and datasets are used with return periods 1 in 100, 1000, and 2475 years, the hazard category can be understood as:

- **High**: user should be highly aware of the effect of earthquake for its project, since potentially damaging shaking (exceeding Intensity VI according to EMS) are likely to occur in the area, on average, once every 100 years.

- **Medium**: user should be aware of the effect of earthquake for its project, since potentially damaging shaking (exceeding Intensity VI according to EMS) are likely to occur in the area, on average, once every 1000 years.

- **Low**: user should be aware of the effect of earthquake for its project in some cases, since potentially damaging shaking (exceeding Intensity VI according to EMS) are likely to occur in the area, on average, once in 2475 years.

---

[7] G. Grünthal, A Levret, European macrosismic sclae 1998, Cahier du centre européen de géodynamique et de séismologie, vol 19, 2001.
[8] Atkinson and Sonley, 2000

- **Very Low**: available data suggest that potentially damaging shaking are unlikely to occur, on average, within 2475 years. This does not mean that minor shaking (intensity lower than VI) could not occur.

- **No Data Available**: it means that no dataset covering the chosen location was acceptable for processing by the tool.

3.1.1 General procedure (hazard data with gridded intensity values)

The procedure to derive hazard level from raw raster data is described here. For the purpose of a versatile and simple tool, it was decided that the following general procedure is more appropriate:

1. Consider hazard maps for different three return periods T_r, ordered in the increasing order: $T_{r,H}$, $T_{r,M}$, $T_{r,L}$

2. Define thresholds of interest I (e.g. Y, corresponding to the intensity level above which damage becomes possible), for each hazard category: High (I_H), Medium (I_M), Low (I_L).

3. Cells without values are flagged "no data", they are ignored in the next part of the procedure.

4. Then apply an incremental procedure:

    a. If I_H is surpassed at the location on the hazard map (a pixel of the raster) for the shortest return period ($T_{r,H}$) then the hazard is categorized as **High**,

    b. If I_H is not surpassed on the hazard map for the shortest return period ($T_{r,H}$), then consider the map for $T_{r,M}$, and the threshold I_M

        i. if I_M is surpassed at the location on the hazard map for the return period ($T_{r,M}$) then the hazard is categorized as **Medium**,

        ii. if I_M is not surpassed on the hazard map for the return period ($T_{r,M}$), then consider the map for $T_{r,L}$, and the threshold I_L

            1. If I_L is surpassed at the location on the hazard map for the longest return period ($T_{r,L}$) then the hazard is categorized as **Low**,

            2. If I_L is not surpassed on the hazard map for the shortest return period ($T_{r,L}$), then the hazard is categorized as **Very low**,

In the procedure, the most important parameters are the ordering of the return periods and of the intensity thresholds. A priori, three return period and three intensity thresholds are needed. These are discussed in the following sections.

3.1.1.1 Hazard intensity parameter

Table 3-1 summarizes the intensity units and thresholds of the data processed by the tool, for each hazard, at global and local scale. Conservative thresholds are used in *ThinkHazard! version 1*, because this version focuses on generic (not project-specific) recommendations in International Development Association (IDA) countries, in which investments may be more vulnerable (in relation to construction and/or availability of resources for recovery).

Table 3-1: Intensity parameters, units and thresholds used in the categorization process.

Hazard	*Intensity parameter*	*Unit of the intensity parameters*	*Value of the intensity threshold*		
			High(I_H)	*Medium(I_M)*	*Low(I_L)*
Earthquake	Acceleration (PGA)	g	0.2	0.1	0.1
Flood	Mean value of water depth (local data)	cm	100	100	100
	Mean value of water depth (global data)	m	0.5	0.5	0.5
Cyclone	Mean wind speed per Admin 2 unit	km/h	80	80	80
Tsunami	Coastal maximum amplitude (wave height)	m	2	0.5	0.5
Coastal Flood	Inundation depth at the coast	m	2	0.5	0.5
Volcanic*	Volcanic explosive index	-	4	4	4
Water Scarcity	Water available per capita	m^3/capita/year[3]	1700	1000	500
Landslide*	Landslide Susceptibility Index	-	5	5	5

** These thresholds are not used because data is preprocessed. When raster data are available, other thresholds will be used.*

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 (see section 5.2), so 0.1g is the intensity threshold associated with PGA for very low, low and medium levels. For high levels, it was decided to fix the threshold to a value typical of more damaging events (causing damage to structures), so shaking exceeding intensity VII (0.2g) are used as the threshold. In doing so, the number of high hazard zones is reduced compared to if a threshold of 0.1g had been used. This serves to preserve the importance high hazard level by not distributing 'high' hazard widely.

Tsunami uses a similar 'stepped' threshold on that basis that previous damage has shown an increased level of tsunami damage when flow depth exceeds 2m; hence this value is used to define the high hazard intensity threshold. On the contrary, for river flood only one value of threshold is considered, since it is considered that main impact of flooding is related to flooded/non flooded status of the area, rather than the inundation depth.

3.1.1.2 Frequency

There are not 'community standard' frequency classes (return periods) for all hazards. For earthquake there are commonly used return periods, previously stated. For flooding, however, there are no *de facto* return periods that are considered in all studies – the most appropriate return periods to use in a study depend on the level of exposure (for example, a country with high populations and economic assets in flood-prone areas may use a longer return period than less-developed regions or regions with small populations). Consequently, it is necessary to allow a range of return periods to be used in *ThinkHazard!* to increase flexibility of which datasets can be incorporated into the tool, and ultimately increasing the number of the number of datasets used.

ThinkHazard! is set up to, by default, use the return periods given in Table 3-2. These are based upon the hazard datasets available during development of version 1. In cases where datasets are not available for a particular return period, the administrator can select datasets with return period closest to the suggested value. More guidance is provided in the peril-specific sections.

Table 3-2 : Suggested return periods for each hazard. [A] For flood, if longer return periods are available, use those; [B] Return periods not applicable to data currently contained in the tool.

Hazard	*Return period associated with hazard categories (years)*		
	High (T_H)	*Medium (T_M)*	*Low (T_L)*
Earthquake	100	475	2475
River Flood	10	50	100[A]
Cyclone	50	100	1000
Tsunami	100	500	2500
Coastal Flood	10	50	100
Volcanic ash	*N/A*[B]		
Water Scarcity	5	50	1000
Landslide	*N/A*[B]		

The justification return periods may be based on likely project or population lifetime experience of a hazard. For example, Table 3-3 provides a justification for river flood hazard in Germany[9]:

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

Name	Return period	Rationale

[9] Federal Office for Citizen Protection and Disaster Support (2010)

High	10 years	In many cases a River Flood event of once per year or once every five years already causes considerable damage. This return period could pose a difficulty because maps for such a short interval are unlikely to be available in all River Flood hazard assessments. Therefore, it may be more appropriate to use the next class for 'high' and then split the current 'low' class into two.
Medium	50 years	An event that would, on average, be expected to occur once or twice in a lifetime.
Low	10000 years	An event most people will not experience and will only be remembered by previous generations. Often the 10.000 year return period will not be available. We propose to use the return period with the highest length (and longer than 50 years) for the 'Low' hazard class.
Very low	No hazard calculated based on the dataset with the maximum return period.	No floods expected based on current climate, current models and data. However, some uncertainty remains.

An alternative approach would be to interpolate between maps for different return periods to obtain a map for the desired period. This would, however, require spatial interpolation for hazard contours bringing in uncertainties in the linearity of event boundaries and intensity distributions, in addition to being computationally expensive. Therefore, using a range of return period as limits appears a better solution. In *ThinkHazard!* version 1, return period have been selected to rank data files' relevance to a hazard category and to be reasonably consistent with the general text of the recommendations. This text is scaled on the preferred value (Table 3-2), but alternative values have been accepted to extend the set of available data. Guidance will be provided to the administrator to help him identify suitable datasets to be processed by the tool.

3.1.2 Preprocessed hazard data

Preprocessing of data into a format suitable for *ThinkHazard!* may be needed if original hazard data are not probabilistic (i.e., do not provide intensity at multiple hazard levels) or not available in raster format. For example, volcanic hazard is not widely communicated probabilistically, so global hazard levels can only be derived based on databases of eruptive histories. Regional and global coastal data, such as coastal flood or tsunami, are often only provided as data points offshore[10], where the available hazard data do not intersect administrative units, preventing use of the general procedure.

Two types of preprocessing have been applied to enable additional data to be included in *ThinkHazard! version 1*:

1. Full Preprocessing
    a. **Non-probabilistic data.** Converts non-spatial data containing non-probabilistic hazard information into raster file(s) containing hazard levels. Data may not contain the hazard intensity and frequency parameters required for the general procedure. The data are fully preprocessed outside the tool, to generate maps showing hazard level (High, Medium, Low, and Very Low). The

[10] Due to the high computational expense of simulating inundation over large area.

preprocessed layers are imported to the tool, which manages the association of hazard levels to administrative units (Mapping). An example of this is the processing of global volcanic location and eruptive history data into volcanic hazard level. See section 5.7 for further details.

b. **Probabilistic data.** Converts spatial data containing probabilistic hazard information into raster file(s) containing hazard levels. Data should contain hazard intensity and frequency parameters required for the general procedure, but may require additional processing for the data to be processed using the general procedure. The preprocessed layers containing hazard layers are imported to the tool, which manages the association of hazard levels to administrative units (Mapping). An example of this is the processing of tsunami data, which converts wave height value offshore (in polygon or point files) into hazard categories by buffering the data to intersect with onshore areas, then comparing wave height values with ground elevation. See section 5.6 for further details.

2. Partial preprocessing
   a. **Probabilistic data.** This option exists for data that only requires translation of intensity and frequency parameters to onshore areas, in order that the general procedure can be used. It converts spatial files containing probabilistic data that do not intersect administrative units, into raster files containing hazard intensity parameters that *do intersect administrative units*. The general procedure can be applied on the partially processed files. Buffering analysis is one method that can be used to extrapolate data at the coast to onshore areas, however, it is often prudent in coastal hazards to account for the effect of topography, as applied in 1b. above.

3.2 Mapping

This step enables the user to view the hazard level for each administrative unit in and around the queried location, on the tool user interface. By clicking on the map the user can select neighboring or hierarchically-associated administrative units.

Mapping is achieved by intersecting a vector polygon layer containing administrative unit level 2[11], with a raster dataset containing hazard level. The raster dataset to use is determined on the basis of hazard, spatial coverage of the data[12], and the data quality procedure (section 2.4). *ThinkHazard!* computes the hazard level for each unit in the global boundary dataset, every time hazard data is refreshed, i.e. imported from Geonode. The

[11] Administrative units are classified with three levels, Administrative level 0 (country), 'Admin' 1 (e.g., state, province) and 'Admin' 2 (e.g., district, county). Administrative unit data is sourced from the FAO Global Administrative Unit Layers (GAUL) dataset (http://www.fao.org/geonetwork/srv/en/metadata.show?id=12691&currTab=simple)

[12] Global data are considered valid for all administrative units. Local data can only be relevant for a limit number of administrative units.

updated hazard levels are stored in the tool processing database, and copied into the public database at a given date for user queries.

Mapping of hazard level is conducted at Admin 2 level. The tool calculates the appropriate hazard level for each cell in the queried raster, identifies the highest hazard level within each administrative unit, and assigns that hazard level to the administrative unit. The hazard level of a higher associated Admin units (1, or 0) is defined as the maximum hazard level in all lower units that it contains (referred to as ‘upscaling’). In the example illustrated in Figure 3-2, the region in Morocco named *Center* (Admin 1) is constituted of eight Admin 2 units. At the Admin 2 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 Admin 1 region Center, is classified as ‘High’.

If only one sub region in a country was categorized as ‘High’ then the whole country would be classed as ‘High’. This is obviously conservative but is applied because the tool (and potentially the user) does not know where in the country the development project is situated. The tool, however, provides a map showing the hazard classes for the administrative level one degree higher than the chosen level, e.g. Administrative level 1 when the user has requested a report for Administrative level 0. This provides the user with information on how the hazard varies over the region of interest.

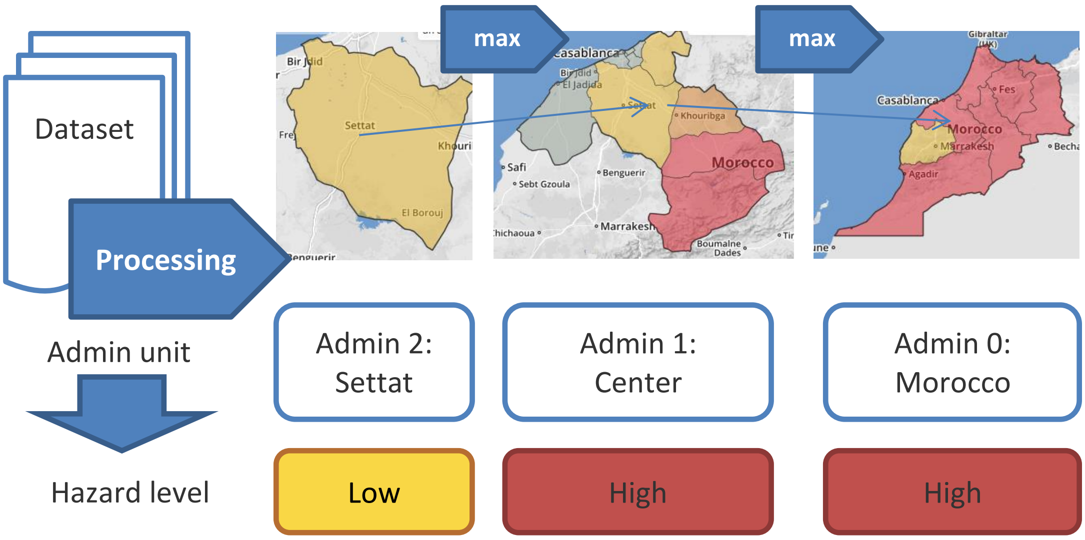


Figure 3-2: Principles of the mapping process.

4 Disaster Risk Reduction (DRR) guidance

One of the most important outcomes of *ThinkHazard!* is guidance provided to the user, to help reduce disaster risk to their projects. This guidance consists of recommendations on how to reduce the risk to projects and links to further resources, such as data sources, event reports, or previous risk analysis relevant to that hazard and/or location. These are described in this section.

4.1 Recommendations to reduce disaster risk

Hazard-specific recommendations comprise three parts:

- A general text, the 'core recommendation' that explains the meaning of the hazard category name and encourages the user to take into consideration the risk in an appropriate way.

- A brief statement on any impact that climate change might have on the hazard. This statement is linked to the hazard, and the country (or country containing the location) searched by the user. This captures regional variation in climate change advice.

- Technical recommendations, consisting of a list of actions to be taken to reduce risk to the project. Each technical recommendation may be repeated for more than one hazard category, but higher hazard categories generally has a greater number of recommendations, comprising the more serious actions that are required to reduce the highest level of risk. 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 recommendations have been developed to provide guidance that minimizes technical language, and provides usable actions specific to each hazard. In *ThinkHazard! Version 1*, recommendations are not project specific, so they can be used for a variety of sectors. This component can be updated in other versions of *ThinkHazard!* to accommodate sector-specific recommendations. Users are able to propose additions or amendments to resources via the user feedback function in the tool. Updates to recommendations are then made by the *ThinkHazard!* administrator.

Initial recommendations were developed by the project team, which comprises hazard experts. This stage included development of the general text and short summary technical recommendations. Following review of the recommendations by the project team, GFDRR contracted external hazard experts to:

1. Review the existing recommendations,
2. Develop new short recommendations as required to remove gaps in coverage, and
3. Develop more detailed, 1-2 page recommendations to provide users with additional detail about actions to take. These remain non-specific to any project or sector.

Recommendations comprise static text, and can be updated by the administrator using the administration interface.

4.2 Additional Information

The user interface includes a section providing further resources to the user. Further resources comprise links to documents, tools, and websites that provide information relevant to the hazard and country (or country containing the location) searched by the user. The resources may provide information on previous projects in the area, general advice on managing risk to a particular hazard, or provide a link to relevant agencies and tools for further analysis. The displayed resources are first filtered by hazard and location, then displayed in chronological order.

All further resources are stored on the same Geonode as hazard data layers, and retrieved periodically for use in *ThinkHazard!*

5 Hazard specific classification methods

5.1 River Flood

This section describes the hazard data, intensity thresholds and return periods used in classification of river flood hazard levels in *ThinkHazard!*.

5.1.1 Hazard data source

River flood hazard is available either from global flood models (GLOFRIS[13], GAR) or 1D/2D flood modeling on a local scale. Flood hazard is typically expressed as water depth per spatial entity (e.g. in the form of a raster cell). The resolution varies between some kilometers for global models up to 10 meters for local flood hazard models.

Flood hazard return periods are typically expressed as 1 event in 5, 10, 25, 50, 100, 500, 1000 and 10000 years with corresponding exceedance probabilities of 1/5, 1/10, 1/25, 1/50, 1/100, 1/500, 1/1000 and 1/10000 per year, respectively.

5.1.2 Intensity

Flood hazard intensity is expressed as the water depth at a specific location. For local models, sometimes flow velocity is used to describe flood intensities additionally. However, as *Think Hazard!* is focusing on a global application, only the water depth is considered as a significant parameter for flood hazard categorization.

The most frequent application of water depth classifications can be found in flood hazard maps, as such maps are made to 'communicate' the severity of the hazard to stakeholders such as people prone to the flood hazard, decision makers and emergency offices. Therefore, in the following, a brief review of thresholds as applied in flood hazard maps is provided.

The recently finalized European joint research project RiskMap investigated several aspects of flood hazard mapping and recommends four water depth classes as follows[14]: < 0.5 m / 0.5 – 1.0 m / 1.0 - 2.0 m / > 2.0 m.

The existing classifications are grounded on the vulnerability of assets which are prone to the flood hazard. For example, in Figure 5-1 three classes are used for an assessment of flood related damages and risks for buildings. The values represent typical thresholds for which significant changes in damages on buildings occur:

- 0.5 m: Flood mitigation by sand-sacks and other preliminary measures not any longer possible. Typical height of tables and light switches.
- 2.0 m: The first floor and its interior are completely flooded.

[13] Winsemius et al. (2013)
[14] Meyer et al. (2011)

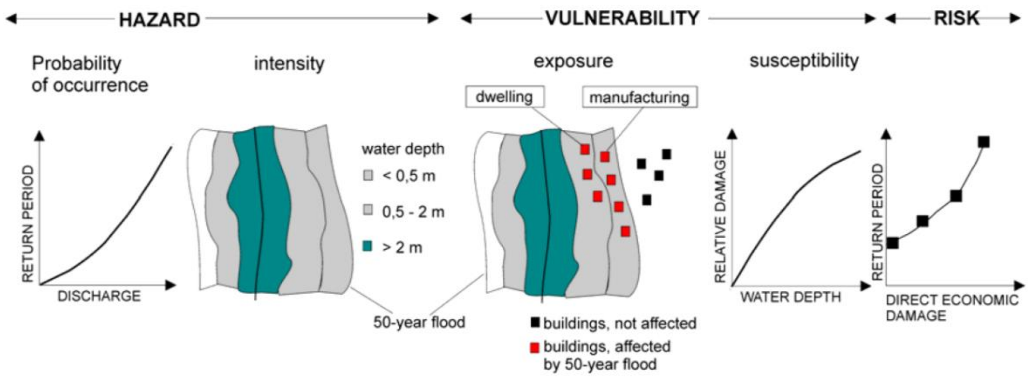


Figure 5-1 : Flood hazard classes in a flood risk study for buildings[15].

Hence, a flood hazard intensity threshold between 0.5 meters and 2.0 meters is recommended for the river flood hazard classification in *Think Hazard!*. The threshold will be chosen based on the sensitivity of the classification and the compliance between the flood hazard and the result of the classification.

In a next step, different flood hazard intensity thresholds have been applied iteratively in order to investigate the results with respect to the hazard classification. For this purpose, local data with a resolution of 90m has been used.

In order to ensure that permanent water such as lakes or reservoirs is not included in the hazard classification, the river flood hazard datasets need to be masked. This is necessary because typically in flood hazard maps (both at global and local scales) the river bed itself also has a water depth, and this flood depth will typically be higher than the water depth in the floodplain. River bed depths, consequently, do not reflect the characteristics of the flood hazard in the flood plain and thus lead to a significant overestimation of the hazard level.

The masking approach focuses at removing grid cells with permanent water from the flood hazard maps. The requirement for applying this approach is availability of high-quality maps showing permanent water bodies. Therefore, the 5 year period flood extent for the global data as a mask for the other return periods. More generally, it is proposed to exclude the river area by replacing the flood depth there by zero (no data) values using the shortest return period available.

Figure 5-2 shows the hazard classification for an intensity threshold of 50cm. Generally there is a good agreement between flood hazard and the classification. However, for one administrative unit (Errachidia), the classification is not correct. In this area, a limited flood extent leads to the exceedance of the classification threshold, while there is no further evidence of a significant flood hazard in this area.

[15] Merz and Thieken (2004) in: Lindenschmidt et al. (2006)

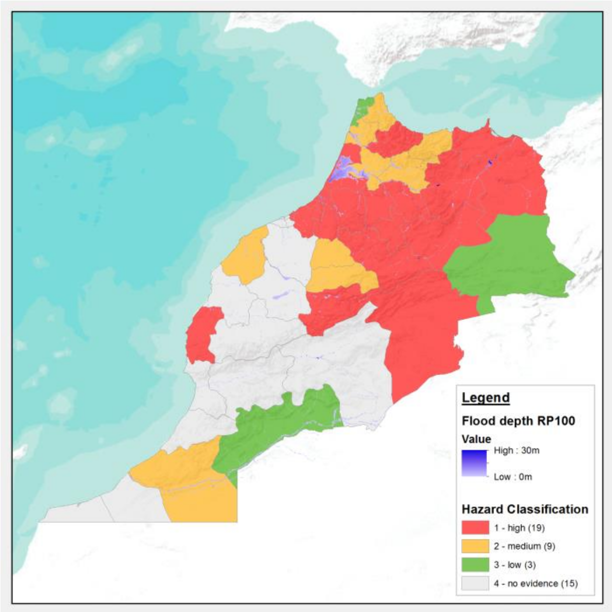


Figure 5-2 : Flood hazard classification using a 50cm threshold for Morocco.

Figure 5-3 shows the comparison between the intensity thresholds of 100cm (left), 150cm (centre) and 200cm (right). It can be seen, that the number of administrative units with medium or high hazard level has decreased significantly.

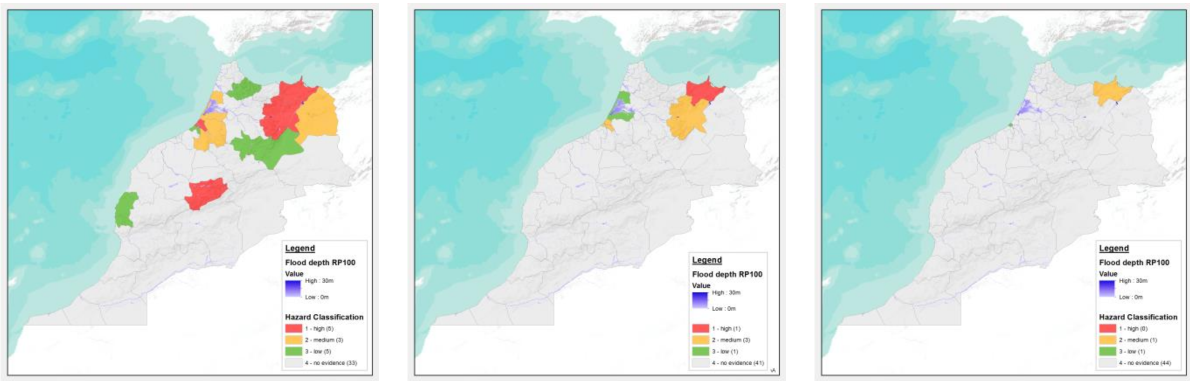


Figure 5-3: Comparison of the hazard classification for 100cm (left), 150cm (centre) and 200cm (right) intensity threshold.

Therefore it is recommended to use the 50cm intensity threshold for the classification of river flood hazards for implementation on the *Think Hazard!* platform.

After masking the rivers in the flood dataset, the mean water depth per admin2 unit was calculated and finally an intensity threshold for water depth (comparable to the 0.1g PGA threshold for earthquakes) was applied of 0.5 meters for the local dataset. This threshold is kept the same for each return period.

The application of the hazard classification methodology using global data focused on achieving comparable results as if using local data, to make sure the results are comparable for areas where either only global or local data is available. It is important to emphasize that, due to the use of different elevation models, hydrological models and hydraulic models as well as differences in the resolution of the models, the flood hazard usually shows different patterns and depths if comparing local and global flood data.

Therefore, the results of the classification have been compared for the intensity thresholds 50cm, 100cm and 150cm with the classification using local data and a threshold of 50cm.

The results are shown in Figure 5-4. It can be seen that a threshold of 50cm leads to an overestimation of the hazard classification, while the threshold of 150cm leads to a significant underestimation. Consequently, a threshold of 100cm provides comparable results, though there are still major differences due to the abovementioned differences between global and local flood models. Among other aspects, the lower resolution of the global model generally leads to higher mean values, as the balance between shallower and deeper water as not as natural as in the local model. Furthermore, the underlying Digital Elevation Model of the global model already has an inaccuracy of about 30cm.

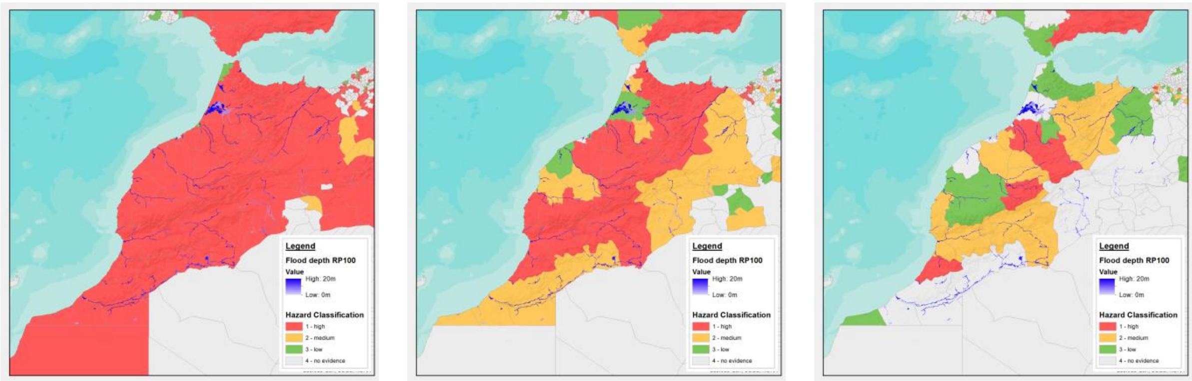


Figure 5-4: Comparison of the hazard classification for 50cm (left), 100cm (centre) and 150cm (right) intensity threshold based on global data.

Therefore, the intensity threshold for each return period is set as 1.0 meter for global models for river flood hazards.

5.1.3 Frequency

Community standards do not exist for frequency classes (return periods) for river flooding. Thus, there are no *de facto* return periods that are considered in all flood studies, as these depend on the level of exposure in a country. For example, a country with high populations and economic assets in flood-prone areas may consider longer return periods (thus less likely events) than less-developed regions or regions with small populations. Beside the higher damage potential, developed countries usually have longer time series of observations they can extrapolate from. Consequently, it is necessary to use a range of return periods to ensure that a hazard map is available which represent a certain hazard class.

For *Think Hazard!* it is recommended to use conservative thresholds, as *Think Hazard!* is focusing on International Development Association (IDA) countries in which investments may be more vulnerable (in type of construction applied and/or availability of resources for recovery in case of a hazard).

For example, the German Federal Office for Citizen Protection and Disaster Support (BBK) (2010) uses the following classes to classify flood hazards (cf. Table 3-3).

Therefore, the following frequency classes are proposed (cf. Table 3-2):

- **high:** If the mean water depth per Admin2 level is exceeded every 10 years, the area is classified as *high hazard* level.
- **medium:** If the mean water depth per Admin2 level is exceeded every 50 years, the area is classified as *medium hazard* level.
- **low:** If the mean water depth per Admin2 level is exceeded every 100 years, the area is classified as *low hazard* level.
- **no evidence:** If the mean water depth per Admin2 level is not even exceeded every 100 years, the area is classified as *no evidence*.

If longer return periods are available, it is recommended to apply those for the low/no evidence hazard classification.

5.1.4 Categorization procedure

The categorization procedure is based on a geo-statistical analysis of the mean water depth per Admin2 level. The mean water depth for the return periods as specified in Table 3-2 is calculated.

The exceedance of the intensity threshold is analyzed using a combined IF-statement, which identifies if the mean water depth is exceeded for a particular return period. If this is true, the hazard classification score is applied. If this is false, the next hazard level is checked. The hazard classification score is joined with the administrative boundaries for the level Admin2. Figure 5-5 shows the results of this analysis based on global data.

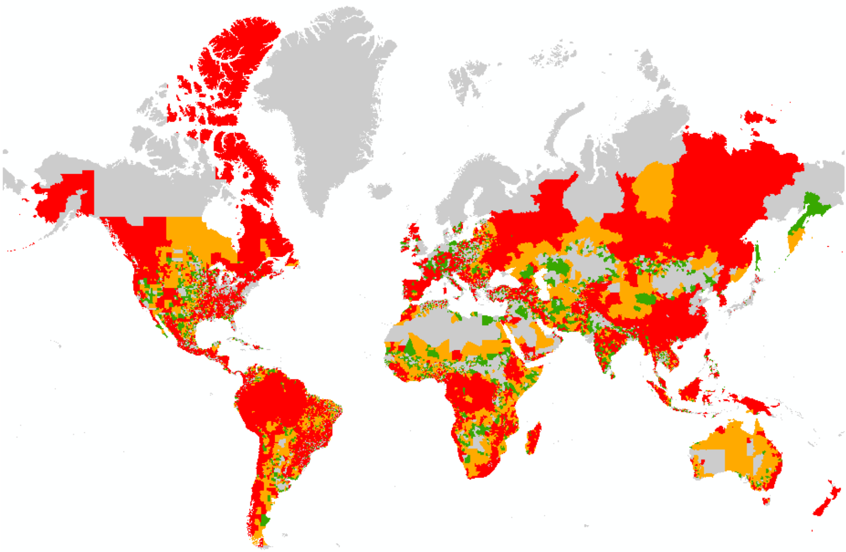


Figure 5-5: Preliminary river flood hazard classification based on Admin2 level.

Next, the largest hazard score per Admin1 level is derived from all underlying Admin2 areas. The same procedure is applied for the Admin1 level to Admin0 (national) level.

5.2 Earthquake

This section describes the hazard data, intensity thresholds and return periods used in classification of earthquake hazard levels in *ThinkHazard!*.

5.2.1 Hazard data source

Earthquake hazard maps generally consist of a grid of the expected peak ground acceleration (PGA) with a 10% chance of being exceeding in a 50-year interval, which translates to a **return period** of 475 years. 50 years is often considered as a standard lifespan for infrastructures. For example, the following map was recently published by the SHARE European project (Figure 5-6, http://www.efehr.org:8080/jetspeed/portal/hazard.psml).

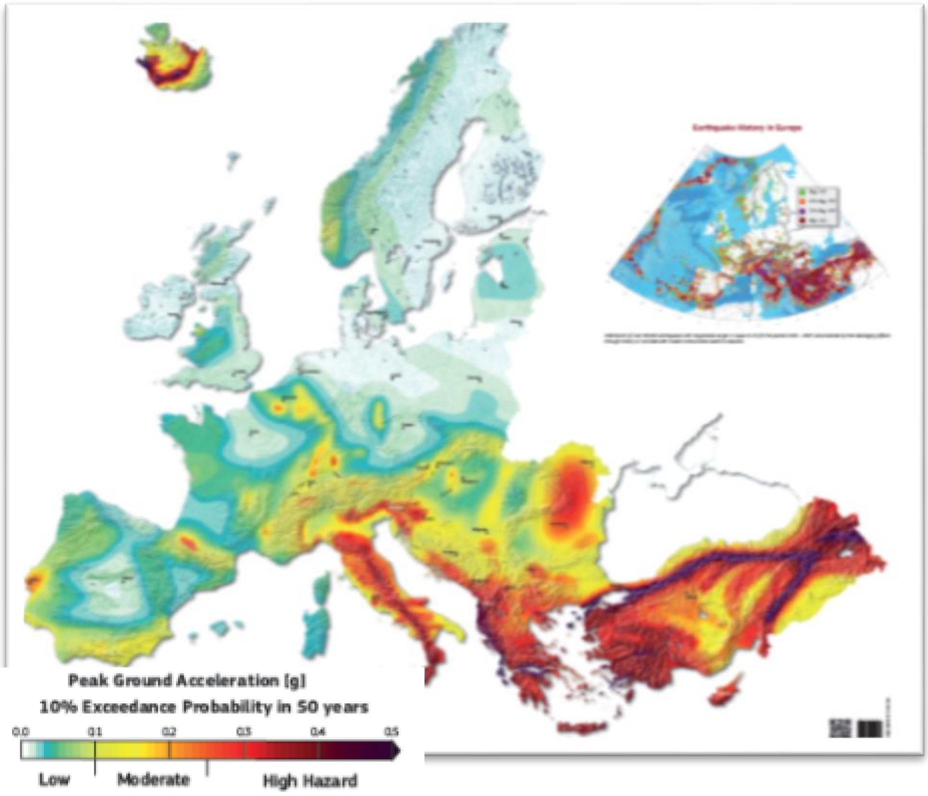


Figure 5-6: Example seismic hazard map (source: FP7-SHARE project).

For different types of infrastructure, such maps are generally produced for different return periods, e.g. for earthquakes return periods of 475, 975 and 2 475 years are quite common. For example, 475 years 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). This is because longer return periods correspond to having a smaller chance that the design level will be surpassed during the lifetime of the structure and hence the risk of damage is lower. Hazard maps are produced by slicing hazard curves calculated for all points in the area at a certain return period (or probability of exceedance).

Standard approach fixes the return period (or probability of exceedance for a given lifespan) and grades the hazard according to the severity of the expected damages via the PGA value.

This is quite easy for civil engineering project, since damages can be simulated from the dynamic behavior of the structure and lifespan is part of the design parameters.

5.2.2 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.

To determine suitable hazard thresholds in *ThinkHazard!* EMS-98 intensity levels were studied in relation to what project planner should be aware in development projects. EMS-98 intensity VI is considered to cause slightly damaging effects to structures, like fine cracks in plaster, and can be felt by most people (Table 5-1). This intensity is considered damaging for most development projects (see section 5.1), so was considered as the threshold for Very Low, Low and Medium hazard. High hazard warrants a higher intensity threshold, to distinguish events that cause significantly greater damage. Intensity VII was selected, because at this intensity people are frightened; cracks appear in buildings, and chimneys start collapsing. It should be noted that the selected thresholds are conservatively low, because of the global and project-generic nature of *ThinkHazard! version 1*. Version that focus on, for example, construction of critical facilities, should consider using higher thresholds more relevant to damaging effects on seismically-engineered structures.

Table 5-1 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

In order to use widely available hazard data, which provide instrumental parameters rather than intensity, a link to instrumental parameters is required. Many studies have explored relations between intensity and instrumental parameters of the ground movement. One commonly used parameter, in particular for earthquake engineering, is peak ground acceleration (**PGA**), which is equal to the maximum acceleration of the ground measured during earthquake shaking at a location. According to widely acknowledged correlations[16] between intensity and peak ground acceleration, intensity VI corresponds approximately to

[16] Atkinson and Sonley, 2000

a PGA of 0.1g and intensity VII to a PGA of 0.2g. Hazard assessment studies generally generate PGA distribution map for a set of return period. We use EMS98 scale to calibrate suitable intensity threshold at the macrocosmic scale, then use the PGA maps to run the categorization procedure.

Hazard dataset provide seismic hazard different units. To enable those datasets to be included in *ThinkHazard!*, a reference table is included in the tool. For example, for earthquake all units are converted to PGA in terms of *g*, for execution of the categorization procedure (Table 5-2).

Table 5-2 Conversion table for units of earthquake data.

Parameter – Unit	**Thresholds**	
	I_H	I_M, I_L
PGA-g	0.2	0.1
PGA-g-per	20	10
PGA-gal	196.133	98.0665
PGA-cm/s^2	196.133	98.0665
PGA-m/s^2	1.96133	0.980665

5.2.3 Frequency

The earthquake field has standard frequencies for presenting earthquake hazard, for which maps are available from many projects. *ThinkHazard!* leverages this availability and consistency in setting the used return periods to maximize the data that can be incorporated. Therefore, *preferred values for return period* associated the hazard levels are:

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

If 2475 year return period data are not available, 1 in 975 years 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. 2500 years is also suitable, and is often available in hazard data produced for the financial sector.

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. As for the 1 in 2500-year hazard, 250 and 500 year hazard data are often available in datasets produced for the financial sector.

5.2.4 Categorization procedure

Earthquake hazard level is categorized within *ThinkHazard!* following an iterative procedure based first on the intensity (PGA) at each cell in a hazard raster layer (e.g., Figure 5-7) and then on the return period of the available layers (values given in Table 5-3). The procedure is illustrated with the preferred threshold values, given above.

Table 5-3 Preferred values for intensity and return period.

	Low	Medium	High
Intensity (PGA), I	$I_L = 0.1$ g	$I_M = 0.1$ g	$I_H = 0.2$ g
Frequency (Return period), T	$T_L = 2475$ years	T_M = 475 or 500 years	$T_M = 100$ years

Incremental procedure:

c. Consider the **100 year** return period hazard map. If **PGA is more than 0.2g** at the location on the map (a pixel of the raster), then the hazard is categorized as ‘**High’**,

d. if PGA is less than 0.2g at the location on the hazard map, then consider the map for ***475 year or 500 year*** *return period*, and the intensity threshold I_M*=0.1g.*

   i. if **PGA is more than 0.1g** at the location on the hazard map, then the hazard is categorized as ‘**Medium’**,

   ii. if PGA is less than 0.1g at the location on the hazard map, then consider the map for ***2475 year*** *return period*, and the intensity threshold I_M*=0.1g.*

      1. if **PGA is more than 0.1g** at the location on the hazard map, then the hazard is categorized as ‘**Low’**,

      2. if **PGA is less than 0.1g** at the location on the hazard map, then the hazard is categorized as ‘**Very low’**.

The hazard categorization procedure is run on each location of the hazard map (e.g., raster cell), to derive the hazard level per cell.

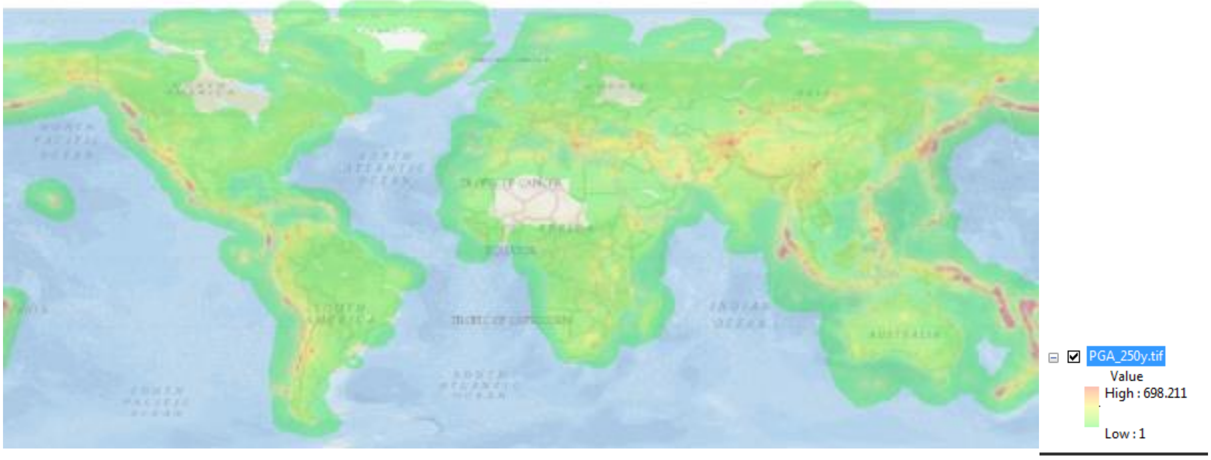


Figure 5-7 : Global seismic hazard at 1 in 250y, PGA in cm/s². red areas show highest hazard. This dataset is an example of input raster data, providing hazard distribution for the hazard categorization procedure (source: UNISDR Global Assessment Report 2015).

5.3 Water scarcity (in relation to drought)

This section describes the data, intensity thresholds and return periods used in classification of water scarcity levels in *ThinkHazard!*.

In *ThinkHazard!* drought refers to hydrological drought. Water scarcity is a shortage of river runoff, in relation to the population density. See chapter 4.1 for the definition.

5.3.1 Data source

The data available at the global scale is provided by VU-IVM and represents 'water availability per capita per year' or Water Crowding Index (WCI) (Veldkamp et al., 2015). Water availability in this case only includes surface water (rivers and lakes) and not soil moisture or groundwater. The data are available both as grids and 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). These data are the mean of outcomes of an ensemble of five global circulation models (CMIP5), run for 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.

The advantage of the use of hydrological drought over meteorological drought is that these data include flow effects. Furthermore, the use of the data per water province has the advantage that the effects of flow within a water province are incorporated and the possibility of solving drought related issues, by distributing water, within one province has been assumed. The flow of water between water provinces however is not included in this assessment. Water stress risks may in reality therefore be lower than represented by these data. Additionally, groundwater and soil moisture are not taken into account in this indicator. In a future version, this is recommendable. Developments on available global datasets related to the different aspects of drought are currently going fast.

5.3.2 Intensity

Falkenmark's "water stress index" (Falkenmark et al., 1989) is based on estimates of water requirements in the household, agricultural, industrial and energy sectors, and the needs of the environment. It was estimated that 1700 m^3 of renewable water resources per capita per year was a threshold, below which a country would experience water stress. For this tool it was decided to use this threshold of 1700 m^3 of renewable water resources per capita per year and not the threshold of 1000 or 500 as the purpose of the tool is to warn end-users of potential hazards. Being on the safe side is key.

5.3.3 Frequency

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 5-8 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.

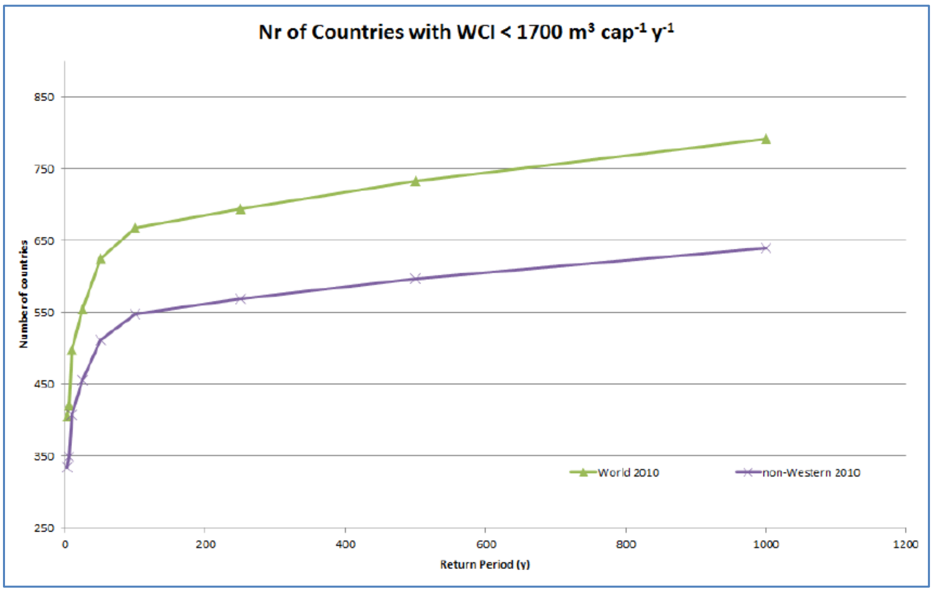


Figure 5-8 : Graph representing the number of countries with a WCI below 1700 m^3 per capita per year in relation to the return period.

5.3.4 Categorization procedure

Based on the return period of water scarcity, a categorization was made of countries, indicating if they experience low, medium or high water scarcity risk. The same methodology is used as for the other indicators, a higher risk and lower risk meaning respectively a higher and a lower frequency.

The rationale for the categorization choice is given in the Table 5-4. Based on the return periods delimiting each category, the "best case" chance (1/best case Return Period) could be obtained.

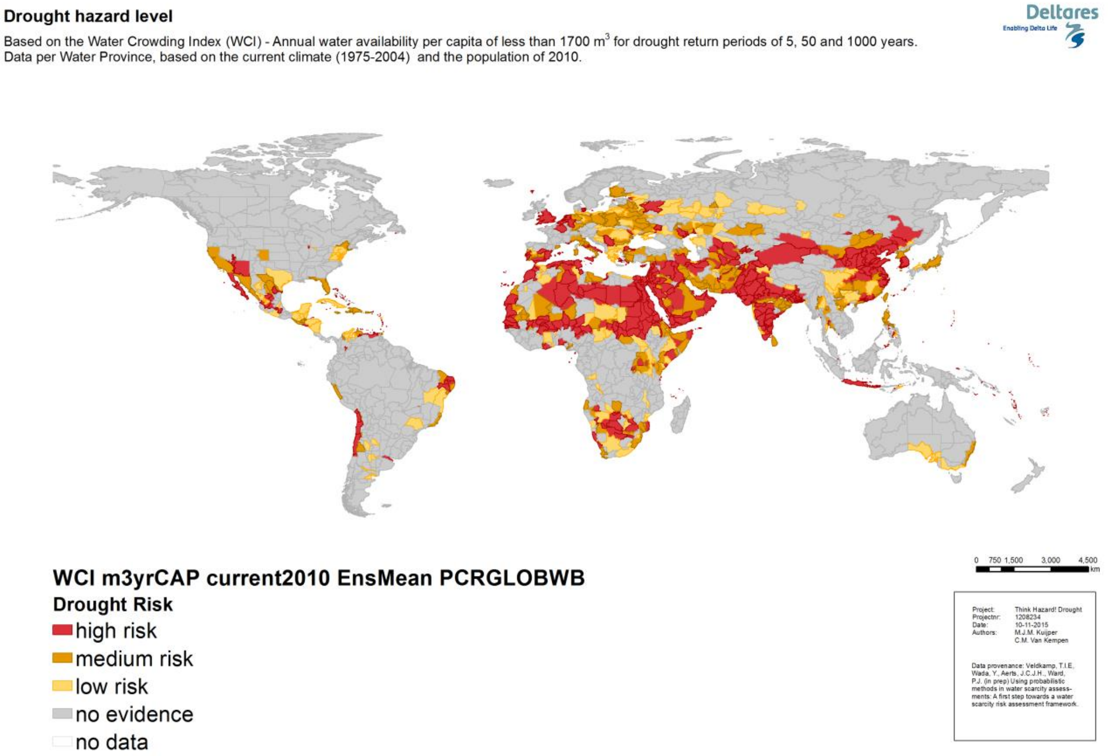
For each category of water scarcity hazard (and return period), the average WCI in the 'admin unit' was subsequently calculated. A trend line and corresponding (power) equation was fitted to this data, allowing to obtain an equation based on which the average return period and the average chance of water stress occurring (1/average return period) could be calculated.

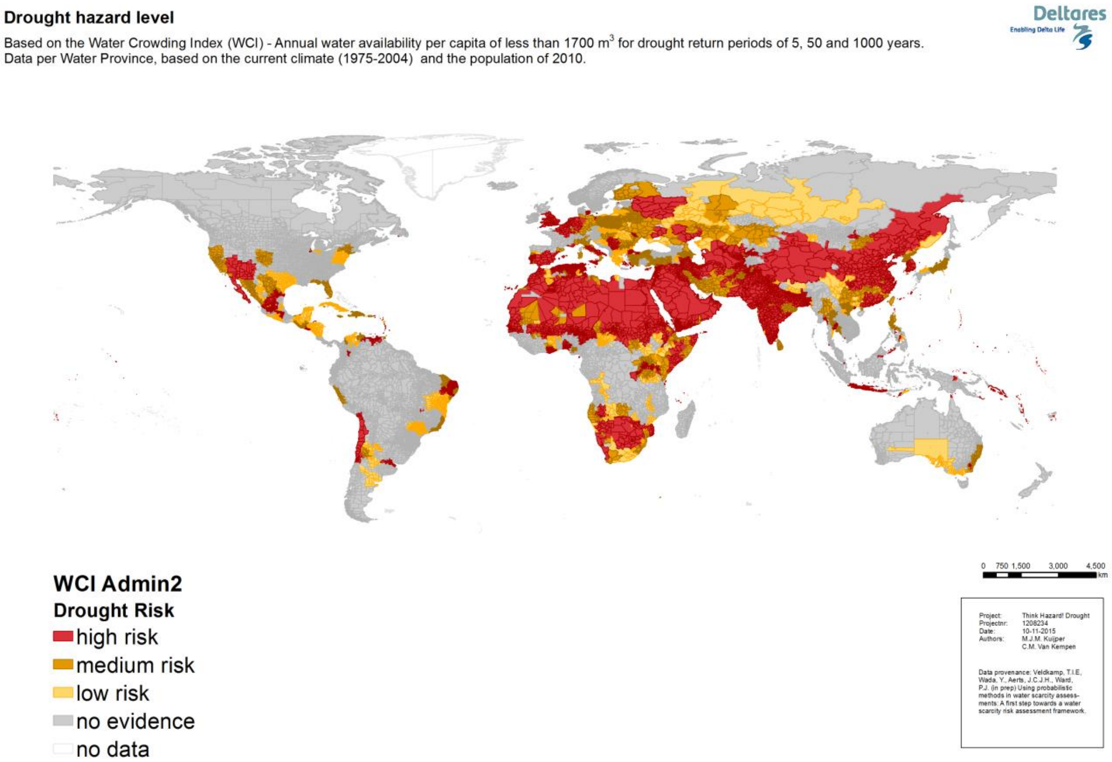
The average and maximum chance are related much like average and standard deviation: both indicate how often water scarcity will occur in a certain water scarcity level category, but one represents the average chance of a drought event occurring while the other represents the chance of a drought event occurring in the "best case" scenario where an area experiences droughts much less frequently than the average country of that same category".

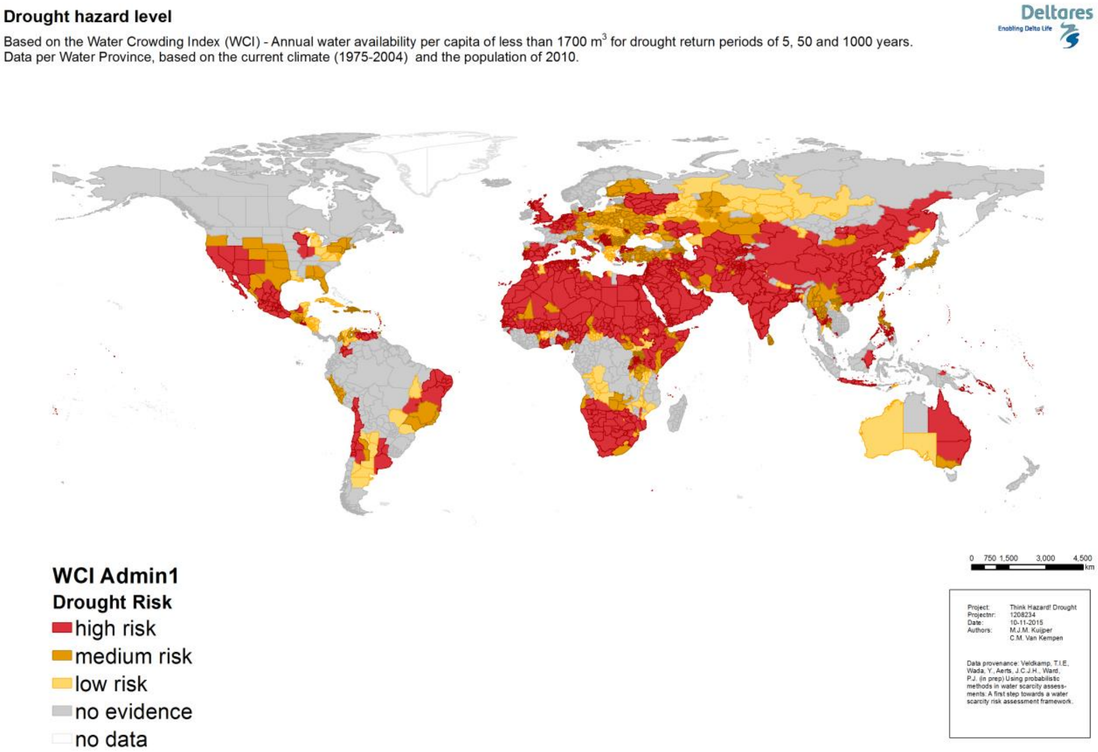
Table 5-4 Water scarcity risk category based on return period and its rationale.

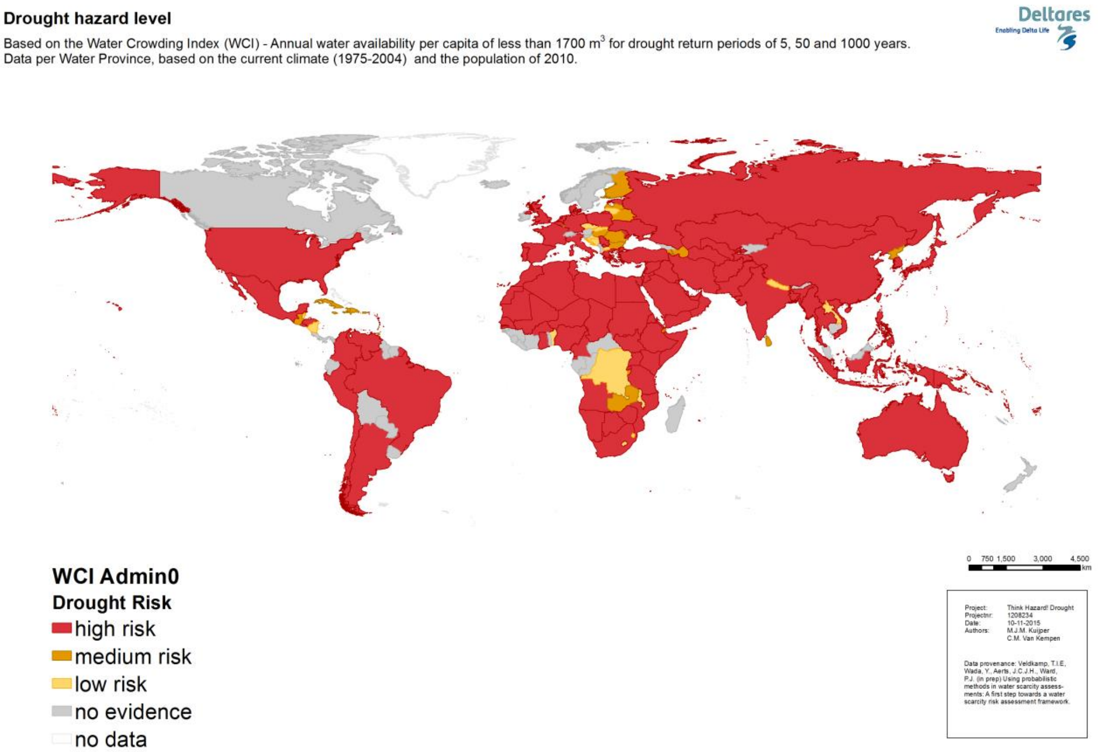
Name of category	Best Case Return Period	Average Return Period	Best case Chance	Average Chance	Rationale
High	5 years	2 years	20%	50%	This category represents areas where water scarcity is common but do not occur every year. Many countries in this category actually experience water stress every other 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	1 000 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	>1 000 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.

The resulting water scarcity map per water province is shown in the figures below, followed by maps of Admin 0, Admin 1 and Admin 2 level boundaries showing the highest risk that occurs within their respective jurisdictions. 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.


Drought hazard level
Based on the Water Crowding Index (WCI) - Annual water availability per capita of less than 1700 m³ for drought return periods of 5, 50 and 1000 years.
Data per Water Province, based on the current climate (1975-2004) and the population of 2010.
Deltares
Enabling Delta Life
WCI m3yrCAP current2010 EnsMean PCRGLOBWB
Drought Risk
high risk
medium risk
low risk
no evidence
no data
0 750 1,500 3,000 4,500
km
Project: Think Hazard! Drought
Projectnr: 1208234
Date: 10-11-2015
Authors: M.J.M. Kuijper
C.M. Van Kempen
Data provenance: Veldkamp, T.I.E, Wada, Y., Aerts, J.C.J.H., Ward, P.J. (in prep) Using probabilistic methods in water scarcity assessments: A first step towards a water scarcity risk assessment framework.


Drought hazard level
Based on the Water Crowding Index (WCI) - Annual water availability per capita of less than 1700 m³ for drought return periods of 5, 50 and 1000 years.
Data per Water Province, based on the current climate (1975-2004) and the population of 2010.
Deltares
Enabling Delta Life
WCI Admin2
Drought Risk
high risk
medium risk
low risk
no evidence
no data
0 750 1,500 3,000 4,500
km
Project: Think Hazard! Drought
Projectnr: 1208234
Date: 10-11-2015
Authors: M.J.M. Kuijper
C.M. Van Kempen
Data provenance: Veldkamp, T.I.E, Wada, Y., Aerts, J.C.J.H., Ward, P.J. (in prep) Using probabilistic methods in water scarcity assessments: A first step towards a water scarcity risk assessment framework.


Drought hazard level
Based on the Water Crowding Index (WCI) - Annual water availability per capita of less than 1700 m³ for drought return periods of 5, 50 and 1000 years.
Data per Water Province, based on the current climate (1975-2004) and the population of 2010.
Deltares
Enabling Delta Life
WCI Admin1
Drought Risk
high risk
medium risk
low risk
no evidence
no data
0 750 1,500 3,000 4,500
km
Project: Think Hazard! Drought
Projectnr: 1208234
Date: 10-11-2015
Authors: M.J.M. Kuijper
C.M. Van Kempen
Data provenance: Veldkamp, T.I.E., Wada, Y., Aerts, J.C.J.H., Ward, P.J. (in prep) Using probabilistic methods in water scarcity assessments: A first step towards a water scarcity risk assessment framework.


Drought hazard level
Based on the Water Crowding Index (WCI) - Annual water availability per capita of less than 1700 m³ for drought return periods of 5, 50 and 1000 years.
Data per Water Province, based on the current climate (1975-2004) and the population of 2010.
Deltares
Enabling Delta Life
WCI Admin0
Drought Risk
high risk
medium risk
low risk
no evidence
no data
0 750 1,500 3,000 4,500
km
Project: Think Hazard! Drought
Projectnr: 1208234
Date: 10-11-2015
Authors: M.J.M. Kuijper
C.M. Van Kempen
Data provenance: Veldkamp, T.I.E., Wada, Y., Aerts, J.C.J.H., Ward, P.J. (in prep) Using probabilistic methods in water scarcity assessments: A first step towards a water scarcity risk assessment framework.

5.4 Cyclone

This section describes the hazard data, intensity thresholds and return periods used in classification of cyclone hazard levels in *ThinkHazard!*.

5.4.1 Hazard data source

Cyclone data is provided from GAR15[17]. 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 km per hour. It is important to note that no return period more frequent than 50 years is available from GAR15.

5.4.2 Intensity

The intensity of cyclones is described by the wind speed, e.g., Figure 5-9. Based on the literature review, an intensity threshold of 80km per hour is applied, which corresponds to the 50-60 miles/hour 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 – first damages occur".

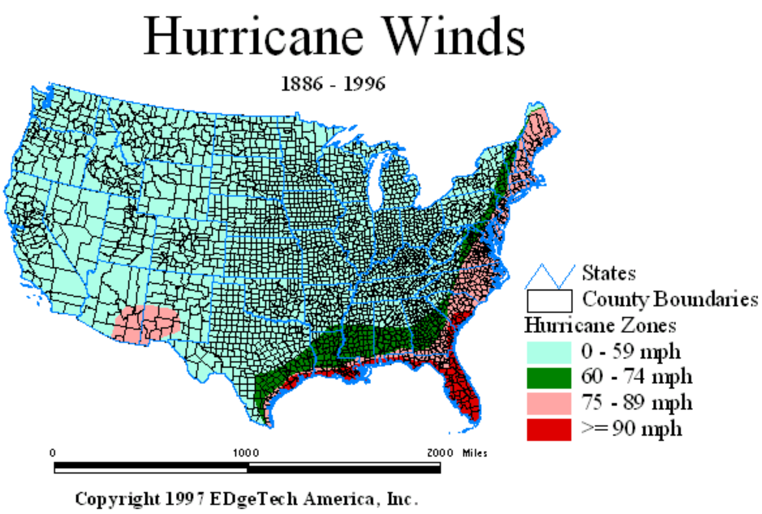


Figure 5-9: Hurricane zoning in the United States, ranging from 60 miles per hour (first damages) up to higher than 90 miles per hour (most severe damages) (Source: http://www.edgetech-us.com/map/MapHurriZone.htm).

[17] http://www.preventionweb.net/english/hyogo/gar/2015/en/home/index.html

5.4.3 Frequency

For cyclones, there are no community standards with respect to the frequency classes or return periods. Therefore, for Think Hazard! it is recommended to use the frequency classification as applied for river floods (cf. Section **Erreur ! Source du renvoi introuvable.**).

However, as there is no dataset with a return period more frequent than 50 years, the classification shifts as follows:

- Exceedance of intensity threshold in 1/50 years: high
- Exceedance of intensity threshold in 1/100 years: medium
- Exceedance of intensity threshold in 1/1000 years: low
- No exceedance of intensity threshold: very low

5.4.4 Categorization procedure

The categorization procedure is based on a geo-statistical analysis of the input data per Admin 2 unit in order to calculate the mean value per administrative unit.

The exceedance of the intensity threshold is analyzed using a combined IF-statement, which identifies if the mean wind speed is exceeded for a particular return period. If this is true, the hazard classification score is applied. If this is false, the next hazard level is checked. Figure 5-10 shows the results of this analysis based on the GAR15 data.

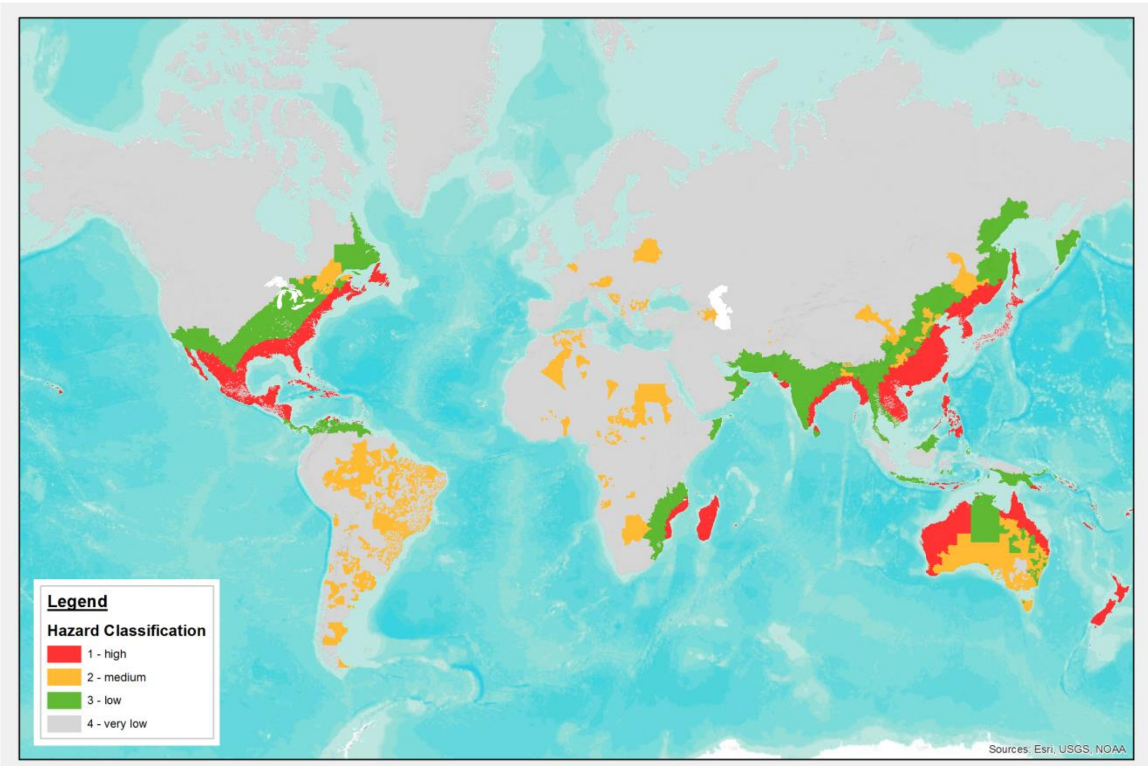


Figure 5-10: Preliminary cyclone hazard classification based on GAR15 global data.

The analysis is carried out on admin2 levels, with a very promising global result:

It can be seen, that the cyclone hazard is mainly threatening the coastal areas both north and south of the equator, where major damages due to high wind speeds have to be expected. It is important to highlight that cyclone hazards are also connected to storm surge induced high water levels, which needs to be taken into account in the *Think Hazard!* recommendations.

As before, the results for Admin 1 are based on the results for Admin2. There, the lowest value (thus the highest hazard classification) is taken as the significant value for the Admin 1 area.

In addition it was tested how the results would look based on maximum wind speed per Admin 2 level. Figure 5-11 shows the results of this analysis. It can be seen that more areas are classified as High hazard, as the intensity threshold is exceeded more frequently. The same observation applies for the other classes. The use of the maximum wind speed thus would lead to a more conservative hazard classification, which should be carefully discussed with regard to a possible overestimation of the hazard level in place.

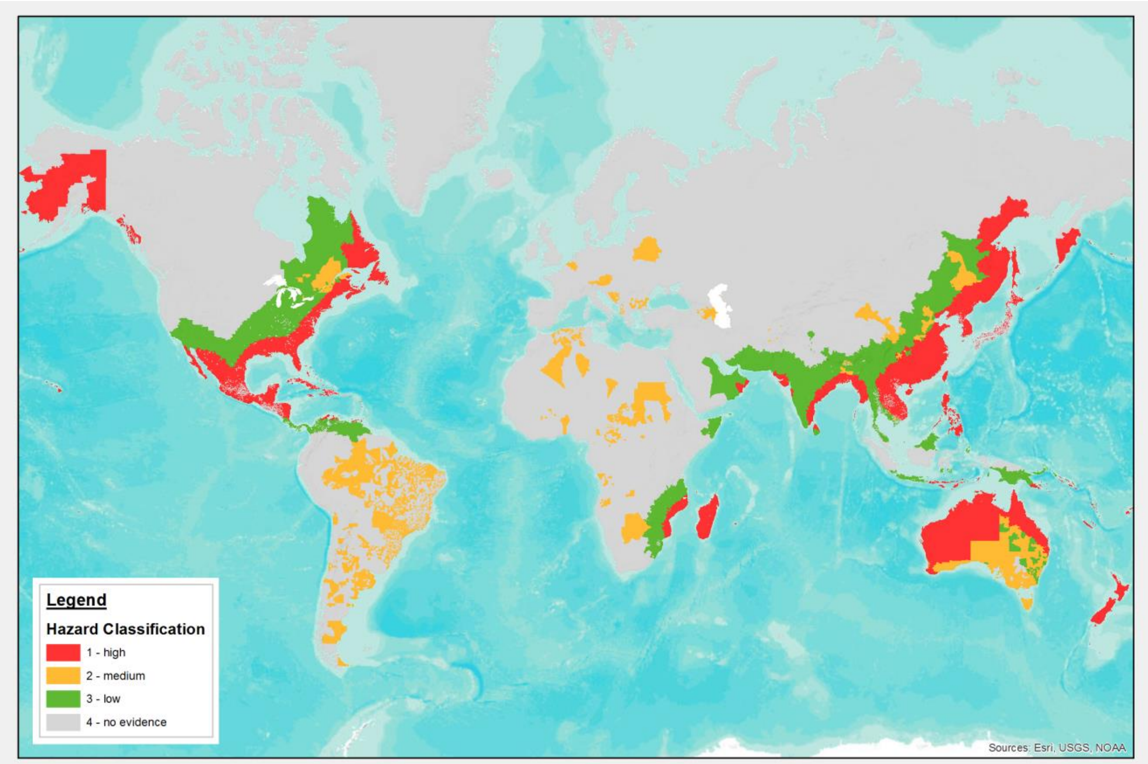


Figure 5-11: Cyclone hazard classification using the maximum value, based on GAR15 global data.

5.5 Coastal Flood

Coastal flooding is caused by a combination of different phenomena:

- Tide, which is natural cyclic sea level variation,
- Storm surge, which is water pushed toward the shore by wing during a storm or a cyclone, on addition to a rise of the sea level due to atmospheric depression,

- Wave set-up resulting from the energy transferred from offshore weaves to the water column at coast.

The maximal height reached by the water is called run-up. It results in two types of threat for coastal area, resulting in potentially intensive flooding:

- The overflowing of protective infrastructures or natural defenses,
- The overtopping of protective infrastructures by waves, which result in large quantities of water shipped over the built protection.

The extent of inundated area and the depth of water depend both of the meteorological mechanisms and topography. Extensive studies and simulation are needed to determine inundation parameters form wave data, since it requires at least an elevation model, bathymetry, and to take into account defenses. In its current state of development, *Think Hazard!* requires inundation maps or inundation depth at the coastline to process the hazard categorization procedure.

Categorization procedure

The categorization procedure is based on the inundation depth calculated at the coast by storm models. Disregarding the nature of the land (urbanization, elevation, defenses) in coastal area, hazard level is assessed from the inundation depth data only.

In GAR15 data, the parameter quantifying storm surge is the **coastal flooding height (h)** obtained by subtracting the simulated run-up at the shore line to the local topographical elevation in-land. Data (Simulated) are available for 10, 25, 50, 100 and 250 year return periods. 10, 50 and 100 year return period data files are selected. For medium and low levels, a Boolean approach is relevant (flooded/non-flooded). Regarding the precision of data, the threshold was fixed to 0.5m, which indicates that land and buildings are flooded (Table 5-5). For “High” level, a more life-threatening threshold was targeted. 2m is chosen, regarding common river flood risk analysis. In this case first floor of building is totally flooded and walls hardly sustain water pressure gradient between indoors and outdoors. These thresholds are fixed without taking into account the dynamics of water (speed, impact).

Table 5-5 Preferred values for intensity and return period of coastal flood.

	Low	Medium	High
Intensity (h), I	I_L = 0.5 m	I_M = 0.5 m	I_H = 2 m
Frequency (Return period), T	T_L = 100 years	T_M = 50 years	T_M = 10 years

Incremental procedure:

a. Consider the **10 year** return period hazard map. If **h exceeds 2m** at the location on the map (a pixel of the raster), then the hazard is categorized as ‘**High’**,

b. if h is less than 2m at the location on the hazard map, then consider the map for ***50 year** return period*, and the intensity threshold I_M*=0.5m.*

iii. if **h exceeds 0.5m** at the location on the hazard map, then the hazard is categorized as '**Medium'**,

iv. if h is less than 0.5m at the location on the hazard map, then consider the map for ***100 year** return period*, and the intensity threshold I_M*=0.5m.*

1. if **h exceeds 0.5m** at the location on the hazard map, then the hazard is categorized as **'Low'**,

2. if **h is less than 0.5m** at the location on the hazard map, then the hazard is categorized as '**Very low'**.

The hazard categorization procedure is run on each location of the hazard map (e.g., raster cell), to derive the hazard level per cell.

5.6 Tsunami

Tsunami hazard results from wave generated by submarine earth movements, caused by earthquakes, landslides or volcanic eruptions. Tsunami is common threat among coastal area that claimed many lives and devastate local economy and environment. The produce huge inundations and the energy of waves is particularly damaging for seafront infrastructures and buildings.

Tsunami data are available as

- A coastal maximum amplitude at discrete points along the coast (wave height) (Figure 5-12),
- Inundation maps resulting from complex models.

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, at the 100 m isobath (indicating water depth). On the other hand, availability of inundation maps is likely to be limited, since models have very high technical and computational requirements.

Categorization procedure

In tsunami data for Indonesia, the parameter quantifying tsunami is the **maximum coastal amplitude (h)**. Data (Simulated) are available for 100, 500, and 2500 year return periods. For medium and low levels, it is decided that a threshold of 0.5 m should be used for onshore hazards, which indicates that land and buildings are flooded. For "High" level, a higher threshold of 2.0 m is used (Table 5-6). This corresponds to evidence of increased damage ratio, from post-disaster surveys in Japan, 2011 (MLIT, 2012).

Figure 5-12: input shapefile showing coastal maximum amplitude locations for 100y return period.

Table 5-6 Preferred values for intensity and return period of tsunami.

	Low	Medium	High
Intensity (h), *I*	I_L = 0.5 m	I_M = 0.5 m	I_H = 2 m
Frequency (Return period), *T*	T_L = 2500 years	T_M = 500 years	T_M = 100 years

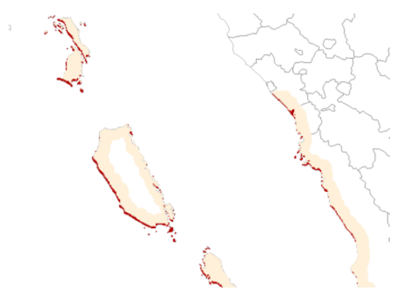
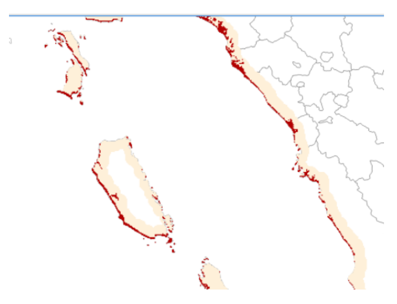
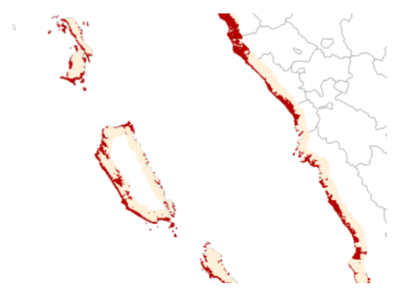
Partial pre-processing of local tsunami data is required in order to take into account the tsunami run-up and inundation process onshore. Due to the associated computational requirements, it is not feasible to conduct full inundation modelling to define inundation height and extent. It is proposed to apply a simple method which is to do for each wave height point, a ground elevation analysis along the coast in the first 10km inland. This distance inland is used, as it is expected to incorporate the maximum inundation of any tsunami.

For an automatic importation into *Think Hazard!* tool, the input data should be provided in multiple-return period raster format.

Pre-processing procedure:

a. Collect SRTM30 plus tiles and extract the pixels covering the coast zone
b. Buffer polygons around the data points from the specified distance.
c. Create a raster format data file aligned on SRTM cells, which assign maximum amplitude of the buffered polygons that intersect a SRTM cell
d. Compare SRTM and data raster to neutralize cells where elevation is higher than the wave height (value set to 0). The produced raster file (one per return period) is the hazard map imported to *ThinkHazard!* (Table 5-7).

Table 5-7 Example of pre-processed rasters for tsunami, Indonesia (Sumatera barat). Red zones show regions were wave height exceeds both elevation and the threshold on the buffered coastline.

		
RP100 years Red: h>2m	RP500 years Red: h>0.5m	RP2500 years Red: h>0.5m

Incremental procedure:

a. Consider the **100 year** return period hazard map. If **h exceeds 2m** at the location on the map (a pixel of the raster), then the hazard is categorized as ‘**High’**,

b. if h is less than 2m at the location on the hazard map, then consider the map for ***500 year*** *return period*, and the intensity threshold I_M*=0.5m.*

   v. if **h exceeds 0.5m** at the location on the hazard map, then the hazard is categorized as ‘**Medium’**,

   vi. if h is less than 0.5m at the location on the hazard map, then consider the map for ***2500 year*** *return period*, and the intensity threshold I_M*=0.5m.*

      1. if **h exceeds 0.5m** at the location on the hazard map, then the hazard is categorized as ‘**Low’**,

      2. if **h is less than 0.5m** at the location on the hazard map, then the hazard is categorized as ‘**Very low’**.

As a result of the procedure, Sumatera barat region (Admin 1) is affected a high tsunami hazard level, but Agam (Admin 2) is affected a medium level.

5.7 Volcanic

A volcano is a place where magma comes to the surface[18], during volcanic activity. Volcanoes present potential threats to people and property, due to phenomena triggered by a volcanic eruption:

- 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.

[18] Most of the general explanations about volcanic hazard are taken from leaflet “Volcanoes and their risks” released during the European project MED-SUV.

    - Pyroclastic flows are a mixture of hot gases and volcanic materials that moves downhill, destroying most things it comes into contact with and depositing large quantities of volcanic debris in its path.
    - Lahars represent a flow of water and volcanic materials that destroy most things they make contact with, and deposit large quantities of volcanic debris. They can occur without an eruption during wet weather when rainfall mobilizes loose volcanic material.
- Distal hazards, affecting larger distances
    - Ash (also called tephra fall), consist of small fragments projected in the air during eruption, and which drop over large area. They can cause health problem, disturb services and agriculture, and their weight can cause roofs to collapse.
    - Gases released into the atmosphere.

5.7.1 Hazard data source

Eruptions can be relatively rare events, and records show that volcanoes can remain quiet for several hundreds of years between to eruptions. Even a volcano that has never erupted in living memory may be a threat. Consequently, much available data on volcanic hazard are not suitable for a classical risk assessment approach based on return period. Ashfall is an exception to this; there are gradually more data becoming available that show potential depth of ash at multiple return periods.

Global databases record the coordinate location, dates, type and magnitude of past eruptions. The Smithsonian Global Volcanic Program (GVP)[19] and LaMEVE[20] databases contain a volcanic eruption index (VEI)[21] which quantifies the eruptive magnitude of past events. These data require full preprocessing to manually classify data into hazard levels for incorporation into *ThinkHazard!* via direct import (see section 5.1.2)

When available, probabilistically simulated ashfall data can be obtained as maps of ash depth, at multiple return periods, enabling their incorporation into *ThinkHazard!* using the general procedure described in section 5.1.1.

5.7.2 Intensity

5.7.2.1 Non-probabilistic data

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 databases have an associated VEI value. Where this

[19] http://volcano.si.edu/
[20] http://www.bgs.ac.uk/vogripa/view/controller.cfc?method=lameve
[21] 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.

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 VEI < 3), are rated as 'Low' hazard. Volcanoes with maximum VEI > 5 are rated High, and those with a maximum recorded VEI 3-5 are 'Medium'.

5.7.2.2 Probabilistic data

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 a number of return periods. Recommended ash depth thresholds for hazard levels are: 0.5 mm, 10 mm, and 50 mm. These values are selected on the basis of 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).

5.7.3 Frequency

5.7.3.1 Non-probabilistic data

Frequency of non-probabilistic data is based on eruptive history (dates of previous eruptions). 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. Conversely, some volcanoes erupt with surprising regularity – monthly or even daily. As a result, it is difficult to reliably and consistently compare frequency across all volcanoes in the GVP database. The most basic measure of eruption frequency is date of last eruption.

In *ThinkHazard!*, the current frequency thresholds used are 2000 years and 10000 years. If a volcanoes last eruption was before 10000, it is considered 'Low' hazard; if more recent than 200 years ago, it is 'High' hazard. Volcanoes with a last eruption date between these values is 'Medium'. Note again that this only applied if no maximum VEI is available for a volcano.

5.7.3.2 Probabilistic data

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.

5.7.4 Categorization procedure

5.7.4.1 Categorization of non-probabilistic data

The categorization procedure for data obtained from eruption databases is based on the VEI index value when available, and according to a date of last eruption threshold as a secondary factor. Based on these data, each volcano is assigned a hazard level 'Low' to 'High'.

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 $\mathrm{VEI} \geq 5$, then hazard level is **high,**
  - If $5 > \mathrm{VEI} \geq 3$, then hazard level is **medium,**
  - If $\mathrm{VEI} < 3$, 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 5-13. This map is uploaded to the *ThinkHazard!* datamart. 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.

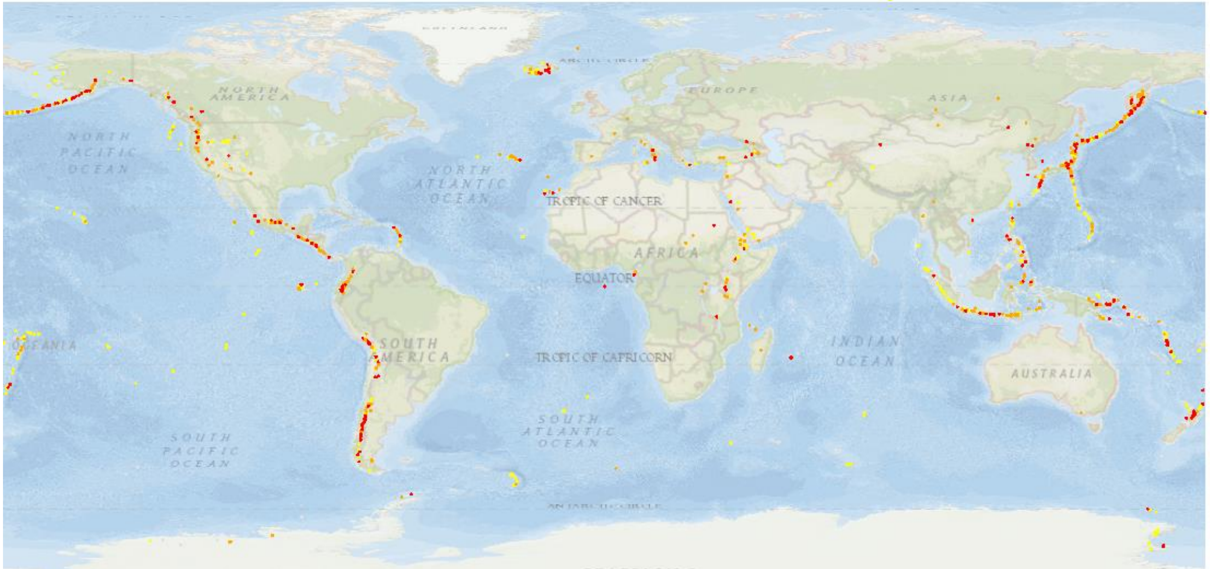


Figure 5-13 : GIS view of the maximum hazard class raster at all locations within 100 km of a volcano.

5.8 Landslide

5.8.1 Hazard data source

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. Landslide triggers include precipitation and seismic activity.

NASA SEDAC 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 Global Assessment Report (GAR) 2013[22] (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.

These global data are consistent with availibility 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.

In both cases, the hazard index is categorized to hazard levels consistent with the *ThinkHazard!* methodology, and imported to *ThinkHazard!* as pre-processed datasets.

[22] http://preview.grid.unep.ch/index.php?preview=data&events=landslides&evcat=1&lang=eng

5.8.2 Categorization procedure

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. The categorization procedure applied to each of the available global datasets is described below:

5.8.2.1 NGI GAR13 data

1. The data are supplied as two raster datasets. Each raster has a range of values (0-15000) indicating the annual frequency of landslide (Table 5-8) due to:
    a. EQ-trigger
    b. Rainfall (precipitation) trigger
2. The two rasters are combined, by taking the maximum annual frequency per grid cell.
3. The resulting raster is reclassified to four hazard levels in GIS, as defined by NGI (2013) as follows:
    a. <180 to Very Low;
    b. 180-320 to Low;
    c. 321-750 to Medium;
    d. >750 to High

Table 5-8 Annual frequency and corresponding raster score for landslide, based on NGI (2013)

Hazard class	Annual frequency per km2 (1x10-4)	Annual frequency per km2	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

5.8.2.2 SEDAC data

1. A single raster dataset is provided, giving landslide hazard as an index value per grid cell,in the range 5-9, adjusted to 6-10 for consistency with other datasets. The index values represent the decile of global landslide hazard.
2. 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:
    a. 6 = negligible (translated to very low for *ThinkHazard!*)

b. 7 = Low
c. 8 = Medium
d. 9 = High
e. 10 = High

6 Further development

ThinkHazard! version 1 will be administrated by GFDRR, using the GFDRR Innovation Lab Geonode for data curation. New data, resources, and recommendation will be accepted on an ongoing basis, with adjustments made by administrators on the basis of data availability and feedback by users in the user interface.

6.1 Obtaining user feedback

User feedback is a vital component of ongoing improvements and updates to *ThinkHazard!*. Users are able to provide feedback on any topic concerning the tool, via the feedback form available on the user interface. Feedback is delivered to the administrator, who will action any required changes and log requests for new features. If the feedback concerns new data for use in the tool, the administrator will follow up to review the data suitability for *ThinkHazard!*.

6.2 Developing new versions

ThinkHazard! uses open-source code, available at https://github.com/GFDRR/thinkhazard. Versions specific to an organization or sector can be developed, in order to provide coverage of particular hazards, or to tailor recommendations more specifically to sector requirements. Sector-specific versions of the tool 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.

New versions can be developed using the open-source code as a basis, by including new recommendations and branding. Further, new functionality can be developed as required, and the tool linked to different data repositories.

7 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

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

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

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

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.

Winsemuis, 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.