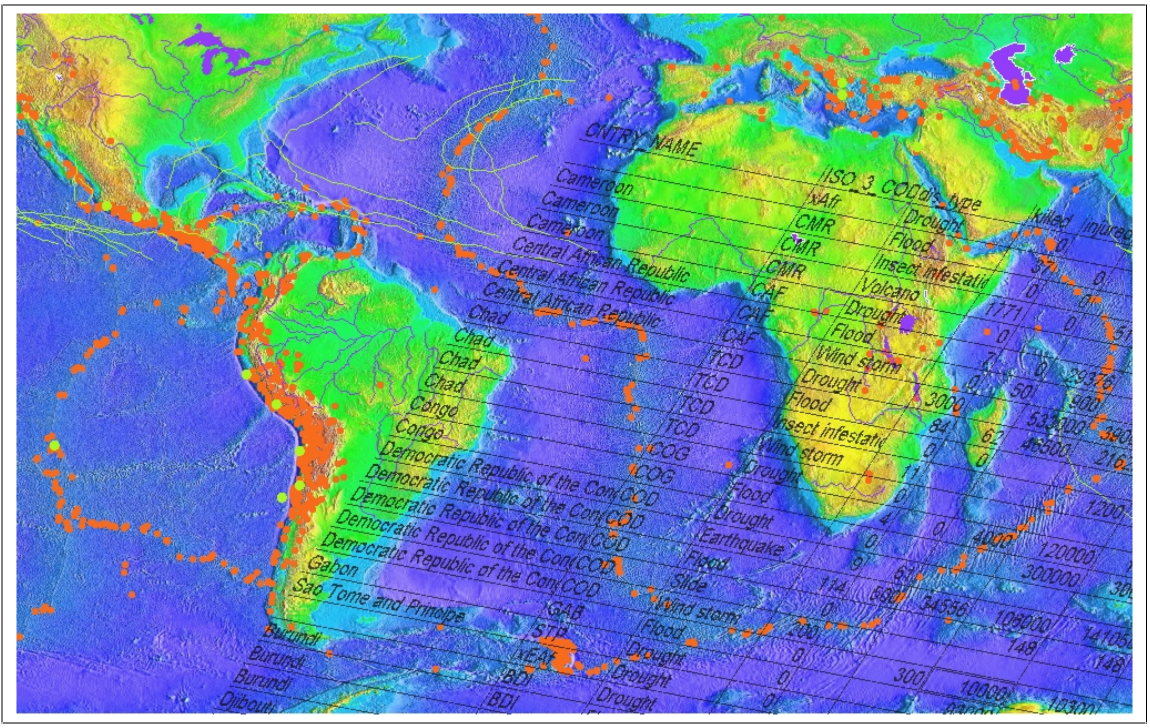
– Feasibility Study Report –
On
Global Risk And Vulnerability Index –Trends per Year
(GRAVITY)

From
The “GRAVITY-Team”
United Nations Environment Programme
Global Resource Information Database - Geneva
(UNEP/DEWA/GRID-Geneva)

For
United Nations Development Programme
Emergency Response Division
UNDP/ERD



Geneva, June 2001

Pascal Peduzzi, Scientific Coordinator (Natural Disasters)
Hy Dao, Scientific Coordinator (Population and Database)
Christian Herold, Spatial Analyst
Damien Rochette, Statistical Analyst
Haris Sanahuja, Scientific Collaborator

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

AVHRR	Advanced Very High Resolution Radiometer
CDIAC	Carbon Dioxide Information Analysis Centre
CNSS	Council of the National Seismic System
CRED	Centre for Research on Epidemiology of Disasters
ERS SAR	European Remote-Sensing Satellite - Synthetic Aperture Radar
FAO	Food and Agriculture Organisation
GDP	Gross Domestic Product
GIS	Geographical Information System
GRAVITY	Global Risk And Vulnerability Index Trend per Year
GRID	Global Resource Information Database
GSHAP	Global Seismic Hazard Assessment Program
HDI	Human Development Index
IDNDR	International Decade for Natural Disasters Reduction
IFRC	International Federation of the Red Cross
ISDR	International Strategy of Disaster Reduction
NGDC	National Geophysical Data Centre
NOAA	National Oceanic and Atmospheric Administration
OCHA	Office of Coordination of Human Affairs
OLS	Ordinary Least Squares
OFDA	Office of US Foreign Disaster Assistance
PGA	Peak Ground Acceleration
UNDP/ERD	United Nation Development Programme, Emergency Response Division
UNEP	United Nation Environment Programme
USGS	United States Geological Survey
VEI	Volcanic Explosivity Index
WFP	World Food Programme
WHO/OMS	World Health Organisation
WMO	World Meteorological Organisation

INTRODUCTION

Disasters and development

On 14 April 1912, the “Titanic” drawn after hitting an iceberg. Out of the 2227 passengers, only 32% were rescued by the ship “Carpathia”. It appeared that the percentage of victims varied according to the travelling class of the passengers, as only 25% of third class passengers survived. Several hypothesises were attempted to explain such difference. They ranged from higher consideration of the crew toward first class passengers, which may have been prioritised (socio-economical factors), to the proximity of the first class passengers from the lifeboats (spatial factors) or by the fact that some of the people in the third class could not speak any English and therefore may not have received the appropriate instructions (access to information, education).

The Titanic victims are of course not the subject of this study, however similarities can be drawn when assessing the number of victims while confronted to a natural disaster. For the same scale of event, a significant discrepancy in the number of victims can be observed, depending if the disaster is located in a developed or less developed country. The countries with least economical power present a lack of means for setting appropriate preparedness, or for improving their capacities of both response and mitigation. The reverse statement may also be true: as high exposure toward natural disasters could lead to slow economical development by scaring the investors away, or by requesting costly infrastructures. Over-arching development differences, countries are not equally exposed to natural hazards. Differences in geophysical factors (slopes, elevation, proximity from the shore or geological fault, inter-tropical location, …) are parameters leading to higher occurrence and severity of hazards. Consequently, the risk of human losses varies extensively according to the differences in population density, level of development, and geophysical parameters enhancing the human vulnerability, the occurrence or the severity of hazards.

If some infrastructures can be designed to decrease some disasters occurrence (slopes stabilising against landslide, dam to contain sudden floods,…) nothing can be done to avoid the occurrence of earthquakes or cyclones. Although geophysical factors should be taken into account when planing human settlements, means for improved preparedness and capacity of response should be developed. Hence the need for developing a culture of risk and to achieve a better understanding of the vulnerability concept through the identification of its different components, thus enabling to provide more specific aid and development for reducing human vulnerability.

The World Disaster Report

The purpose of the World Disaster Report Project from United Nation Development Programme, Emergency Response Division (UNDP/ERD), aims to encourage governments to undertake actions for reducing their population vulnerability. This could be achieved by improving their preparedness, capacities of response and mitigation. It is hoped that the creation of an index, where categories of country from least to best prepared can be ranked, will stimulate the competition between countries in order to improve their status. It could also be used to identify the countries with the highest needs. This index should be derived using appropriate available factors identified through a scientifically sounded method, in order to reach a broad acceptance. It will also request the elaboration of indicators enabling to measure the efforts of the countries.

Risk can be approached in several way, insurance’s and banks have their own way of computing the risk, which take into account financial and insured lives losses. However in the view of a development or aid organisation, human vulnerability need to be approached primarily in terms of human losses and secondly in terms of financial losses only if this affect or threat populations.

If, intuitively, one can understand that improved infrastructures and building constructions, lower density of population, access to information and other favourable socio-economical or geographical factors are leading to a lower vulnerability. Deriving a formula that directly connects the number of victims to the contextual factors is a much higher step. It requests comprehensive set of spatial data to derive the exposure of the different country (occurrence, severity and size of areas affected), as well as precise socio-economical factors which can be correlated with the number of victims recorded.

Objectives of the present study: Statistical analysis

The objective of the present study is to provide UNDP/ERD with a statistical analysis, that aims to gather, prepare the information in order to highlight factors related to human vulnerability and allows the comparison between countries (thus requiring extensive normalisation). The tasks undertaken include the identification and treatment of the most accurate available data sets on geographical, and geophysical parameters to determine the exposure of the countries. As well as socio-economical factors that are connected trough a thorough analysis with the information recorded on human losses per event extracted from the International Disaster Database from CRED[1]. All data sources were registered using meta data information to ease further research and provide references for accuracy assessment.

The results provided in this study delineate a strong potential of these data to approach the concept of human vulnerability. However, they also highlight numerous pit falls and limitations, which will be extensively discussed further. It appears that trends exist, but caution should be taken as the broad precision of the original data is leading to significant margin of error. UNEP/DEWA/GRID-Geneva team has explored patterns, tendencies and possibilities for data extrapolation and margins of error. The outputs achieved are promising, given the time constraint, and include a ready to use database and spatial layers of information, the identification of most relevant factors. A method for modelling an index is proposed in these pages and will only need choices from end users in order that the index could fit with UNDP strategies and goal. Several factors were identified and if selection of criteria and weight are implemented in the model (i.e. imply the choice of how the risk is measured). A multi-criteria analysis could then be extended to drought and other type of disasters if appropriate data are provided. A list of recommendations can be found at the end of the study. They include methods for validation and for connecting affected population and victims in a more precise (but time consuming) way.

---

[1] Collaborating Centre for Research on the Epidemiology of Disasters (CRED) EM-DAT: The OFDA/CRED International Disaster Database – www.cred.be/emdat – Université Catholique de Louvain - Brussels – Belgium.

1. GENERAL MODEL

1.1. *Concepts*

A multitude of definitions and terms related to risk and vulnerability can be found in the literature. Although most of the definitions are similar, some divergences exist. The following definitions do not have the pretension to be THE definite appropriate terminology, but are provided to ensure that the concepts used in the present study are well perceived by the readers.

Figure 1. Definitions

Capacity:	*The ability to protect one's community, home and family and to re-establish ones livelihood (Anderson and Woodrow; 1989).*
Disaster:	*A sudden calamitous event producing great material damage, loss and distress. (Webster's Dictionary, definition found in Carter 1991)*
Early warning:	*A process that provides timely information so that communities are not only informed, but sufficiently impressed, that they take preparedness actions before and during the anticipated hazardous event. (IDNDR)*
Hazards:	*Potential threat to humans and their welfare (Smith, 1996).*
Intensity:	*When direct measurement are not possible, indices incorporating human variables of destruction are used to approach magnitude. (Tobin & Montz, 1997)*
Livelihood:	*Is the command an individual, family or other group has over an income and/or bundles of resources that can be used or exchanged to satisfy its needs (Blaikie et al., 1996)*
Magnitude:	*Magnitudes of geophysical events rest primarily on scientifically based measures of the strength of physical process (e.g. wind speed, energy released by earthquakes,…). (Tobin & Montz, (1997 p.52-53)*
Mitigation:	*Actions that reduce damage and loss. These include measures to reduce the physical hazard, to provide structural and non structureal mitigation and increase preparedness… Mitigation can be viewed from the point of view of vulnerability, vulnerability reduction, and popular attempts at coping and self-help (Blaikie et al, 1996).*
Natural hazard:	*Represents the potential interaction between humans and extreme natural events. It represents the potential or likelihood of an event (it is not the event itself). (Tobin & Montz 1997)*
Physical Exposure:	*Reflects the range of potentially damaging events and their statistical variability at a particular location. (Smith, 1996).*
Prevention:	*Prevention is saying no to the hazard. Mitigation is saying no to vulnerability. (Wilches Chaux, 1989)*
Preparedness:	*Reflects the degree of alertness immediately before the onset of the hazard; for example, arrangements for emergency warnings to be issued and the effectiveness with which public officials can mobilise an evacuation plan. (Smith, 1996).*
Risk:	*A measure of the expected losses due to hazard event of a particular magnitude occurring in a given area over a specific time period.). (Tobin & Montz 1997)* *Note: in this study, only human losses (killed and injured persons) were taken into account.*
Severity:*	*Terms that refer to both size and strength of a given event. (Peduzzi, 2001)*
Size:*	*The area affected by an event, it represent the spatial component of the severity (e.g. ponctual, small scale or large scale event,…). (Peduzzi, 2001)*

Strength:*	*Either measured in magnitude or in intensity, this term will refer to the force released by a natural disaster and represent the energy component of the severity. (Peduzzi, 2001)*
Vulnerability:	*The extent to which a community, structure, service or geographic area is likely to be damaged or disrupted by the impact of a particular hazard. (Tobin & Montz 1997)*

** Concepts that had to be specifically developed for the present study.*

1.2. General approach

Hazards, Vulnerability and Risk

Hazard

In hazard is taken into account the type of hazard (floods, earthquakes,…), the probability of occurrence (expected average number of events per year) and the severity (strength and size of events).

The hazard allows the comparison between countries that may be affected by small strength but numerous events, with countries that are affected rarely by significant strength events, at the end – for a same vulnerability - the casualties may be comparable. In the hazard, the size of the area is also taken into account. Usually the size of the area affected is connected with the strength of the disaster: significant strength usually affects larger area. However, this may not be always the case.

To model the population affected by hazard, databases of realised risk (events that have occurred) were downloaded from reliable organisations websites. This was made for earthquakes, volcanic eruptions, tsunamis, floods and wind storms (including cyclones, typhoons,…). The size of affected areas and strength was taken into account when available.

Vulnerability

The vulnerability can be separated into different components, namely geophysical, socio-economical parameters and mitigation capacities.

The first component can be described as the extent to which geophysical factors are enhancing a potential threat for a specific population, for example a low elevation along the sea shore are geophysical factors enhancing vulnerability for Tsunami, independently if the area is exposed or not to the hazard. However, geophysical factors should be taken into account when planning human settlement. The lack of consideration for them can be seen as an indicator of a need for improving local/national policies. The data requested for approaching the geophysical factors of vulnerability, are of high precision (detailed scale). This explain why they can not be taken into account in a global index, and therefore in this study.

The second component, socio-economical factors, is the most significant in a development perspective. It includes various and numerous parameters depending on cultural, technical and economical factors of the society itself. Some of the parameters (i.e. even most of them) can not be easily measured (e.g. the quality of infrastructures, the political willingness, the capacity to achieve appropriate level of planning) and need to be approached in an indirect way. It is the aim of the statistical analysis to identify them with all appropriate data available. To approach the vulnerability, the following data sets were identified and analysed. It includes: Gross Domestic Product (GDP), GDP growth, literacy rate, life expectancy, level of corruption, population, population density, population growth, and urban population growth.

There is a need to find indicators for measuring mitigation, which is the level of actions taken to reduce the population vulnerability. At the moment these indicators are not available and still need to be defined and measured.

Physical exposure

The hazard multiplied by the population provide information on the number of person living in an exposed area. This is called the physical exposure. The computation of physical exposure was performed using a model of population crossed with frequency and identification of areas threaten by hazards derived from several sources of information.

Risk

A multitude of definitions and formulae can be found in the literature that define the concept of risk. It depends mostly on what type of risk is observed (i.e. financial, human losses,…). Here the study concentrates on the risk faced by population, in terms of wounded and killed while confronted to natural disasters. To approach the above definition of risk, the expected losses over a given time period need to be computed. This lead to the formulation of risk that takes into accounts several components. The probability of occurrence and severity of a specific hazard for a given area and length of time, the vulnerability of the population and the capacity of mitigation, this last could be introduced in the vulnerability or taken separately, depending on authors.

Formula

For a hazard i,

$$\text{Risk}_i = (\text{Hazard}_i - \text{Prevention}_i) \times [\text{Population} \times (\text{Vulnerability}_i - \text{Mitigation}_i)]$$

Where: *Hazard depends on frequency and strength of a given hazard,*
Prevention, is the level of actions undertaken for decreasing the frequency or the strength of hazards (e.g. dam to prevent floods, consolidated slopes against landslides,…).
Population is the number of person living in a given area.
Vulnerability depends on socio-politico-economical parameters of this population, geophysical parameters of the area concerned which needs to be taken into account for human settlements and Mitigation represents the level of actions taken to decrease the population vulnerability (e.g. better aid services, improved information and appropriate strength of building parameters…).

As no information where available for the moment on both mitigation and preparedness, the model was based on a simplified version of the previous equation:

$$\text{Risk}_i = \text{Hazard}_i \times \text{Population} \times \text{Vulnerability}_i$$

The Hazard in the formula reflects the frequency of occurrence and the severity of a specific hazard. It can be computed as the sum of events, multiplied by the severity (size of area affected and strength -measured either in magnitude or in intensity) divided by the length of time taken into account.

$$H_i = (\Sigma\, E_i \times S_i)/ t$$

Where: *H is the Hazard for a specific type of event*
E represents the events
S is the severity of each individual event
t is the length of time (e.g. number of year for a yearly measured of hazard)

The hazard multiplied by the population is the **physical exposure**, i.e. the number of person living in an exposed area. The risk can then also be expressed as:

$$\text{Risk} = \text{Physical exposure} \times \text{Vulnerability}$$

If the number of realised risk is known, as well as the physical exposure, then the vulnerability of a population can easily be derived:

$$\text{Risk} / \text{Physical exposure} = \text{Vulnerability}$$

If the probability of occurrence is null or if the expected strength is too low to affect a human community, then the risk equal zero:

$$\text{Risk} = 0 \times \text{Population} \times \text{Vulnerability}$$

In the same way, if the population is null in a given area, then the vulnerability is null and the risk is null.

$$\text{Risk} = \text{Hazard} \times 0$$

The expected losses due to natural hazards are equal to the sum of all types of risk faced by a population in a given area.

$$\text{Risk}_{\text{Tot}} = \Sigma\,(\text{Risk}_{\text{Flood}} + \text{Risk}_{\text{Earthquake}} + \text{Risk}_{\text{Volcano}} + \text{Risk}_{\text{Cyclone}} + \ldots + \text{Risk}_{n})$$

Providing the total risk for a country induces the need to estimate the probability of occurrence and severity of each hazard, the number of person affected by them, the identification of population vulnerability and mitigation capacities. This is of course not possible in absolute, however the aim is to provide indicators which will be refined years after years in order to approach the concept of risk.

Spatial units

The spatial definition of vulnerability and risk is a crucial topic. First, a distinction must be made between display units and observation units. In the context the present study, display units are countries: the vulnerability/risk evaluation is presented on a country by country basis, according to UNDP requirements. But collection of data is not only performed at the country level; for instance, data on a particular event refers to the area affected by this event, not to the country as a whole. In the context of the present study, several observation units were considered:

Table 1. Spatial units

Observation units	**Remarks**
Countries	Most of the socio-economic vulnerability factors are only available at this resolution (GDP, literacy rate, life expectancy, HDI, etc.
Areas at risk (=all potential areas of disaster, where probability of occurrence > 0)	Defined by the probabilities of occurrence of disaster types. Allows for the evaluation of population/areas that can potentially be affected by disasters
Area of a particular event	Extent of a particular event. Losses reported in the CRED database implicitly refers to this type of area, but in most cases no information on the disasters location or extent is provided (see chapter on disaster data, p. 48)
Pixels	Some data like population density, probabilities of cyclones are already available according to this type of regular grid

The spatial definition of risk/vulnerability differs if only damages are considered or if causal factors are to be explored. Spatial circumscription of damages (realised risk) is in principle relatively easy to depict, although it may also depends on the time frame considered (direct or indirect, induced damages). On the contrary, the spatial extent of causal factors does not necessarily coincide with the observed damages: for example, illegal occupation of exposed slopes by migrants in a region may be caused by the disastrous economic situation in an other region.

In the methodology developed to estimate vulnerability from socio-economic indicators, the figures used were only available at the country level (except for population), which might be not sufficient or even not relevant. On the other hand, hazard data, originally raster grids or vector coverages, have been aggregated to produce figures on a country by country basis.

The correct use (i.e. appropriate scale, pixel size, type of representation, …) of data at various spatial resolution is a major concern when performing environmental modelling, both from the GIS and from the geographical point of view.

Temporal units

Like the spatial definition of risk/vulnerability, its temporal definition is subject to discussion. First, the temporal circumscription of a disaster is very different if only direct losses are considered or if longer term and/or indirect effects are also included.

The periodicity of disaster types (centuries, decades, years, …) is also a very important aspect. Considering vulnerability, repetitive disasters have an influence on the future capacities of response and recovery of a country. Considering data availability, the access to information on natural disaster (number of events, number of victims,…) has considerably risen in recent years following the significant improvement in telecommunication technologies, however such rise is not uniform in all the regions world-wide. Furthermore, the 20-30 years long time-series of more or less complete records provided by various databases may not be adequate to depict geological or climatic phenomenons following trends over hundred or thousand years.

Finally, when exploring relations between disasters and vulnerability, it may appear that causal factors may be shifted in time as compared to observed disasters: the actual vulnerability of a country may be caused by past economic situations.

In the present study, the model of vulnerability factors is based on the analysis of direct human losses on a disaster by disaster basis (as reported by CRED, with all its inherent limitations including the definition of the disaster type). The observed losses (realised risk) over a 21 year period are then used to validate the estimation of risk provided by the socio-economic factors (see chapter 4.)

1.3. Data availability, precision and limitations

Vulnerability and hazard have been the subject of numerous researches from local to global scale. Extrapolations from local researches to global scale are rarely applicable as data may not be of comparable formats or simply not available. The difficulty when trying to approach the risk at global scale is to find relevant and available indicators allowing a comparison between all countries. If a model requests a large amount of inputs, the chances that such model will never be used, by lack of data or by too fuzzy data, are significant. On the other hand a model based on too few parameters will lead to large gap between observed facts and expected figures. The “art” of modelling consists on identifying appropriate data at relevant scale and precision for expected aim. The UNEP office Global Research Information Database (GRID) has been collecting global database for decades and has extensive experience in dealing with such data and on how to combine them. Given time limitations, some of the global data could not be used in this present research (drought, extreme temperature) other data could not be used by lack of detailed data (tsunamis) or by lack of global coverage (landslides).

Socio-economical and/or political parameters are very difficult to measure and needs to be approached indirectly by other available indicators. However, is it possible to approach the quality of infrastructures by the Gross Domestic Product? Is it possible to derive the access to information through level of education and literacy rate?

Finally, is it possible to overlay information provided at country level with more precisely localised information? Approximations, subjective choices had to be made, however the purpose of this study was not to provide a final Global Risk And Vulnerability Index Trends per Year (GRAVITY) or a final country ranking, but was to ascertain whether available identified data with their approximation could be used for such purpose. How accurate are the calculated risks when compared to observed facts? What actions, indicators and other studies should be designed and/or undertaken before an accepted model could be derived. All of these constitute limiting factors, which will be thoroughly discussed further in this report.

2. DATA SETS COMPLETENESS AND POSSIBILITY OF RISK INDICATORS

2.1. Introduction

In order to model the expected risk faced by each country, data on realised risk and on frequency and severity of hazards needed to be found and connected.

Using UNEP/GRID-Geneva access to data sets through the Project of Risk Evaluation Vulnerability Indexing and Early Warning (PREVIEW) and UNDP/ERD contacts, numerous data sets on geophysical and socio-economical parameters were found. To connect this data with expected human losses, the compilation of victims from disasters made by the Centre for Research on Epidemiology of Disasters (CRED) was used.

Since 1988, the CRED is maintaining an Emergency Events Database - EM-DAT. The main objective is to “serve the purposes of humanitarian action at national and international levels. It is an initiative aimed to rationalise decision making for disaster prevention, as well as providing an objective base for vulnerability assessment and priority setting” (OFDA/CRED 2001).

The CRED database is the most complete database publicly available, which contain records of victims, injured persons per type of events with world coverage.

The objective of the GRAVITY feasibility study is to ascertain whether correlation can be found between the number of victims (killed and injured) as recorded in CRED database and socio-economical, geophysical factors. This would enable to approach the concept of human vulnerability through statistical and spatial analysis. However, there are some limitations when using these data sets which will be discussed in this chapter.

2.2. "Realised" risk (from CRED)

The first analysis of the data relevance and completeness concentrates on the data provided by CRED. The collaborators from CRED have achieved a tremendously useful job by compiling all this information. However, some intrinsic structures of the CRED database are causing difficulties for the statistical analysis. Possibilities and limitations of CRED data are described in this part.

Data sources

The data were kindly provided by CRED directly sent by email from CRED collaborator to UNEP/GRID-Geneva. In this study the files used were joined to provide information from 6 January 1900 to end of December 2000. This was the latest and most complete information available at CRED at the beginning of the study on 1st March 2001.

CRED contributors

CRED compile the information from various sources, ranging from UN agencies, non-governmental organisations, insurance companies, research institutes and press agencies.[2]

- UN agencies (FAO, OCHA, UNDP, UNEP, UNICEF, WHO/OMS, WMO, AFRO, WFP,…).
- US agencies (CDC, FEMA, NOAA, OFDA, USGS, …)
- International Decade for Natural Disasters Reduction (IDNDR)
- International Federation of the Red Cross (IFRC)
- LLOYDS casualty week,
- Society of reinsurance (MunichRe, SwissRe),
- Press agencies (AFP, International Herald Tribune, Mode, Reuters)

Data quality

Classification of disaster types

The natural disasters followed by CRED are:

---

[2] See detailed list at http://www.cred.be/emdat/srcelist.htm

Droughts; Earthquakes; Epidemics; Extremes temperatures; Famines; Insect infestations; Floods; Slides; Volcanoes; Wave/surges (tsunamis); Wild fires and Wind storms (including cyclones).

CRED identifies the primary cause of the event. In some cases, a disaster can cause another one (e.g., a cyclone often causes floods and/or landslides). In these cases the primary cause is identified (e.g. cyclone) and then the secondary or other consequences are described in the “comments” field. This is producing a statistical “skew” as the primary cause is not necessarily the one leading to human lives losses. Secondary effects are often the cause of casualties, e.g. earthquakes or volcanic eruption leading to floods, …

Only major disasters are considered in CRED database. Some report such as *LA RED “Sistema de Inventario de Desastres*” suggests that less severe but more frequent events may lead to a higher number of victims as compared with more severe but less frequent events. Once again, CRED’s information may not be the most appropriate but is still the best global database, case studies will be required to see if the restriction to major disasters is too limiting.

Evaluation of damages

The following figure explains how CRED is doing the evaluation

Figure 2. CRED’s Definitions and Evaluation of Losses

Killed: Persons confirmed as dead and persons missing and presumed dead (official figures when available).

Injured: The number of injured is entered when the term "injured" is written in the source. Injured people are always part of the affected population. Any related word like "hospitalized" is considered as injured. If there is no precise number like "hundreds of injured", 200 injured will be entered (although it is probably underestimated). Any other specification will be written in the comments field.

Homeless: They are always part of the affected population. Reporting from the field should give the number of individuals that are homeless; if only the number of families or houses is reported, the figure is multiplied by the average family size for the affected area (x5 for the developing countries, x3 for the industrialised countries, according to UNDP country list). Any other specification will be written in the comments field.

Specific examples:

Number of houses destroyed = 50 x 5 = 250 homeless (although it is probably underestimated), If the value ranging from a minimum to a maximum : take the average. Thousands of homeless = 2000 homeless (although it is probably underestimated)

Affected: People requiring immediate assistance during a period of emergency ; it can also be displaced or evacuated people. Any other specification will be written in the comments field.

Total affected: Sum of injured, homeless, and affected.

sources: http://www.cred.be/emdat/guide.htm

As indicate in the table above, the computation method is very careful and leads to an underestimation of the losses. The positive consequence is that there are no exaggerations and the result is the minimum of what should be expected. The method is not perfect but at least is applied to the all world, which should make it comparable.

Completeness

One of the first task for the statistician analyst was to evaluate the completeness of the CRED's database, in order to ascertain what could be derived from it and which disaster types could be taken into account. The level of details provided varies extensively from one event to the other. The following table shows the number of events in connection with the level of information. The robustness of a model being highly depending on the number of inputs – the fewer the entrees the weaker the model – following this principle only disasters with relevant significance were introduced (see recommendations in parts 2.5 p.31). The Table 2 describes the number of records found in CRED database. Important: the **number of records** including killed person should not be confused with the **number of killed** person!

Table 2. Number and types of events recorded in CRED database (1980-2000)

Type of disaster	Number of records total	Number of records including information on:			
		Killed	Killed and total affected	Killed + Injured + Info	Killed + injured
Drought	748	405	239	35	387
Earthquake	1087	920	691	410	707
Extreme temp	250	229	63	59	200
Flood	2295	2053	1460	466	1821
Insect infest.	75	67	4	51	62
Slide	451	442	275	106	378
Volcano	191	159	131	34	122
Wave/surge	45	42	13	9	24
Wild fire	269	197	97	45	174
Wind storm	2535	2207	1222	629	1777
total	7946	6721	4195	1844	5652

Killed = Number of records including information on killed persons, Injured = Number of records including information on injured persons, Info = information provided on additional factors that could be used.

Access to information and time scale

Due to the significant rise of access to information, the number of events reported in CRED's International Disasters Database present a large discrepancy. While 2227 natural disasters were reported in the 90ies, only 78 were introduced in the database for the period between 1900 and 1909. This difference can be explained by the tremendous increase in information availability, a deeper interest for the subject and eventually partly by a rise in the number of natural disasters. However, the rise in information access is not uniform worldwide. Differences must be expected at least for disasters that can not be detected remotely.

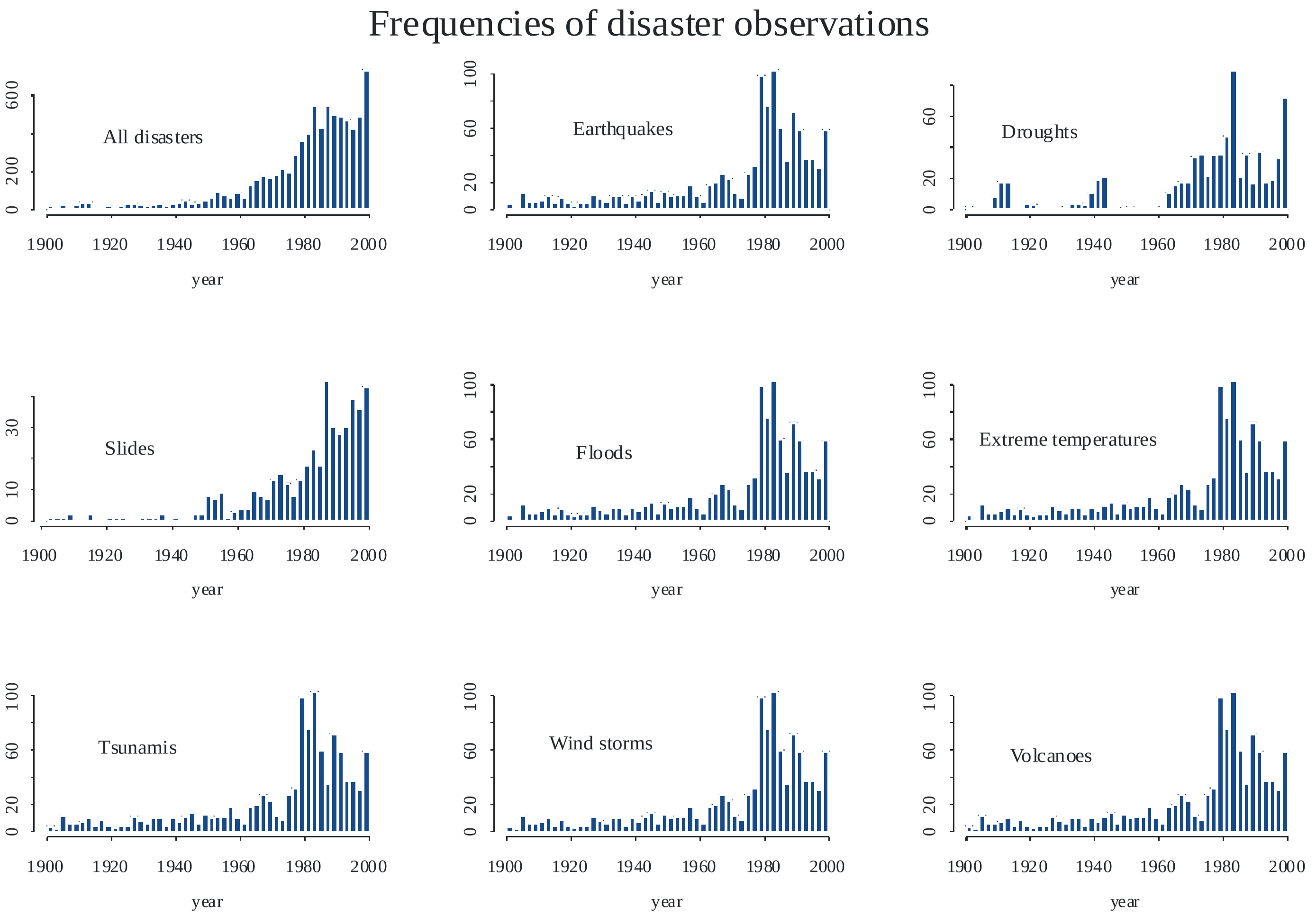
The Figure 3 on page 19 illustrates the variation of events recorded in CRED database. Trends seems to stabilise since the 80th where for some types of event decreases can be observed, which could tend to prove that the level of access to information was sufficient. To overcome these difficulties, only the most recent data were used for calculations. The time span was set to take into account events from 1980 to 2000, period during which the access to information is still not uniform and comprehensive, but at least does not present a tremendous rise of events.

CRED database and georeference

A georeferenced data, is a data that includes an indication of location attached to it. In the best case it consists on two columns with latitudes and longitudes, but it could also be the name of the village/city, province or country. In the case of CRED data, with the exception of 250 earthquakes locations provided with latitude and longitude indications, the resolution of the georeference information is limited to the country. A more precise location would have allowed the connection of the total number of population living in the affected area and the percentage of casualties. This could have been directly used to extrapolate the vulnerability. In absence of such precise locations, the average density of the country was taken into account, this is affecting a great deal the precision achievable by the statistical analysis.

The 250 earthquakes locations were used for a refined analysis (see p.35 in the spatial analysis and p. 52 for the statistical analysis). There is a possibility for tropical cyclones, as names are included in CRED database and could be attached to another file including coordinates of each cyclone for 8 years (see chapter 5 p.63 for recommendations). For the other disasters, average information and approximation had to be used to connect the events with the other data.

Figure 3. Trend over time of information access and variation of number of events



2.3. *Proxy for Vulnerability and Choice of Risk indicators*

Proxy for Vulnerability

The approach of infrastructure quality by Gross Domestic Product (GDP)

Wealthy countries present a much lower rate of losses than poorer countries. However, the method that consists on taking into account the wealth of a country as an indicator, such as Gross Domestic Product (GDP), presents some inconsistencies. Indeed, intuitively it could be easily discerned, that it is not the average amount of money per capita that help in preventing losses, but the amount of investments made in health, education, appropriate planning of infrastructures and development of emergency tools and procedures. Moreover, some of the infrastructures which help rising the GDP, are also factors leading to higher vulnerability, for example, chemical factories, nuclear power plants or extreme forest exploitations, while augmenting the GDP, are producing a higher vulnerability toward earthquakes, floods, landslides,… and create complex disasters. It underlines the need for a more comprehensive approach and highlight the significant role of appropriate development for preventing human losses.

Availability of Global Socio-economical Data sets

As already mentioned in the introduction, the purpose of this study was not to provide a ready to use model, but to test what could be used. The following data sets, which includes comparable information between countries were identified and introduced in the model.

Table 3. Socio-economical variables tested

Variable	Information derived	Years available
Gross Domestic Product	Absolute and augmentation	1960 – 2000
Population	Absolute, density and augmentation	1960 – 2000
Life expectancy	Number of year	1960 – 2000
Literacy rate	Percentage of pop.	1960 – 2000
Country area	Surface in Km^2	217 countries
Human Development Index	Index per country	1998
Urban population	Absolute, %age and augmentation	1960 – 2000
Corruption	Index per country	2000

Except for the information on population, which was provided by pixel of 30’’ resolution (~ 5 x 5 km at the equator), the other information were only available per country. This requested the aggregation of geophysical parameters per country.

The choice of Risk Indicators

The sizes of countries

In accounting the casualties by country, some spatial inconsistencies are produced. The significant discrepancy of country sizes and population call for a neutral spatial reference. How to compare large and small countries high and low populated countries? What is the most relevant feature? The percentage of the population affected (or killed) or the raw number of victims. This question may look trivial (at least at first glance) but has in fact some drastic consequences:

1) If the total of persons affected is chosen:

The advantage is that all human are treated equally, (e.g. 1 Chinese = 1 person from Honduras). However, the real impact of small countries may be underestimated, as the sum of victims of several small countries constituting an equal amount of population may be equal. One thousand deaths spread over several small countries may not be taken into account as the number is divided by the number of countries. But the same amount of victims would have been taken into account in a specific large country (e.g. China or India).

2) If the percentage of persons affected by country is chosen, when comparisons are computed, 1 person from Honduras "weight" more than 160 Chinese!

To overcome these difficulties, a method was tested briefly at GRID-Geneva last year, it consists on computing a density of victims by dividing the number of victims by the area exposed of the countries. This is the most neutral indicators for comparing countries. The vulnerability is then derived according to the number of persons living in exposed areas in association with their intrinsic socio-economical parameters and then multiplied by the frequency and strength of the hazard. The risk for a country is equal to the sum of the risk in all the areas that could be affected. This approach has numerous advantages but need to be explored in a more thoughtful way. This task is time consuming and may be proposed for a future study.

At a country level, a straight solution for this spatial consideration cannot be found. Only a subjective decision can decide what weight should be associated with absolute number of victims, percentage of population affected or with the density of victims. In this study, all the cases will be presented.

As already discussed in the introduction, comparison of risk at country level, arise certain difficulties. When ranking them, most populated countries will appear in the first rows, whereas several countries including together an equivalent risk of losses may not appear in first lines.

Several indicators could reflect the risk:

Table 4. Risk indicators and their relative advantages/inconveniences

Indicators	**Advantages**	**Inconveniences**
Number of persons killed	Show the expected human losses. Every inhabitant is equal.	Advantage populated countries, without taking into account an equivalent population split in several countries
Percent of population killed	Allows comparison between countries	1 Honduran = 160 Chinese!
Number of persons killed per 1000 km^2	Allows comparison between countries	Large inhabited countries are disadvantaged.
Number of persons killed per event	Shows expected human losses per event.	Does not take into account differences in severity. Minimise the impacts in events on several countries.
Percent of population killed per event	Approach the concept of vulnerability	Does not take into account differences in severity. Minimise the impacts in populated countries.
Percent of population in the affected areas. Killed / affected area [km^2]	Approach the concept of vulnerability normalised by the number of affectable population.	Does not take into account difference of severity.
Percent of population “affectable” killed. Killed / number of person affectable.	The ratio is normalised. It takes into consideration both areas and population.	The question of severity cannot be taken into account without falling straight to Vulnerability.

Vulnerability can be normalised and thus comparable between countries. However, a population is vulnerable even if no hazard is expected. It does not make real sense of acting to minimise vulnerability if there are no disasters in this area.

The comparison between countries is posing serious problem, normalisation is always removing some significant information.

The Figure 4 illustrates the problem with risk indicators. Here are four countries, A, B, C and D. Country A is large and includes a population of 14. Countries B, C and D are smaller and include a population of 10 all together. The exposed area and population between B, C and D equal the exposed area and exposed population in country A. However, if a population of 3 is killed during a disaster in both exposed areas, the statistics will be drastically different (see Table 5).

Figure 4. The Choice of Risk Indicators

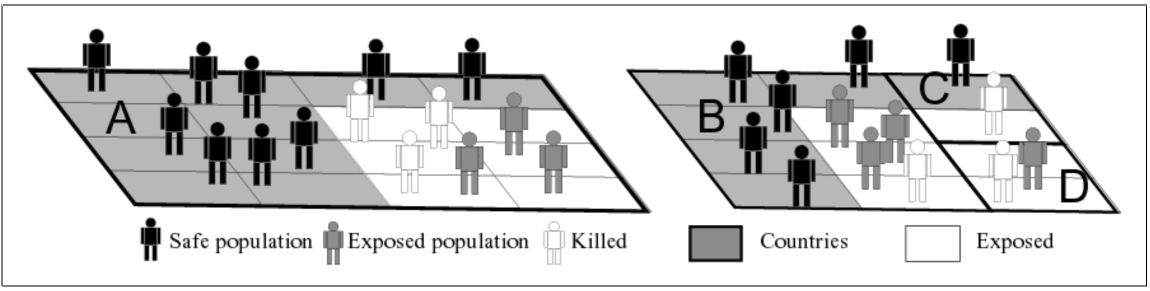


Table 5. Hypothetical results of risk indicators

Countries	Area	Pop.	Killed	% pop. killed	% of affectable pop. killed	Killed per area	Killed/area affectable
A	16	15	3	20 %	50 %	0.1875	0.5
B	8	8	1	25 %	33.3 %	0.375	0.33
C	2	2	1	50 %	100 %	0.5	1
D	2	2	1	50 %	50 %	0.5	0.5

This fictive situation illustrates the complexity of the choice for an indicator of risk. If the number of killed persons is chosen, for a similar area affected, country B, C and D are disadvantaged in terms of consideration. If the percentage of population is chosen, the less populated countries (C & D) are advantaged… a choice has to be made and it will be subjective. No scientific methods can verify what would be the best indicators. A composite indicator including several possibilities may be a solution to take into account two or three parameters such as the number of killed, the number of killed per km^2, …

Some composite ranking were attempted for each hazard types (see in Appendix I).

2.4. Results

CRED outlook

The following figures help to get a better idea of what disasters are causing the most severe losses in terms of killed and where. The analysis is based on 1980 – 2000 period, during which the access of information is considered as being sufficient. It appears that 96.5% of deaths recorded in CRED databases are caused by drought, wind storms, floods, earthquakes and volcanic eruptions; 85% of deaths occur in Africa and Asia-Pacific regions:

Table 6. Deaths per disaster types and per regions

Disaster types	Deaths	% of total
Drought	563'701	46.54 %
Wind storm	251'384	20.76 %
Flood	170'010	14.04 %
Earthquake	158'551	13.09 %
Volcano	25'050	2.07 %
Extreme temp	19'249	1.59 %
Slide	18'200	1.50 %
Wave/surge	3'968	0.32 %
Wild fire	1'046	0.06 %
Insect infestation	0	0.00 %
Total	**1'211'159**	**100%**

GEO regions	Deaths	% of total
Africa	581'391	48.00 %
Asia + Pacific	447'906	36.98 %
Latin America + Caribbean	105'599	8.72 %
Europe	63'274	5.22 %
North America	8'870	0.73 %
West Asia	4'047	0.33 %
Not referenced	72	0.01 %
Total	**1'211'159**	**100%**

Figure 5. Percentage of killed by type of event and by GEO regions

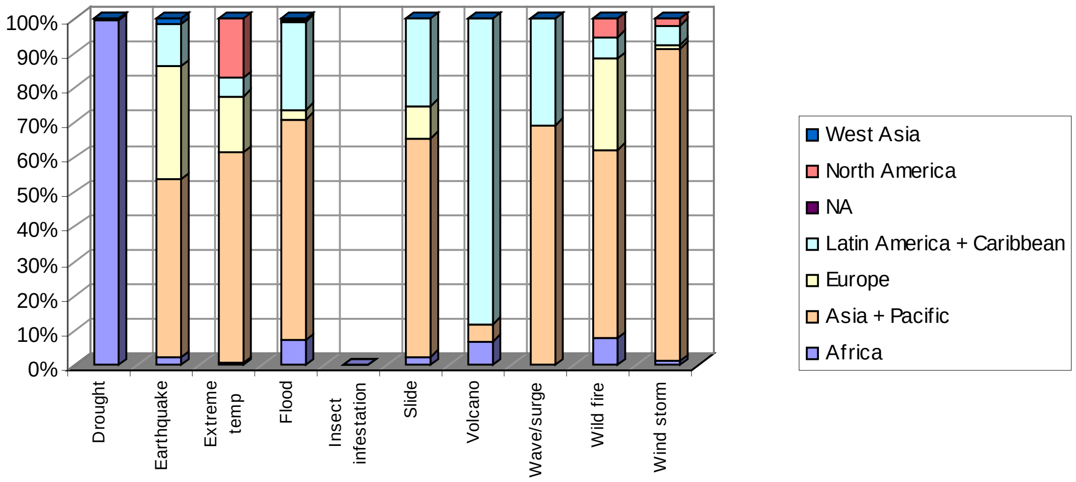


Figure 6. Total killed by GEO regions for 8 disasters type

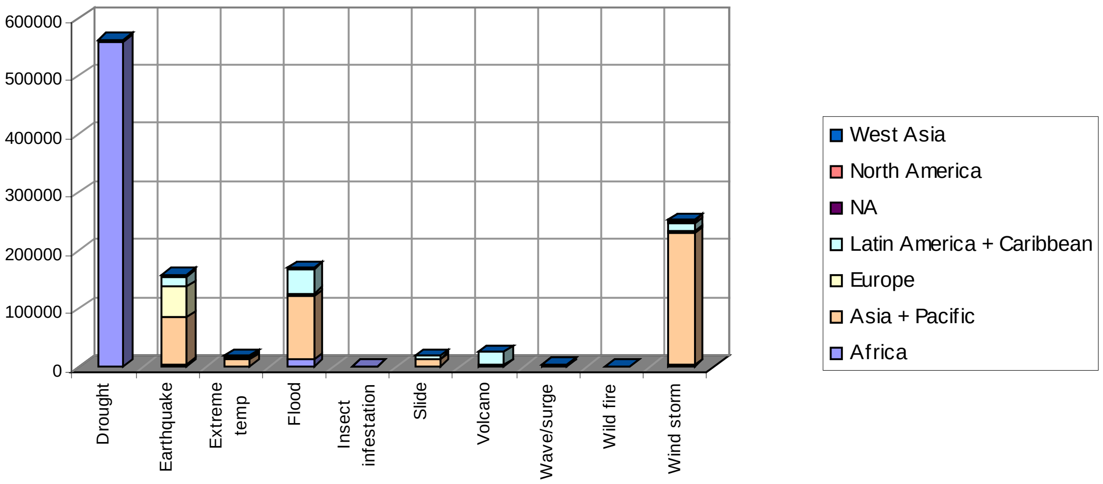


Figure 7. Percentage of killed by disaster type for 6 GEO regions

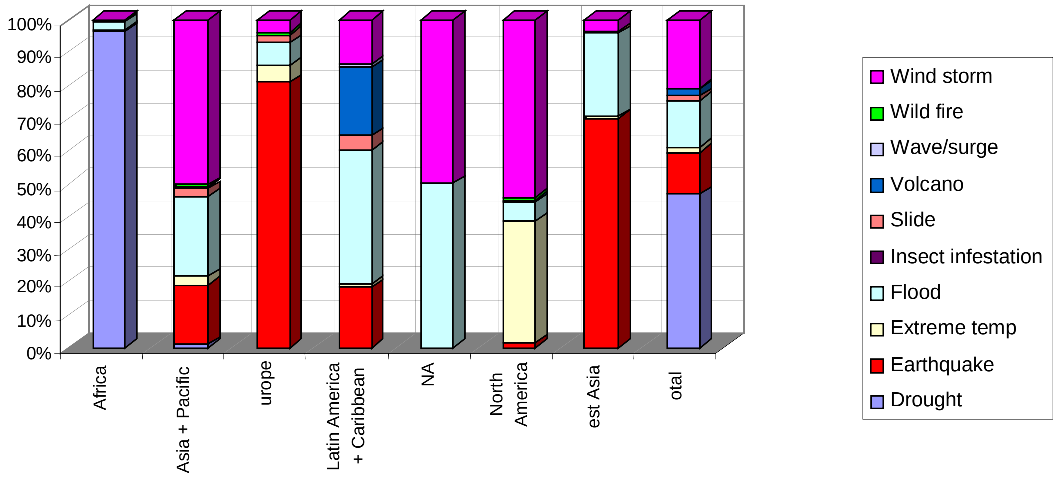
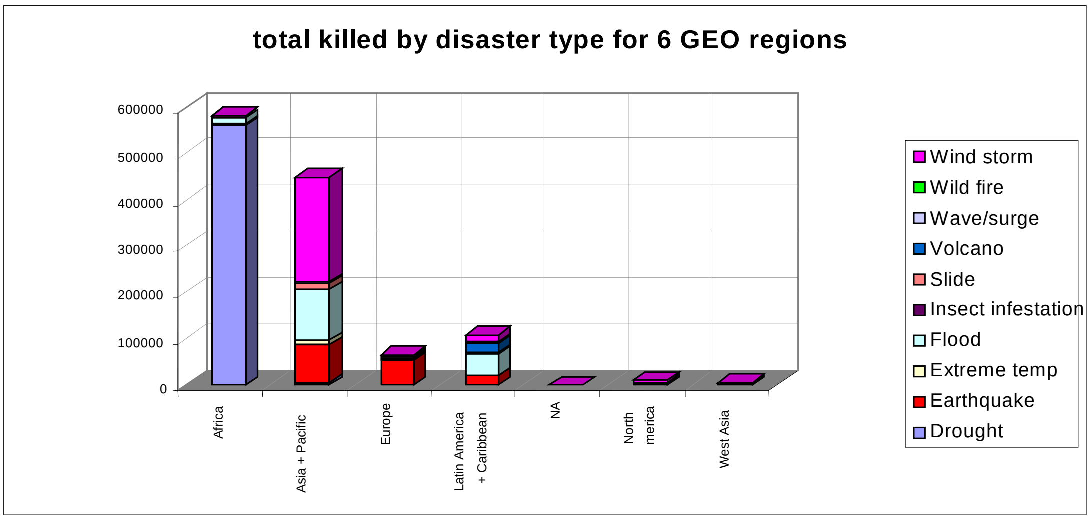


Figure 8. Total killed by disaster type for 6 GEO regions



Variation in scale of impacts

The number of deaths from droughts is incredibly high, most probably because of food shortage. This causes a problem as food shortage is not only related to food production, but also connected with wars and political problems. This is not a natural disaster but a complex disaster. Due to the number of victims it should be introduced into the GRAVITY index, however a lack of data was identified to deal with the complexity of the phenomenon. For the present study, it was not possible to include drought into the proto GRAVITY-model, but recommendations were made to see how such information could be included in a further research (see recommendations at the end of this chapter and in chapter 5 for more details).

In order to simplify an already complex analysis, it was decided to exclude the incorporation of forest fires due to the small amount of victims from this type of disaster. No matter how precise the connection could be made, it would not have much weight on the final risk evaluation. The question may be different if financial cost need to be evaluated, however as already discussed, this was not the focus of this study.

Insect invasion was already not taken into account, following instruction from UNDP/ERD.

The relevance of spatial analysis

Looking at the graphics, the variation of types of events striking the different continents highlights the importance of the spatial analysis. The normalisation of spatial factors in order to compare different situations was of great complexity. If differences in time scale and differences in strength can be approached through modelisation, statistical analysis cannot take into account the geographical differences as well as reflecting the difference of sizes of affected areas.

Tables

see Table 7 p.26

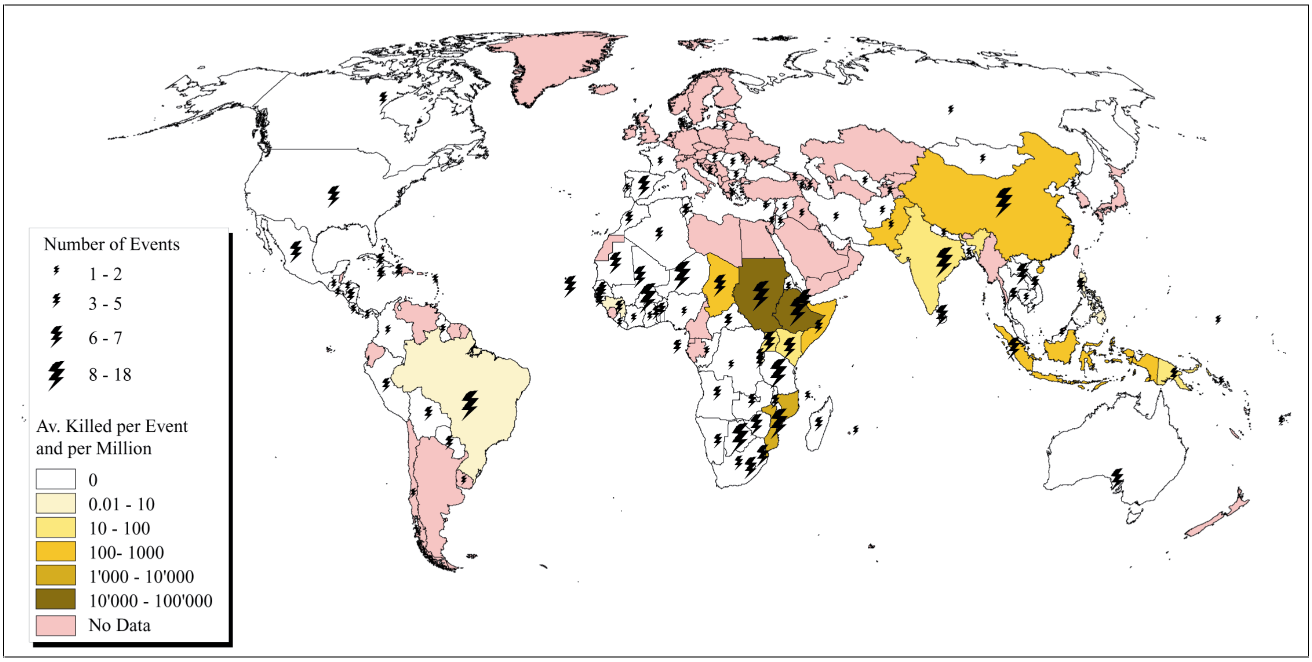
Maps

The following maps delineate the number of events in relation with the average number of killed persons per event and per million of inhabitants. The map underneath provides the number of killed persons between 1980 and 2000 in relation with the density of killed persons (number of killed per thousands Km^2). This information is provided for Floods, Wind storms, Earthquakes and Droughts. Through these maps the geographical patterns and distribution are highlighted, it allows the identification of hotspots for the four main disasters (together they are responsible of 95% of the victims recorded by CRED).

Table 7. Realised risk

Type of disaster	Risk indicators									
	Nb of events	**Killed by event**	**Total killed**	**Average killed per year**	**Killed per 100 million**	**Total affected**	**Average Affected per year**	**Killed/1000 km^2**	**killed/ year/ 1000 km^2**	**Nb countries affected**
Droughts	411	503	563’701	26’843	235	1’044’411’440	49’733’878	4.79	0.23	114
Earthquakes	621	150	158’575	7’551	10	50’623’518	2’410’643	3.8	0.18	88
Wind storms	1616	49	251’440	11’973	72.7	403’648’423	19’221’353	14.24	0.68	162
Floods	1654	61	170’914	8’139	40	2’028’269’338	96’584’254	3.31	0.16	160
Eruptions	107	298	25’050	1’193	189	2’868’015	136’572	15.15	0.72	21
Landslides	315	40	18’200	867	14.4	5’164’626	245’934	---	---	68
Tsunamis	17	50	3’968	189	53	47’269	2’251	0.79	0.04	11
Extreme temperatures	203	43	19’252	917	6	56’670’106	2’698’576	---	---	59

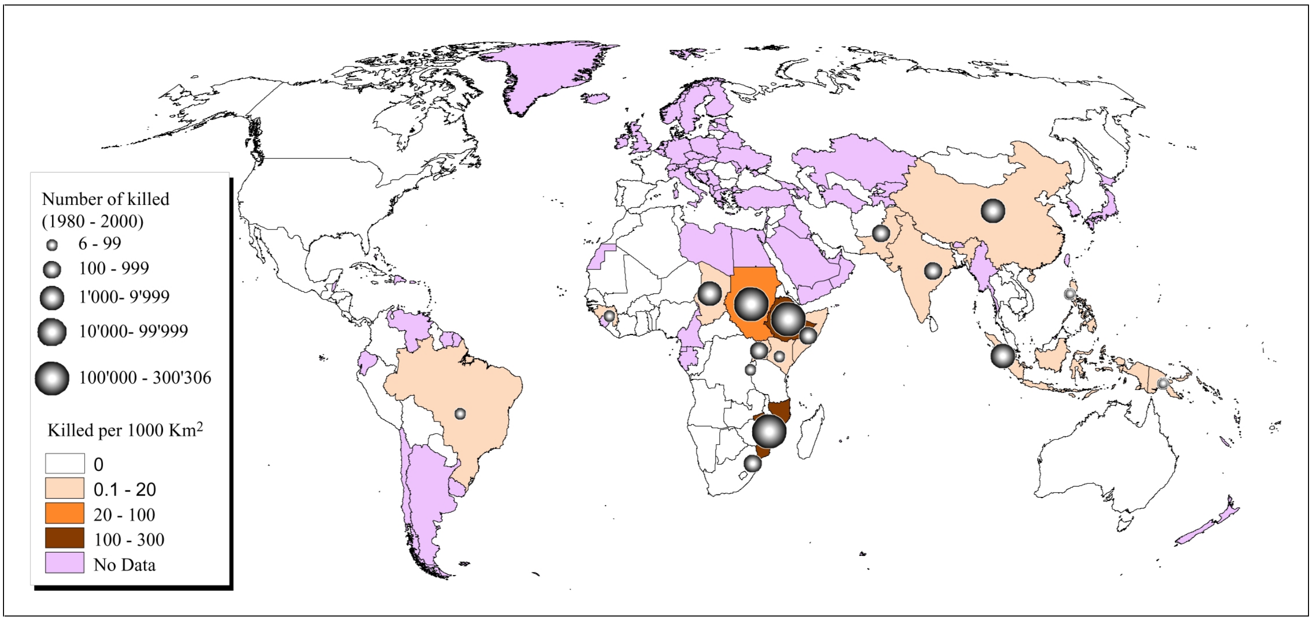
Figure 9. Drought events as recorded in CRED (1980-2000)



Data sources: EM-DAT: The OFDA/CRED International Disaster Database
www.cred.be/emdat – Université Catholique de Louvain - Brussels - Belgium.

Data analysis and cartography: UNEP/GRID-Geneva, June 2001

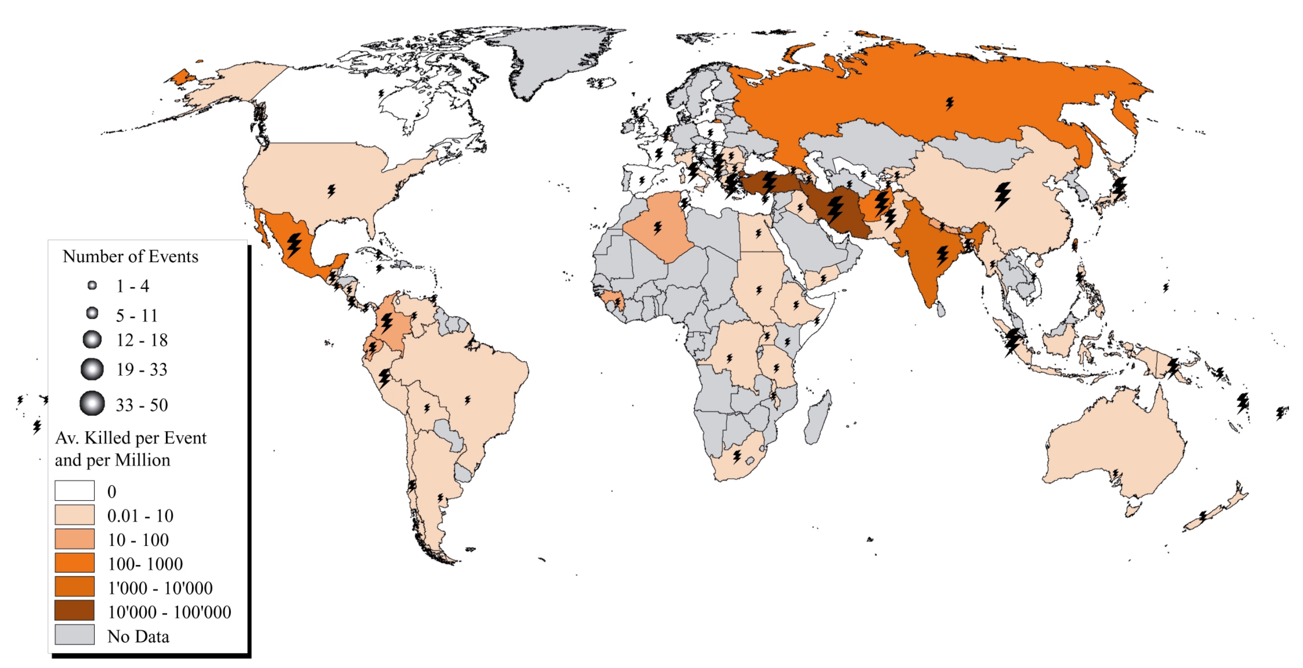
Figure 10. Victims from droughts as recorded in CRED (1980 – 2000)



Data sources: EM-DAT: The OFDA/CRED International Disaster Database
www.cred.be/emdat – Université Catholique de Louvain - Brussels - Belgium.

Data analysis and cartography: UNEP/GRID-Geneva, June 2001

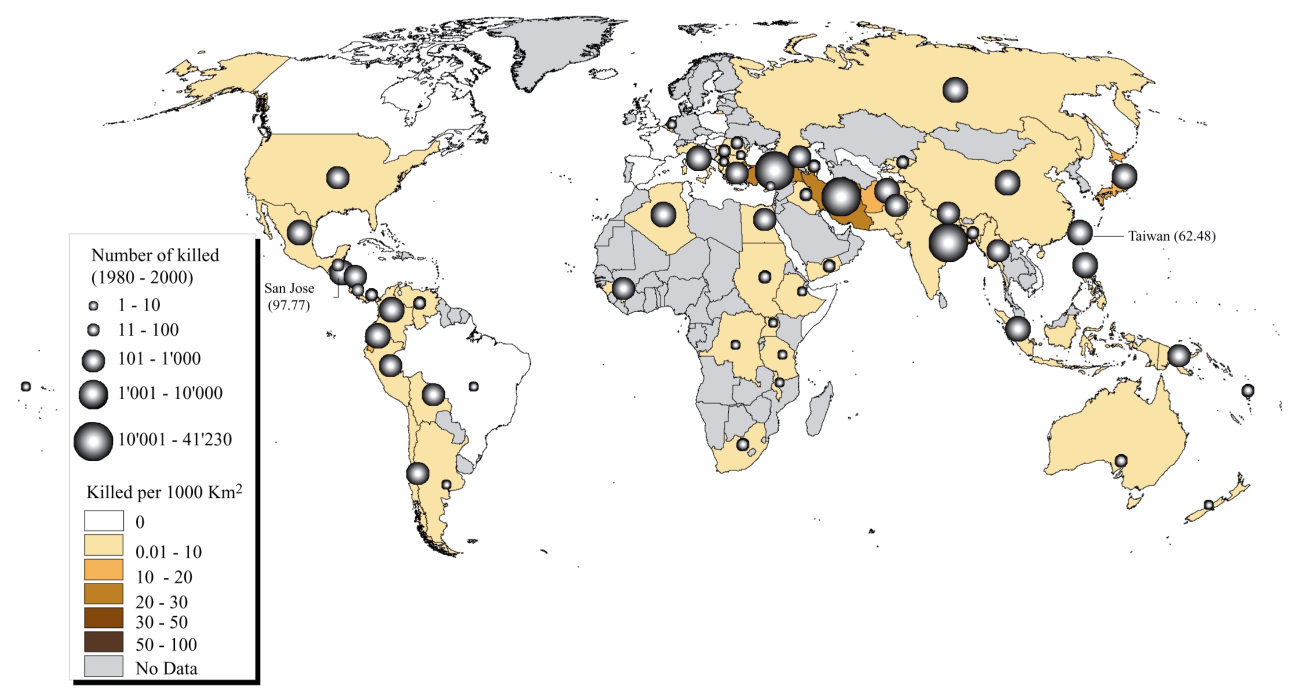
Figure 11. Earthquakes events as recorded in CRED (1980 – 2000)



Data sources: EM-DAT: The OFDA/CRED International Disaster Database
www.cred.be/emdat – Université Catholique de Louvain - Brussels - Belgium.

Data analysis and cartography: UNEP/GRID-Geneva, June 2001

Figure 12. Victims from earthquakes as recorded in CRED (1980 – 2000)



Data sources: EM-DAT: The OFDA/CRED International Disaster Database
www.cred.be/emdat – Université Catholique de Louvain - Brussels - Belgium.

Data analysis and cartography: UNEP/GRID-Geneva, June 2001

Figure 13. Floods events as recorded in CRED (1980 – 2000)

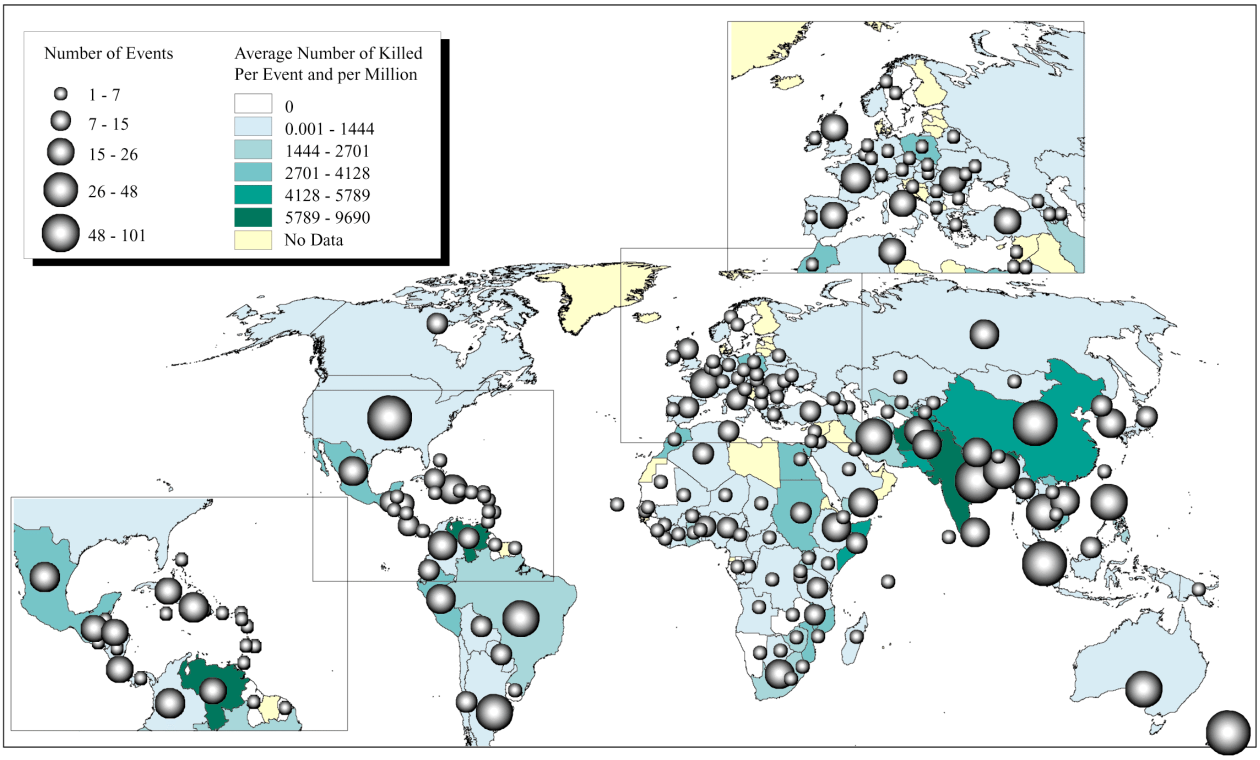
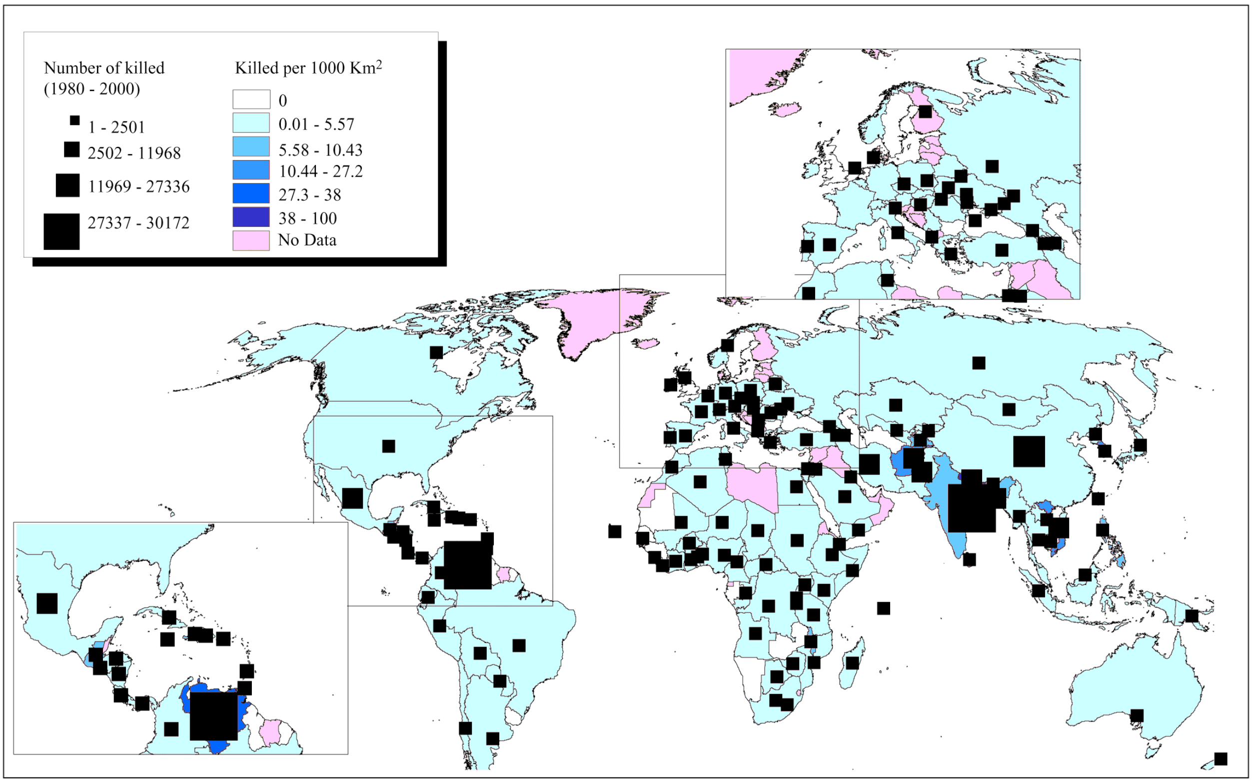


Figure 14. Victims from floods as recorded in CRED (1980 – 2000)



Data sources: EM-DAT: The OFDA/CRED International Disaster Database
www.cred.be/emdat – Université Catholique de Louvain - Brussels - Belgium.

Data analysis and cartography: UNEP/GRID-Geneva, June 2001

Figure 15. Wind storms events as recorded in CRED (1980 – 2000)

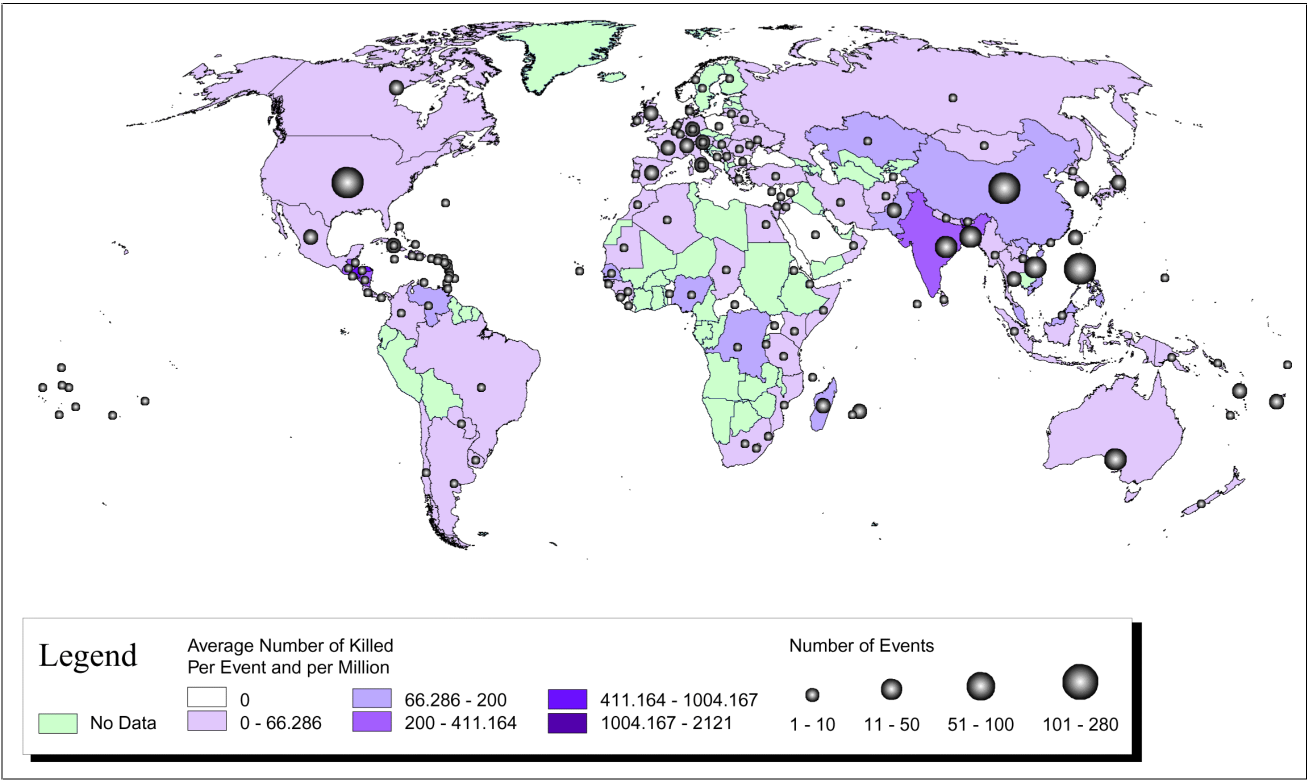
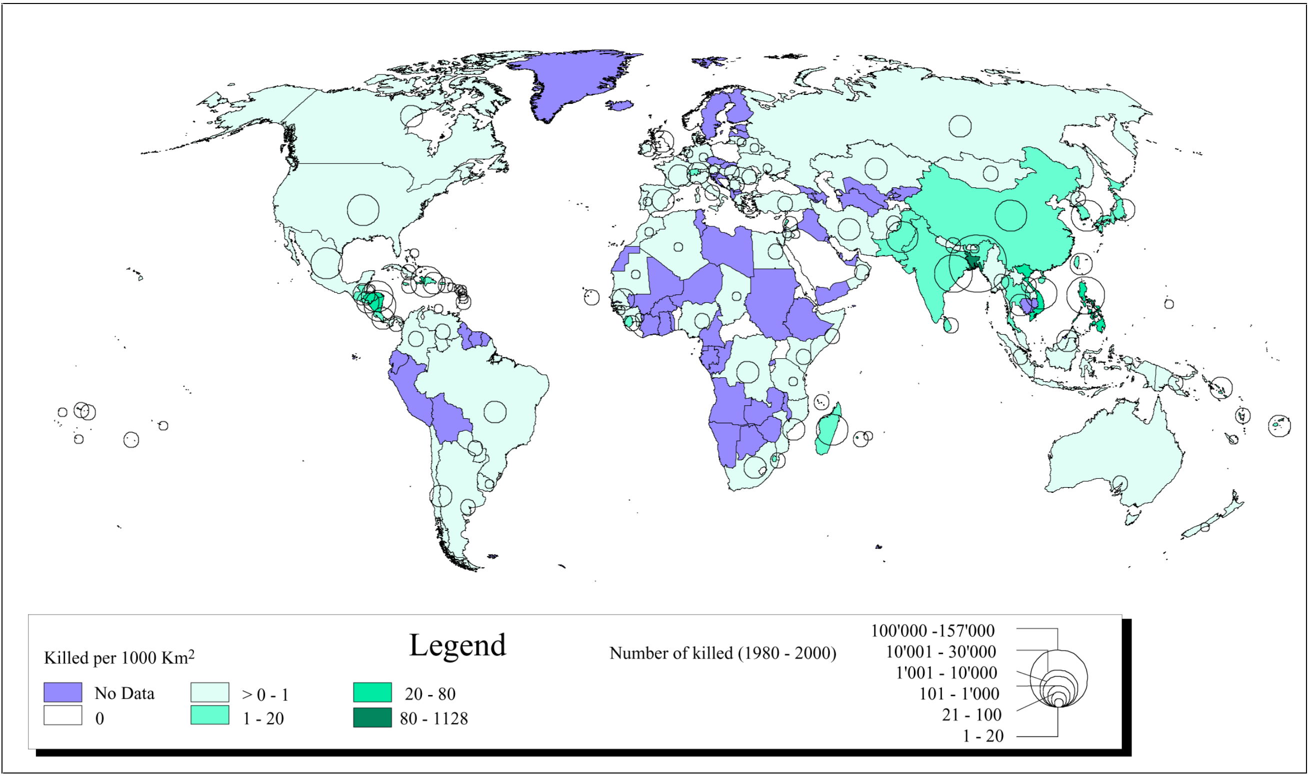


Figure 16. Victims from wind storms as recorded in CRED (1980 – 2000)



Data sources: EM-DAT: The OFDA/CRED International Disaster Database
www.cred.be/emdat – Université Catholique de Louvain - Brussels - Belgium.

Data analysis and cartography: UNEP/GRID-Geneva, June 2001

2.5. *Conclusions and recommendations*

Choice of events to be taken into account

To select what type of hazards should be incorporated into the model, numerous criteria were applied. Some of the recommendations were provided by UNDP/ERD, other were based on availability of the data, significance of the impact in a global way, time constraint for developing the model. Finally 4 disaster types were selected, namely floods, earthquakes, volcanoes and windstorms. They represent 50% of the victims from natural disasters. The drought (with 46.5 % of impact) was left a part for the moment due to the complexity of the phenomenon, however it should be the next one to be incorporated in the model.

Table 8. Disaster selection criteria

Type of Hazards	UNDP/ERD's request if possible	CRED possible ? >5% >20 k / year ?	Geophysical Data available ? Possible ?	incorporated in this study ? Yes / No
Drought	Yes	Yes	Complex	No
Earthquakes	Yes	Yes	Yes	Yes
Epidemics	No	---	---	No
Extrem. temperature	Yes	Not significant	---	No
Flood	Yes	Yes	Yes	Yes
Forest fires	Yes	Not significant	---	No
Landslides	Yes	Not significant	No	No
Locust	No	---	---	No
Volcanoes	Yes	Yes	Yes	Yes
Wave/surge	Yes	No (with exceptions)	lack of detailed data	No
Wind storms	Yes	Yes	Yes	Yes

Instructions from UNDP/ERD for the choice of hazards

UNDP/ERD has already removed some of the hazards that should not be taken into account:

- Locust (insect infestation)
- Epidemics

Restriction from CRED data

Forest and other wild fires

Forest fires were eliminated, as they were never representing a threat for human life in a global way and even not in a national way. Forest fires represent 0.06 % of the victims. There was not a country where this hazard was causing more than 5% and more than 25 victims per year. No matter how accurate and precise it could be modelled, it would not influence the general model in a significant way. It was then decided to simplify the model, which was complex enough with the other inputs.

Landslides

Globally, landslides are causing 1.5 % of the victims. However, it requests very detailed data sets that are not available. Some recommendations are provided in chapter 5, but this could not be taken into account for the present study.

The following table delineates that except these five countries, slides are not representing a major component of risk for the population, at least once aggregated at national level. In chapter 5, some solutions are discussed to see how this could be taken into account but not in a global way which would be too time consuming.

Table 9. Countries where landslides are causing > 20 victims / year and > 5%

Countries	Disasters	Killed	%age of all natural causes
China	Slides	2884	6.1
Indonesia	Slides	1360	13.9
Brazil	Slides	659	20.4
Ecuador	Slides	714	10.2
Peru	Slides	1286	33.0

Extreme temperature

This hazard is causing serious threats to the six following countries, especially Greece and USA. Data sets are available

Table 10. Countries where Extreme temperature are causing > 20 victims / year and > 5%

Countries	Disasters	Killed	%age of all natural causes
China	Extreme temperature	3037	6.4
India	Extreme temperature	6457	8.4
Pakistan	Extreme temperature	708	9.5
Greece	Extreme temperature	1084	67.9
Mexico	Extreme temperature	844	5.9
USA	Extreme temperature	3280	37.6

Wave/surge (tsunamis)

Although Tsunamis are definitely not a global cause of death worldwide (only 0.32 %), it is causing a serious problem in the two following countries. Data sets were downloaded, however some data were missing to extrapolate the exposed area (see ch. 3 and recommendations in ch. 5 for details).

Table 11. Countries where Wave/surge are causing > 20 victims / year and > 5%

Countries	Disasters	Killed	%age of all natural causes
Papua New Guinea	Wave/surge	2182	67.8
Ecuador	Wave/surge	1000	14.3

Restriction from geophysical hazards

Drought

Drought is not the real cause of the victims, the real issue is food shortage. This is not a natural disaster but a mixture of extreme temperature, low precipitation and tensed political situations, which make it a complex disaster. Given the time constraint it would have been impossible to model it in an appropriate way. Moreover, some significant researches have already been carried out by FAO among others. It was decided not to develop this hazard, but to recommend collaboration with an institution having already developed such model. Some possibilities were already identified (see the recommendations in ch. 5).

Landslides and tsunamis

These two disasters were already highlighted as not significant in a global way. Moreover, the requested data were lacking (see ch. 3 and 5 for precision and recommendations).

3. HAZARDS

3.1. *Objectives*

The aim of the spatial analysis was to find data and transform them into information on hazard (i.e. specific locations, severity and probability of occurrence) that could be associated with parameters showing vulnerability. The other aspect was to extract the population "affectable" by hazards.

The spatial analysis includes several tasks:

1) Identify the layers with appropriate spatial resolution for a global survey.
2) Prepare the data for integration into a Geographical Information System (GIS)
3) Find relevant information on modelling the different type of disasters.
4) Extracting the population on exposed area.
5) Evaluate the frequency and the severity of the hazards from previous realised risk.
6) Finally, all this information had to be summarised by countries.

Some choices were made to develop certain methodologies instead of others which were more time consuming. Recommendations at the end of this part and in chapter 5, are providing some insights of what could be achieved in a longer stretch of time.

Global data were found for floods, windstorms (including cyclones), earthquakes, volcanoes and tsunamis. However, data on bathymetry were not of appropriate precision for extrapolating areas affected by tsunamis.

For landslides, numerous data at regional to local scale were found, but unfortunately, no global data seems to have been developed.

The complexity of droughts would have taken more than the three months to approach, although some data sets and methods were found. However, it was also suspected that Food and Agriculture Organisation (FAO) and probably other institutions had already done some significant approaches for modelling drought.

In some cases it was possible to derive the severity of the hazards, for instance, in the case of earthquakes: Peak Ground Accelerations (PGA), magnitude and intensity. On the other hands, it was most of the time difficult, if not impossible, to provide accurate probabilities of occurrence. This is mainly due to the short length of time available globally in the records, while compared with geological or climatic scale, which is measured in thousands or years.

3.2. *Data sources*

The data sources are listed in Table 12, page 34.

Table 12. Data sources on hazards

Type of disaster	From available data					
	Data source	**URL**	**Resolution/scale**	**Spatial unit**	**Intensity**	**Probability**
Droughts						
Earthquakes	Global Seismic Hazard Assessment Program	http://www.seismo.ethz.ch/GSHAP/	1/10 degree	Grid cell	Peak Ground Acceleration	PGA hazard
Wind storms	Carbon Dioxide InformationAnalysis Center	http://cdiac.esd.ornl.gov/	5 degrees	Grid cell	Wind speed	Annual probability
Floods	Dartmouth Flood Observatory	http://www.dartmouth.edu/artsci/geog/floods/index.html	-Large flood events watersheds: based on satellite image and DFO news sources. -Flood limits observed by satellite: 250m-1km.	Polygons	Large flood events: 1-3 scale, based on damages, fatalities, recurrence interval and flood extend.	Frequency 1980 – 2000
	CRED	http://www.cred.be/emdat/	Records on 21 years	Countries		Frequency 1980 – 2000
Eruptions	National GeophysicalData Center	http://www.ngdc.noaa.gov/seg/hazard/volcano.shtml		Points (buffered to produce polygons)	Volcanic Explosivity Index	Frequency VEI 2-3: 1950 – 2000 VEI 4-7: 1500 – 2000
Landslides						
Tsunamis	National Geophysical Data Center	http://www.ngdc.noaa.gov/seg/hazard/tsu.shtml		Points	Maximum runup	Frequency 1900 – 2000
Extreme temperatures						

3.3. Methods

The approach adopted was to define, for each disaster type and for each country, broads hazard zones, in order to evaluate which populations are liable to be affected by specific events.

Earthquakes

The first attempt was to computerise the global seismicity database of the CNSS (Council of the National Seismic System) with ArcView to produce a grid of hazard areas. However, the complexity of the seismic phenomenon led quickly to think that this approach was not appropriate. Indeed, earthquake effects on the earth surface are direct consequences of magnitude, depth of the hypocentre, distance to epicentre and subsoil effects.

The second option was to use a map of intensity based on the Mercalli scale. But this scale, as it describes the effects of earthquakes on inhabitants and buildings, is closely linked to the vulnerability of a specific population and, therefore, not really appropriate to define seismic hazard in a strictly physical way. Furthermore, it depends on the subjectivity of earthquake reports. However, such a map could be eventually integrates in furthers investigations, for it contains, in a particular manner, information on vulnerability (see recommendations in chapter 5).

Finally, the choice was made to use the global seismic hazard map realised by the GSHAP (Global Seismic Hazard Assessment Program, see Figure 18, p.41). It depicts Peak Ground Acceleration (PGA), which is a short-period ground motion parameter, with a 9o% probability that values shown will not be exceeded in 50 years, which corresponds to a return period of 475 years. In other words, this is the "likely level of short-period ground motion from earthquakes in a fifty-year window".

The site classification is rock/firm soil for Canada and the United States, and rock in all others regions.

PGA is particularly useful, for short-period ground motions affect short-period structures, (e.g. personal houses) which are the most common in the world.

The range of PGA values shown in the map may be grouped in four categories, which represent hazard in term of "low", "moderate", "high" and "very high" (Giardini, Grünthal, Shedlock and Zhang 2000).

The original data is a text file of latitude and longitude values in decimal degrees, and the PGA value corresponding. A grid of one tenth of decimal degree resolution has been produced with ArcInfo from this data file.

Physical exposure to earthquakes (Figure 19, p.42) was calculated for the three zones of highest hazard (i.e. moderate, high and very high) with the next equation:

$$P_{exp} = PGA_i \text{ x } Pop_i$$

Where PGA_i *is the mean PGA of a specific hazard zone in a particular country, and where* Pop_i *is the total population living inside this area. This population was extracted for computation of affectable population.*

A more detailed method, which should achieve improved accuracy – time consuming in reason of resolution differences between layers of information – could be carried out if a refine study is conducted.

The second task was to locate events of CRED's database, in order to define socio-economic parameters for each of them. Geographic coordinates in decimal degrees are available for about two hundred fifty events, and all of them happened in a ten-year window (1990-2000). Then, the purpose was to define approximately a zone around each epicentre, beyond which ground motion is considered as insignificant. Here, two parameters were taken into account: peak ground acceleration and duration of shaking.

Considering the PGA on rock soil, the values are unlikely to exceed 0.1 g (~0.98 m/s^2) for distance over hundred kilometres from the source, and for frequencies less than 8 Hz, which presents the largest interest for most of the structures. On the other hand, according to estimations for specific PGA and frequency ranges, duration of ground motions beyond a distance of hundred kilometres is significantly attenuated (Bolt, Horn, Macdonald and Scott, 1975):

Table 13. Bracketed duration in seconds
(acceleration > 0.05 g = ~ 0,49 m/s2, frequency > 2 Hz)

Distance	**Magnitude**						
(km)	**5.5**	**6.0**	**6.5**	**7.0**	**7.5**	**8.0**	**8.5**
10	8	12	19	26	31	34	35
25	4	9	15	24	28	30	32
50	2	3	10	22	26	28	29
75	1	1	5	10	14	16	17
100	0	0	1	4	5	6	7
125	0	0	1	2	2	3	3
150	0	0	0	1	2	2	3
175	0	0	0	0	1	2	2
200	0	0	0	0	0	1	2

Source : Bolt, Horn, Macdonald and Scott, 1975

Bracketed duration is " the elapsed time (for a particular frequency range) between the first and last acceleration excursions on the record greater than a given amplitude level (for example, 0,05 g)" (Bolt, Horn, Macdonald and Scott, 1975).

According to these figures, a buffer of hundred kilometres has been defined around each event to limit area affected by ground motions.

The possibility to attribute different dimensions to these buffers, according to earthquake magnitudes, has not been pursued because of lack of regional data. Indeed, extent of earthquake effects may vary, depending on several regional aspects such as subsoil characteristics. However, such approach could be eventually integrated in furthers investigations to improve precision, given the available data.

Wind storms

The data used to define wind storm hazard areas are produced by the CDIAC (Carbon Dioxide Information Analysis Center). These are Arc/Info coverages, which delineate annual probabilities of occurrence of tropical cyclones (Figure 18, p.41). The spatial unit is a 5x5 decimal degrees cell. Probabilities are based on tropical cyclones activity of a specific record period, except for several estimated values attributed to areas that may present occasional activity but where no tropical cyclones were observed during the record period.

Table 14. Wind speeds and appellations

Wind speeds	Name of the phenomenon
≥ 17 m/ s	Tropical storms
≥ 33 m/ s	Hurricans
	Typhoons
	Tropical Cyclones
	Severe cyclonic storm
	Depending on location[3]
≥ 65 m/ s	Super Typhoons

Saffir-Simpson tropical cyclones classification is based on the "maximum sustained surface wind". With winds of less than 17 m/s, they are called "tropical depressions". If the wind reaches speeds of at least 17 m/s, they are called "tropical storms". If the wind speed is equal to or greater than 33 m/s, they get one of the following names, depending on their location: "hurricanes", "typhoon", "severe tropical cyclone", "severe cyclonic storm" or "tropical cyclone". At last, if the wind reaches speeds of 65 m/s or more, they are called "super typhoons" (Christopher W. Landsea, NOAA/AOML, 2000).

Physical exposure (P_{exp}) to tropical cyclones of each magnitude was calculated for each country with the following equation:

$$P_{exp} = \Sigma P_i \text{ x } Pop_i$$

Where P_i is the annual probability of occurrence of a specific magnitude event in one spatial unit (5x5 decimal degrees cell), and where Pop_i is the total population living inside this area. Total physical exposure in a particular country is the sum of all physical exposure of this country.

The CDIAC coverage includes occurrence probabilities of events of at least "tropical storm" intensity. Then, in order to define hazard zones and probability of events of less intensity, supplementary data are requested. In this intention, several other CDIAC's climatic data could be used. For instance, a similar global database depicting the annual mean number of cyclones, without regard to cyclone type.

Furthermore, in order to define areas affected by specific events, it could be very interesting to pursue a previously initiated investigation, which purpose is to limit area around cyclone tracks, according to wind speed and pressure data. The results could significantly help to enhance correlation between CRED's events and regional parameters. This is a time consuming task as each cyclone's track need to be build, however the information is available and the methodology has already been tested in a previous GRID-Geneva's research (see recommendations in chapter 5).

---

[3] Hurricans:	North Atlantic Ocean, Northeast Pacific Ocean east of the dateline, or the South Pacific Ocean east of 160E);
Typhoon :	Northwest Pacific Ocean west of the dateline,
Severe tropical cyclone:	Southwest Pacific Ocean west of 160E and Southeast Indian Ocean east of 90E, Severe cyclonic storm: North Indian Ocean,
Tropical cyclone:	Southwest Indian Ocean

Sources: NOAA/AOML, FAQ: Hurricanes, Typhoons, and Tropical Cyclones.

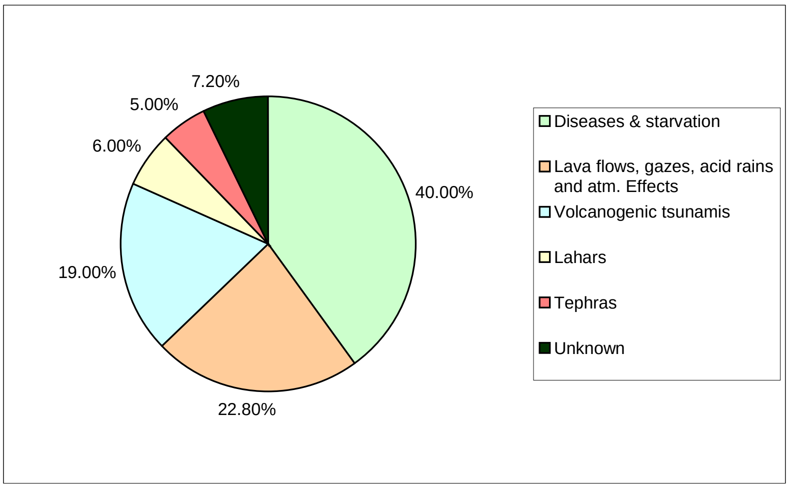
http://weather.about.com/science/weather/gi/dynamic/offsite.htm?site=http%3A%2F%2Fwww.aoml.noaa.gov%2Fhrd%2Ftcfaq%2FtcfaqA.html).

Volcanoes

As risk was defined in term of direct victims, the first task was to determine which volcanic events must be taken into account.

It appears that, for the 1600-1982 period, the major causes of deaths were victims of disease and starvation, pyroclastic flows, volcanogenic tsunamis, lahars, tephras falls and remaining casualties were consequences of lava flows, gases, acid rains, atmospheric effects or seismic action (see Figure 17.).

Figure 17. Deaths causes following a volcanic eruption (based on records 1600-1982)



These statistics are subject to variations, as they are influenced by few big eruptions. Furthermore, considering present-day international communications and first-aid organisations, deaths due to disease and starvation would probably be minimised. However, excepting tsunamis, which are treated as a particular disaster type, pyroclastic flows, lahars, tephra falls and ballistic projectile, remain principal causes of direct deaths (Blong, 1984).

In order to define hazard zones at a global level, the approach used the NGDC database to determine volcanic activity around the world and, broadly, areas that could be affected.

The magnitude unit available in the NGDC eruption database is the Volcanic Explosivity Index (VEI). This is a magnitude measure established by Newhall and Self, integrating quantitative data as well as descriptions of observers. It is a 0-to-8 scale, which describes an increasing explosivity. Each level corresponds, among others, to a particular volume of explosive products, eruptive cloud height and descriptive terms (Simkin and Siebert, 1994).

As the principal causes of direct deaths are linked to explosive events, the first two VEI levels (0 and 1) have been omitted. Then, two groups of magnitudes were defined. The first one corresponds to levels 2 and 3, described as explosive eruptions. The second correspond to levels 4 to 8, described as cataclysmic, paroxysmal or colossal eruptions.

Hazard areas are respectively buffer of 10 and 30 kilometres around the eruptive centre, according to several regional hazard maps.

The records of the last fifty years (1950-2000) have been considered to determine hazard zones and frequencies for the explosive eruptions (VEI 2 and 3). Indeed, histogram of the NGDC database shows that record for explosive eruptions is the most complete for this period. The selection of the events takes into account both the fact that the database completeness is better for the last 150 years and that most of the time intervals between eruptions of levels 4 to 7 are greater than hundred years (Simkin andSiebert, 1994). The records of the last five centuries are to be considered in order to define hazard zones and frequencies. A greater time interval would have clearly underestimated America's hazard (less records).

Physical exposure to volcanism activity was calculated for each country and each group of magnitude with the following equation:

$$P_{exp} = \Sigma F_i \times Pop_i$$

Where F_i *is the annual eruptive frequency of a volcano, based on the last five decades or the last five centuries record, depending on magnitude group, and where* Pop_i *is the total population living in this volcano hazard area.*

This process generates hazard zones in a broad manner and it is clear that areas affected by specific events may vary significantly, depending on regional characteristics. For instance, Lahars are linked to many parameters like pluviometry, seismicity, topography and soils characteristics, among others. Tephra falls are directly influenced by wind dominant direction, and may affect areas hundreds kilometres away from eruption. Ground water access to the magma may produce phreatomagmatic eruption and thus might increase grow significantly the level of explosivity.

Furthermore, as time intervals between eruptions stretch significantly with magnitude (VEI) until tens of thousands of years, long periods without activity are commonly followed by more powerful explosive eruptions. Hence, high volcanic activity determines by historical record does not necessarily imply high volcanic hazard. The fact that some of the most explosive and fatal eruptions of the last two centuries have been from volcanoes without historical activity recorded perfectly highlights this point. Therefore, in attempt to include in records all volcanoes that might erupt in a relatively short-term, modern geological and mineralogical investigations are and will be requested, along with the historical records (Simkin and Siebert, 1994).

However, the general trend of explosive volcanism at convergent plate margins and effusive volcanism at divergent plate margins and hot spots, is well shown by the results, and this was the purpose of this approach.

Floods:

Coverage of large flood events from Dartmouth Flood Observatory were used to broadly define flood hazard zones for the world (Figure 20, p.43). Data is available for height years: 1986-1988, 1993, 1998-2001. Sources are MapInfo files, except for 1998 and 1999. Data for these two years have been downloaded as images and georeferenced with ArcView, then transformed into shapefiles. These zones represent flood-generating watersheds of all large flood events recorded during this eight years. They are based on flood events as recorded by satellite sensors or airborne systems and on various reports.

A database is also available and provides information like location, dates, duration, number of casualties, magnitude, surface flooded. These data could eventually be used on a further development phase (see recommendations chapter 5) to highlight correlation between CRED's events and regional socio-economics aspects.

Global coverage of flooded areas detected by satellites has also been downloaded from Dartmouth Flood Observatory. It delineates water surface changes for the last fifteen years. The resolution depends directly on satellite resolution (AVHRR, Landsat 7, Radarsat Scansar, ERS SAR, MODIS). (Elaine Anderson and G. Robert Brakenridge, 2001).

Annual frequency of flooding was calculated for each affected country and is based on the last two decades records of CRED's database.

Physical exposure to floods (Figure 20, p.43) was calculated for each country with the following equation:

$$P_{exp} = \Sigma F \times Pop_i \,/\, Nwsh$$

Where F is the flooding annual frequency of a country, Pop_i *is the total population living in a watershed generating a particular event, and Nwsh is the number of flood-generating watershed of this country for this eight years record.*

Tsunamis

NGDC Tsunamis database was used to generally locate the affected countries. The last hundred years records (1900-2000) were considered, as it is the most complete. The first attempt to define hazard zones of countries highlighted difficulties to find either local or regional data. As one of the NGDC database variables is the maximum runup (i.e. the maximum height of the water observed above sea level) an area could be delimited using a digital elevation model to get the maximum extent reached by each event on land. However, given the complexity of the phenomenon and diversity of factors generating tsunamis, a hundred-year records is clearly too short a time to get relevant figures of hazard zones. Indeed, tsunamis characteristics and effects on coasts are influenced by the type of sources generating the wave as well as the submarine and coastal topography.

Most tsunamis are triggered by earthquakes vertical movements along subduction zones. These commonly produce two opposite waves, one propagating toward the near shore, generating a local tsunamis, and an other one that may cross the open ocean and generate a remote-source tsunamis up to thousands kilometres from the source. This wave can propagate at a speed of more than 1000 km/h, but, as it is a long-period wave with a maximum height of one meter, it remains unnoticed in open ocean. As soon as it reaches a coast, wavelength is shortened and amplitude of each wave grows progressively producing runup until tens of meters high. It is important to notice here that tsunamis consequences largely depend on regional bathimetry and topography. Indeed, they may rapidly vary along a coastal line.

Two other significant sources of tsunamis are volcanic activity and landslides, the first one having triggered events with terrific consequences on coastline all around.

According to what is mentioned above, most of the last century tsunamis have been generating around the Pacific Basin, some of them propagating toward the opposite coast and affecting large areas of coasline. The specific conditions of the Pacific islands, like Hawaii, is to be mentioned here. Indeed, given their situation, they are particularly exposed to the phenomenon, and should be subject to particular attention.

For furthers studies, considering the complexity of the phenomenon, the best solution to achieve a global tsunamis hazard map, if possible, would probably be to compile regional data produced by specialised national organisations (Pat Lockridge, 2001).

Landslides

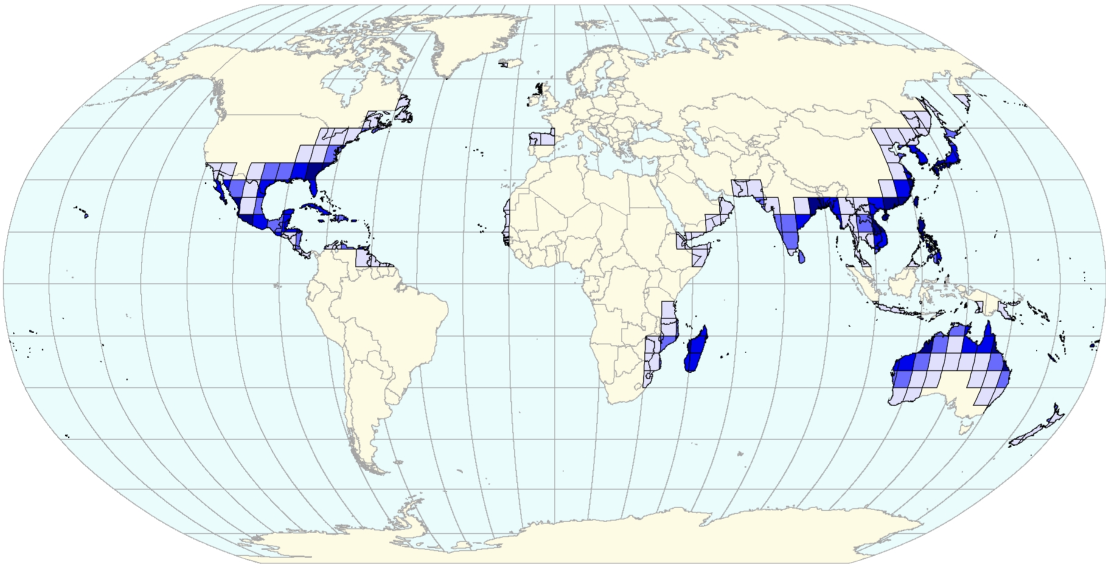
As landslides are generally linked to regional or even local parameters, such as soil features, hydrologic and hydro-geologic characteristics, vegetal cover, type and rate of urbanisation, a general model producing small-scale hazard maps would probably omit a whole range of this regional parameters. Hence, it would generate an output map that would dangerously neglect large-scale areas of specific hazard level.

To achieve a global landslides hazard map, the solution would probably be to compile regional data, although the lack of data for particular region or country would certainly be a problem here. A third solution could be to select several typical cases of landslide disasters, in order to go further into understanding regional particularities, and hence try to extrapolate these first outputs to regions with similar socio-economic and geophysical background.

3.4. *Maps*

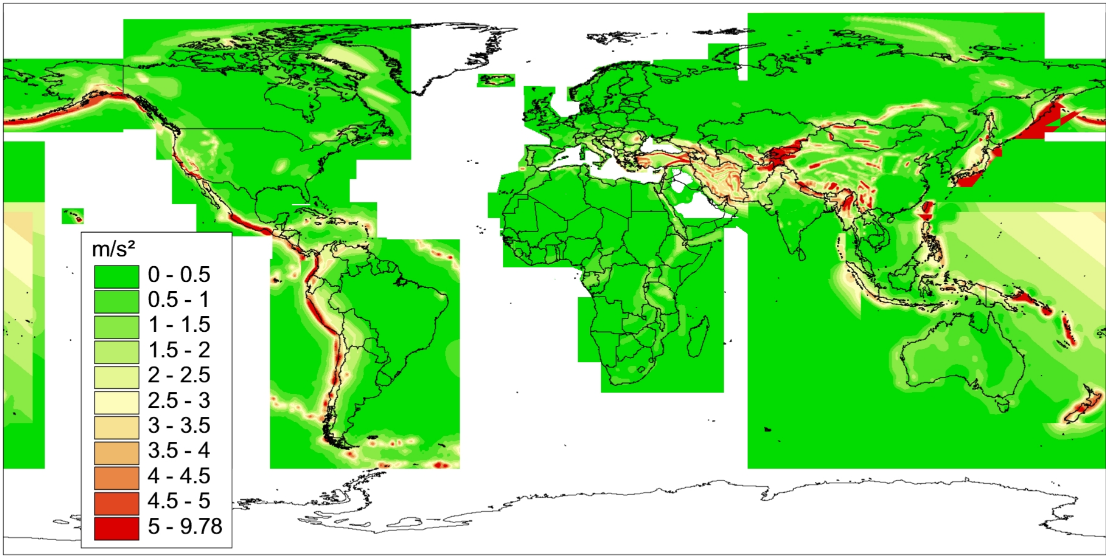
Figure 18. Wind storm and earthquake hazards

Wind storms (tropical storms)



Source: Carbon Dioxide InformationAnalysis Center, http://cdiac.esd.ornl.gov/

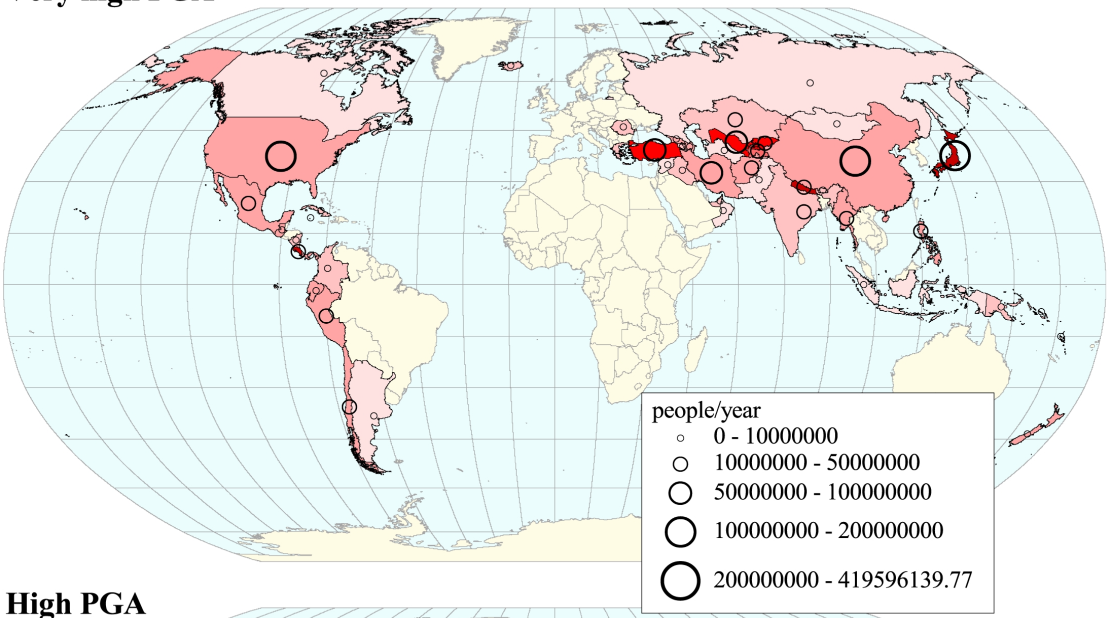
Earthquakes (Peak Ground Acceleration)



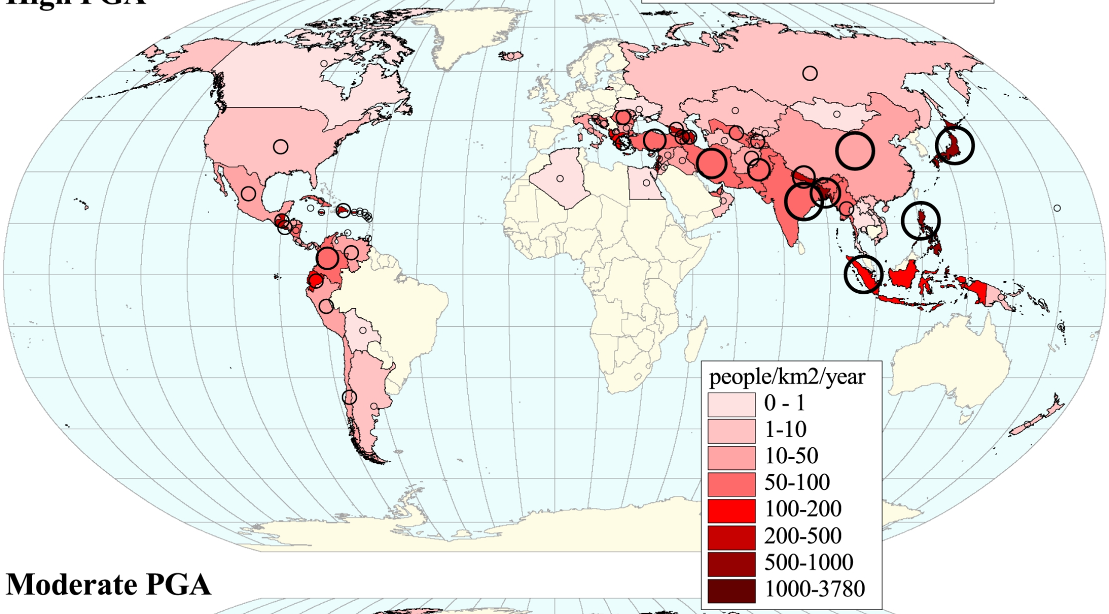
Source : Global Seismic Hazard Assessment Program, http://www.seismo.ethz.ch/GSHAP/

Figure 19. Physical exposure to earthquakes

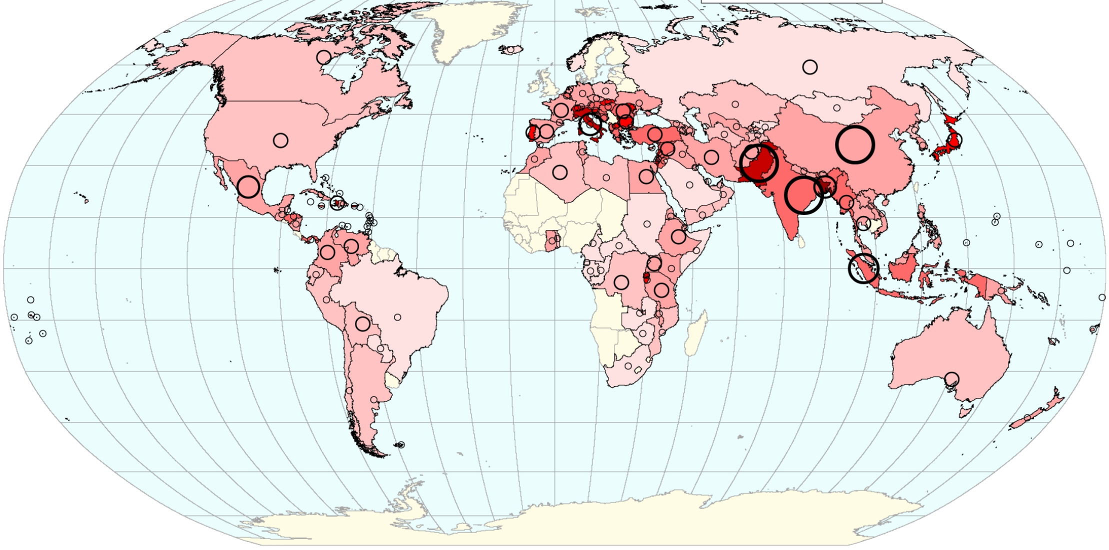
Very high PGA



High PGA



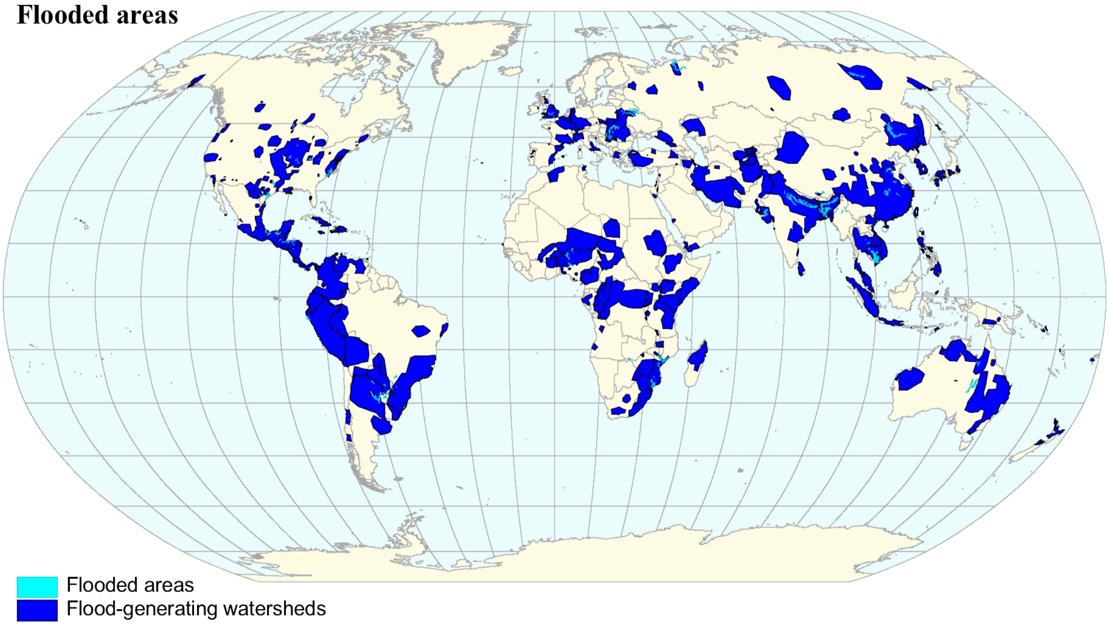
Moderate PGA



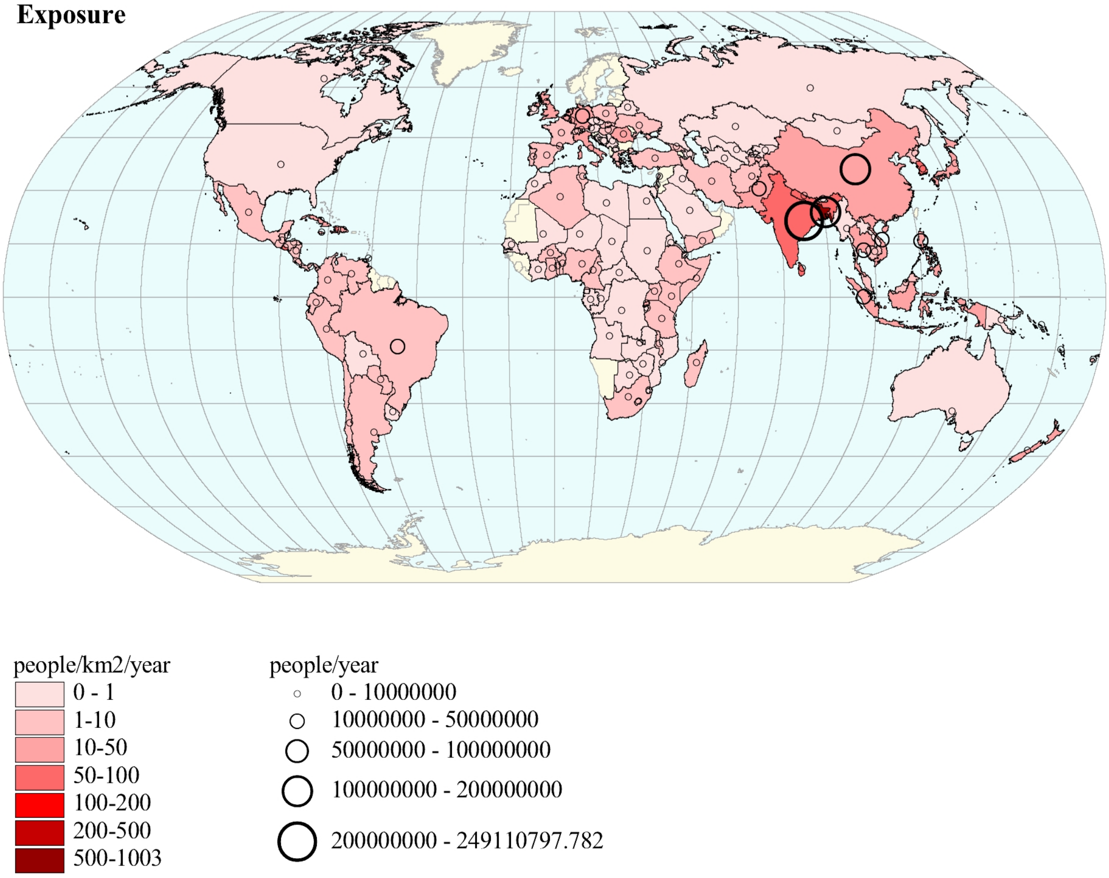
Compiled from : Global Seismic Hazard Assessment Program, http://www.seismo.ethz.ch/GSHAP/

Figure 20. Hazard and exposure to flood

Flooded areas



Exposure



Data sources : Dartmouth Flood Observatory, http://www.dartmouth.edu/artsci/geog/floods/index.html & CRED, http://www.cred.be/emd

4. VULNERABILITY

4.1. *Objectives*

This chapter presents the methodology developed to model vulnerability. Vulnerability, which is defined as "the extent to which a community, structure, service or geographic area is likely to be damaged or disrupted by the impact of a particular hazard" (Tobin & Montz 1997), cannot be directly measured. Vulnerability is estimated by a set of socio-economic variables and compared to actual disaster losses as reported by CRED.

4.2. *Components of the model*

In this section, concepts of hazard and vulnerability (already introduced in chapter 1.) are further defined, from the point of view of the modelisation, in order to ensure an accurate interpretation of the methodology and results.

General model

It is widely accepted that risk is composed of the 4 following aspects[4]:

- Hazard is the occurrence of a physical phenomenon. Hazard is composed of 3 facets: type of hazard, severity (intensity, magnitude, size, …), and probability of occurrence. Example of a hazard : the probability of an earthquake of magnitude higher than 6.5 .
- Exposure measures the quantity of people and physical objects that are subjected to a threat.
- Vulnerability measures how easily the exposed people, physical objects and activities may be affected in the short or long-term.
- Mitigation measures how effectively and efficiently a country can reduce the impact of a disaster through improved preparedness, better mean of intervention, information to population,… While the first three factors measure the country's risk, this factor measures how well the country can counteract the risk.

The definition of risk (and risk index) is crucial: which of the above facets of risk are included in the model and how has to be made clear. In the context of this study, risk is defined as follows:

$$\text{Risk} = \text{Hazard} \ast \text{Vulnerability}$$

In that expression, *Vulnerability* is composed of exposure, vulnerability and mitigation as defined above. Note that mitigation is included in the concept of vulnerability. Therefore, vulnerability is not necessarily a negative thing, but a neutral characteristic. Appropriate mitigation efforts may set vulnerability to a level at which threats are unlike to become disasters.

In the next two sections, main issues about hazard and vulnerability estimation are highlighted.

Vulnerability

The dimensions of vulnerability

Vulnerability is a rather vague notion. However it can be intuitively defined as "what turns a hazard into a disaster". Consequences of disasters are various, and can be considered as different types of vulnerability:

Adapted from B. Tucker, Stanford University, http://pangea.stanford.edu/ tucker/eri/eridesc.html

- *Economic vulnerability* which reflects the consequences of a disaster on the economy of a country. For example, countries based on mono-economy are more vulnerable to a large scale drought.
- *Human vulnerability* which is related to human losses and injuries. Psychological damages are excluded.
- *Social vulnerability* reflects the fact that social structure influences the impact of a hazard in social terms. For example, some individuals may be more affected than other (women, families with only one parent).

It seems natural to assume that *vulnerability is specific to hazard.* For example, a country may be very vulnerable to earthquakes because of poor building quality, but less vulnerable to droughts because of it's early warning capacities.

Furthermore, vulnerability is considered to be *specific to the region*. Different countries or group of countries are not necessarily vulnerable to a particular hazard for the same reasons. For example, a country may be especially vulnerable to earthquakes because of an extremely high population density and another one vulnerable to the same hazard because of lack of access to safe water.

The goal is to provide an estimation of vulnerability for every country. It is possible of course to choose a descriptive approach and base the estimation on observed damages only. For example, one could calculate the average damages per disaster over the past 20 years. That would produce some measurement of vulnerability for countries where disasters were recorded. The major problem being that if no major disaster was recorded in a particular country, there is no way to get an idea of its vulnerability. However, it seems natural to assume that vulnerability depends on certain precise characteristics of countries. That is where statistical induction begins. In the following section, some tracks that may lead to global estimation of vulnerability are presented.

Estimation of vulnerability

Vulnerability is not directly measurable. However, it is possible to induce vulnerability using information about the characteristics of the geographical zone considered. There are multiple approaches to that problem. Three of them are presented here.

A first approach is the *conventional weighting*. It is assumed that vulnerability depends on the socio-economic context of the considered geographical unit. Information about that context is contained in observable indicators. A priori information allows for the selection of indicators that are most relevant to depict every facet of vulnerability; the selected indicators are called *vulnerability factors*. Once selected, the vulnerability factors must be combined to realise a composite index of vulnerability. Both structure (linear / non-linear) and weighting of the factors in the index are determined on a conventional basis. Thus, only a priori knowledge of the phenomenon is exploited to estimate vulnerability.

A second approach is the *inferred weighting*. The major difference between this approach and the conventional weighting is that relevant vulnerability factors are inferred from data, as well as their optimal weighing in the vulnerability composite index. In other terms, relationships between vulnerability and vulnerability factors are examined. The problem is that vulnerability remains non-observable. A solution is to work with a defined *proxy* for vulnerability. For example assumption is made that vulnerability is closely related to observed damages. Using available data on damages caused by disasters, one or several measures of damage (number of human losses, number of houses destroyed, etc.) are selected so that they can be considered as representative of damage in a general sense. Then the relationship between these variables are tested, in order to estimate the function of vulnerability factors that depicts at best the proxies for vulnerability.

A third approach is the *latent variables models*. These models take into account the fact that variables are non-observable ones.

Finally, each one of these approaches enables estimation of vulnerability as a function F of observable vulnerability factors.

$$\text{Estimated vulnerability} = F(\text{vulnerability factors})$$

Given that function *F*, it becomes possible to estimate vulnerability for every country where vulnerability factors data is available.

Hazard

Since hazard definition and estimation is beyond the scope and possibilities of the present study, hazard is simply defined as the probability of occurrence of a certain type of danger of a given intensity range. Such assumption a simplification that certainly will clarify the discussion.

4.3. Data analysis

This section provides a quick overview of the available data and discussion of particularities with direct impact on vulnerability modelisation. The goal here is to highlight choices made in the methodology. A detailed analysis of the databases in previous sections (see p. 15 for disaster databases and p. 33 for hazard databases).

Two types of data are exploited here: *disaster data* and *socio-economic indicators*.

Disaster data

In CRED database, each record refers to a particular disaster. There are 8002 records in the database. The data describe the observed damages and provide specific details about the disaster:

- *killed* : number of killed people in the disaster. Killed is zero for 1923 records and unavailable for 1281 records. Note that it is crucial to consider disasters with a number of killed of zero. Otherwise, a bias would be introduced.
- *injured* : number of injured people in the disaster. Injured is zero for 577 records and unavailable for 2063 records.
- *homeless* : number of homeless people after the disaster. Homeless is zero for 456 records and is only available for 1757 records out of 8002. In other terms, only about 20 percent of the data on homeless people is available.
- *affected* : number of people affected by the disaster. In CRED, the concept of "affected" is defined as the number of people affected by a disaster in another way than death, injury or loss of housing. Affected is zero for 210 records and is only available for 3053 records.
- *disaster type* : 10 disaster types are reported (earthquakes, droughts, land slides, floods, extreme temperatures, wave/surge, wind storms, volcanic eruptions, insect infestations and wild fires). Only natural disasters are considered, which excludes epidemics (insect infestations). Wild fires were left aside, because of their marginal impact in terms of human losses (the weight in the index would be almost zero). It is also important to note that "two or more disasters may be related, i.e. a disaster may occur as a consequence of a primary event. For example, a cyclone may generate a flood or a landslide; or an earthquake may cause a gas line to rupture, causing an ecological disaster. The primary disaster type is recorded."[5]
- *intensity* : a measure of intensity of the disaster. Scale is of course specific to disaster type. Generally speaking, there is little information about disaster intensities. And in some cases, such as extreme temperatures[6] the information can not be used at all. The following table presents basic information on that data :

---

[5] Adapted from CRED, http://www.cred.be/emdat/intro.html

[6] Since a given extreme temperature does not have a comparable impact depending on the region; it is irrelevant to use that measure of intensity when comparing disasters that happened in different regions.

Table 15. Information availability on disaster intensity

Disaster	**Intensity scale**	**Availability**
Earthquake	Richter	25% of records
Drought	Km^2	4% of records
Extreme temperature	Celsius	Unexploitable.
Flood	Km^2	7% of records
Volcano	-	Not available
Slide	-	Not available
Wind storm	Km/h	14% of records
Tsunami	metres	13% of records (6 records only)

- *localisation* : some information about the *precise localisation* of the disasters within the country is also available for some entries. The *name of the cities or regions affected* is available for about 10-30% of the records of each disaster type respectively. It is a fair proportion, but it is not reasonable to attempt to use that information. The reason is that it implies to find the coordinates of every town concerned for several thousands of disasters. However *LatLong coordinates* are available for about 25% of the records of Volcanoes and Earthquakes. Therefore, these records can easily be handled using GIS tools. The interest in mixing use of both Statistical and GIS tools in vulnerability estimation is discussed later.

Socio-economic indicators

The following table presents the national indicators used in the methodology:

Table 16. Socio-economic indicators

Variable code	Description	Sources	Comments
gdp	Gross Domestic Product (constant 1995 millions US$)	UNEP (compiled from World Bank)	1960-2000
pop	Population (persons x 1000)	UNEP (compiled from United Nations Population Division/Dept of Economic and Social Affairs, FAO, WRI)	1960-2000
lifex	Life expectancy (year)	UNEP (compiled from World Bank)	1960-2000
litrate	Literacy rate (%)	UNEP (compiled from World Resources Institute)	1960-2000
surface	Land area (km2)	WHO + ArcWorld (ESRI)	217 countries
hdi	Humand Development Index	UNDP	1998
urban	Urbanisation (%urban population)	UNEP (compiled from World Bank, FAO)	1960-2000
corup	Corruption indicator	Transparency International (TI)	CPI Score 2000

The fact that this data is only available since 1960 and not for all years excludes a number of disaster records from the analysis, because disaster data on disasters must be linked with socio-economic indicators of the corresponding country and time. For example, *corup* and *hdi* are only available for specific years; in these cases, any disaster (whatever the time period) is linked with corup_{2000} and hdi_{1998}. That way of doing is of course is subject to discussion, but it was decided to use anyway the (limited) information contained in these indicators.

Additional variables are derived from the initial variables : population densities, urbanisation growth, ...). The selection and definition of variables to be exploited in vulnerability estimation are discussed in the next section.

4.4. Methodology

This section presents the statistical approach of vulnerability modelling. However restrictions were made to the conceptual model in order to enable estimation. The aim here is to give a detailed account of the process and to highlight these restrictions in order to get an idea of the limits of our results.

Implemented conceptual model

As a first approach to vulnerability estimation, the *inferred weighting* method was chosen. The model is defined by the following set of hypothesis:

Vulnerability being non-obsevable, data on damages were used. In other terms, *observed damages* caused by a disaster were considered as a measure (proxy) of vulnerability.

$$\text{Vulnerability} \approx \text{Observed damages}$$

Only *human vulnerability* (as defined on page 45) is considered. Economic losses are not taken into account for the moment, although they may be the most relevant measure of vulnerability for certain countries and hazards e.g. earthquake in the United States. However, it would not be the case for floods in Bangladesh.

Vulnerability is only seen as *specific to hazard*, i.e. assumption is made that vulnerability structure is the same for every region of the world. This assumption is probably not defendable, but was dictated by the lack of detailed data. Furthermore, the fact of considering different regions implies an (arbitrary?) grouping of countries having common vulnerability structure.

Finally, given the poor information on disaster intensity, *intensities* are ignored in hazard definition, i.e. hazard is considered as a probability of occurrence of a given disaster, whatever its intensity. Intensity playing a key role in risk, this is a heavy assumption that may restrict the significance of the results. However, in order to keep the model as refined as possible in the study of a special case, available intensity data was exploited in the case of earthquakes.

Observed damages data

Observed damages are defined as the number of victims claimed by a disaster. Number of victims is defined as follows:

$$\text{victims} = \#\text{killed (if available)} + \#\text{injured (if available)}$$

3 disaster types are first studied: floods, wind storms and volcanoes. The case of earthquakes is studied separately later, for the reasons presented in the previous chapter. Tsunamis have been left aside because of their low significance: 0.32 % of total killed and 11 countries affected (see chapter 2, page 32 and recommendations, page 63). The choice of disaster types to be included in the risk analysis was limited by the availability of information on hazard.

Vulnerability factors data

Relevant factors in vulnerability modelling are selected from a *set of vulnerability factors*. Based on a priori information, 7 factors were pre-selected, presumably strongly correlated to vulnerability. The first 3 vulnerability factors are the ones presented in Table 16 (p. 49), and the next 4 are transformations of these socio-economic indicators:

- *urban* : an urbanisation indicator was selected in order to include in the model the fact that urban populations may be more or less exposed to a hazard than other populations, depending on the hazard. *urban* is an indicator of **affectable population**.
- *corup* : this indicator of corruption was included in the selection for it might contain information about presence of dangerous situations, e.g. houses built in hazardous areas, etc. *corup* is an indicator of **vulnerability**[7].
- *hdi* : Human Development Index was selected because it seems rather natural to assume that there is a strong correlation between a country's development level and its **mitigation** capacities. Note that nor life expectancy (*lifex*) neither literacy rate (*litrate*) were selected in the set of vulnerability factors. The reason is that *lifex* and *litrate* were strongly correlated, and that HDI provides even more information by itself.
- *popd* : population density is an indicator of **affectable population**. There is an important exposure to a given hazard if population is concentrated. This variable is calculated as follows:

$$popd = \frac{pop_t}{surface}$$

- *gdpcap* : assumption is made that GDP per capita is an indicator of **mitigation** capacities. This variable is obtained through the following formula:

---

[7] vulnerability as a component of the general notion of vulnerability (see page 45).

$$gdpcap = \frac{gdp_t}{pop_t}$$

- *urbang3* : urban growth over last 3 years. Assumption is made that fast urban growth may result in poor quality housing and thus make people more vulnerable. However this assumption may very well be only valid in particular regions. Yearly urban growth was not used because of its high variability. Considering growth over a longer time span is certainly more likely to represent a risky housing situation. In that context, *urbang3* is considered as an indicator of **vulnerability**. This variable was calculated as follows:

$$urbang3_t = \frac{urban_t - urban_{t-3}}{urban_{t-3}}$$

Note that this process suppresses 3 years of observations. Since *urban* is observable for years 1960-2000, *urbang3* is only observable for years 1963-2000.

- *popg3* : population growth over last 3 years. Assumption is made that fast population growth may create pression on housing capacities, and result in risky situations increasing **vulnerability**.

$$popg3_t = \frac{pop_t - pop_{t-3}}{pop_{t-3}}$$

For the same reasons as above, *popg3* is only observable for years 1963-2000.

Regression model

A regression model is defined for each disaster type. For every disaster type, n observations are available. Considering a given disaster type:

let *Y* the vector of n observed damages, each element of vector Y corresponds to a different disaster that happened in a particular country c at a particular time t

$$Y = [victims_{ict}]_{i=1,\dots,n}$$

and let *X* the matrix of vulnerability factors corresponding to the country and time (when possible) of $y_{ict,}$

$$X=[x1_i\ ;\ x2_i\ ;\ \dots\ ;\ x7_i]_{i=1,\dots,n}$$

where:

$$x1=popd_{ct}$$

$$x2=corup_{c2000}$$

$$x3=hdi_{c1998}$$

$$x4=gdpcap_{ct}$$

$$x5=urban_{ct}$$

$$x6=urbang3_{ct}$$

$$x7=popg3_{ct}$$

The following linear regression model is proposed:

$$Y=\beta\cdot X+\varepsilon$$

where β is the vector of parameters:

$$\beta'=[\beta_1\ ;\ \beta_2;\ \dots\ ;\ \beta_7\ ]$$

and ε is a random perturbation satisfying the usual hypothesis of classical linear regression models.

These 4 models (one model for each disaster type) are estimated using Ordinary Least Squares (OLS).

Selection of relevant factors

Each regression model is first estimated with the complete set of 7 vulnerability factors. Then, only the 3 or 4 *most significant* factors with a coherent sign are kept. Models are re-estimated with that second selection of vulnerability factors.

That process is of course not amongst the most refined but considering both nature of data and the need to keep the model as simple as possible (limited number of variables) in order to enable global implementation, this approach seemed reasonable as a first attempt of modelisation.

A special case: Earthquakes

More precise data

Vulnerability to earthquakes is basically estimated in the same way as above but with a finer approach.

First, since it was possible to locate events precisely (LatLong coordinates available in CRED) it seems natural to exploit local information as much as possible. Although most vulnerability factors are still based on national figures, it was possible to deal with *local population density* (variable code: *popdl*). Using GIS tools, the average population density within 100 km from the disaster location was considered. It appears of course more relevant to examine relationship between observed number of victims and population density *in the area where the disaster occurred*, instead of population density at country level.

Second, sufficient information on disasters intensities is available (see Table 15, p. 48). Therefore it becomes possible to use concept of *hazard* as defined in the conceptual model, i.e. considering the probability of occurrence of an earthquake of a *given intensity*. In other terms, low intensity earthquakes and high intensity earthquakes are considered as different hazards to which people are not necessarily vulnerable in the same way. Note that this is only an hypothesis that data might confirm or invalidate.

An intensity threshold

The classification of disasters into low or high intensity earthquakes was performed using the Richter scale and the number of victims as provided by CRED.

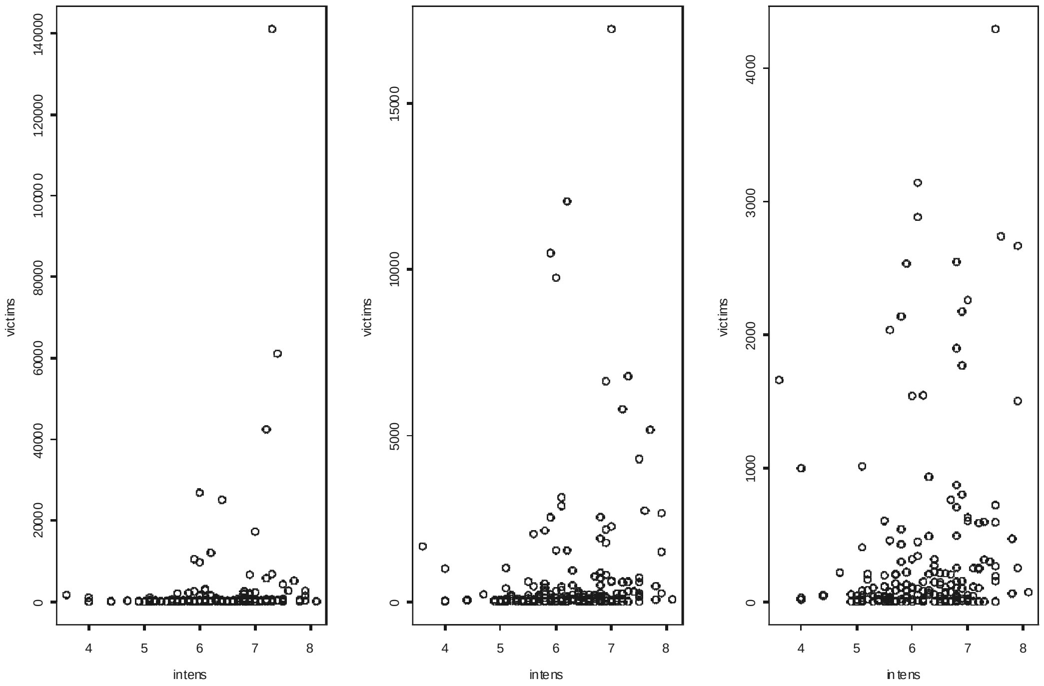
The relationship between intensities of earthquakes (on Richter scale) and numbers of victims were studied in order to determine an intensity threshold above which the number of victims is systematically higher. The following table shows that data on victims is extremely volatile:

Table 17. Summary statistics on earthquake victims

Data: *victims*	
Number of entries	222
Min	0
Median	47.5
Max	406000
Standard deviation	19979.45

In order to highlight a possible threshold, the numbers of victims were plotted against intensities with truncations of victims data. The following figure shows a first plot with the full set of data, a second one with victims data truncated at 20000 victims, and a third one with victims data truncated at 5000 victims:

Figure 21. Observed victims and disaster intensities



This figure tends to show that there is no extreme number of victims under an intensity of 5.5. Besides, variability of data is dramatically higher above that level. Based on this observation, earthquake hazard is defined as follows:

- *Low intensity earthquake hazard*: probability of occurrence of an earthquake of intensity below 5.5 on Richter scale.
- *High intensity earthquake hazard*: probability of occurrence of an earthquake of intensity higher or equal to 5.5 on Richter scale.

From here, vulnerability estimation follows the same process as for other hazard types.

4.5. Results

Vulnerability estimation

The estimated vulnerability functions are shown in Table 18. Figures give the specific weightings for each disaster type. No value means that the vulnerability factor was not selected.

Table 18. Vulnerability factor functions

hazard type	hdi Human Development Index	corup Corruption index	popd Population density	popdl Local population density	gdpcap GDP per capita	urban Urbanisation	urbang3 Urbanisation growth	popg3 Population growth
Flood	-	-	1.3	-	-	-	13653.0	-
Wind storm	-	67.4	1.2	-	-	-9.5	-	6443.9
Volcano	-	-	-	-	-0.02	10.5	-	-
Ethq. low	-1620.3	103.2	3.2	-	-0.03	16.5	-	-
Ethq. high	-	-	-	8.1	-0.09	36.9	-	-

Interpretation

The above results are briefly interpreted. Extreme differences between values of parameters are namely due to variable units. Therefore, it is hazardous to attempt to interpret parameter values. Sign and significance of the parameters is the main focus.

- *Floods* : both population density and urbanisation growth are relevant in explanation of vulnerability to floods. They have a positive impact on vulnerability. The urban population growth could be interpreted as new comers in a city have to settle in the areas left, which might not be the safer ones in many cases. Population density is quite obvious that the more people live in a water shed the higher is the physical exposure.

- *Wind storm* : population density, corruption and population growth all have a significant positive impact on expected number of victims. Besides, the parameter of urbanisation is negative, which tends to show that people living in urban areas are less vulnerable to wind storms.

- *Volcano* : on the contrary, considering vulnerability to volcanoes, results show that urban areas would be more vulnerable. Results show as well that a high GDP per capita is significant in vulnerability reduction.

- *Earthquake high/low intensity* : interpretation is very different considering low or high intensity earthquake hazard. First, results show that a high HDI reduces expected number of victims due to low intensity earthquakes; while it has non significant impact in the case of high intensity earthquakes. Same observation for the positive impact of corruption.

  Second, local population density has a significant positive impact on vulnerability to high intensity earthquakes, while it is population density at country level that has a significant positive impact on vulnerability to low intensity earthquakes. The local population density is expected to be the most relevant in both cases, but maybe that observation shows that both local and global levels should be taken into account.

  The corruption appears in the low intensity earthquakes but not in the high intensity. This could be explained that lack of appropriate regulations applications (for building codes) due to corruption is leading to increase destruction during low intensity earthquakes, whereas while a high intensity earthquake is striking a population, all the buildings are affected whether they follow the regulations or not.

  Finally, GDP per capita has the expected negative impact in both cases. Finally, results show that occurrence of earthquakes in urban areas has a significant positive impact on expected number of victims for both high and low intensity earthquakes.

Statistical adequacy

Most parameters shown in *Table 2* are significant at a 10% level. However some parameters that were not significant on a strict statistical basis (45% level) were kept for they still contain some interesting information to be interpreted.

Generally speaking, linear models assumed in the methodology show a rather poor statistical adequacy (R^2 statistics not higher than 0.4). Which tends to show that either linear structure is not optimal, or that phenomenons meant to be highlighted are not contained in data.

Finally, it must be underlined that all the results are dependant on the fact that residuals ε_i follow a Normal distribution, according to our the set of hypothesis. QQ-plot of residuals show that such hypothesis is probably not valid.

The possible use of alternative estimation methods and the need to focus on data quality are discussed later.

Application of the results: Global Risk estimation

Once vulnerabilities to each danger considered are estimated, it becomes possible to estimate risk itself.

Interpretation of risk

According to our definitions and model, interpretation of vulnerability is the "expected" (in a mathematical sense) number of victims *given the occurrence of a disaster of given type*. Risk is calculated multiplying vulnerability by hazard. Hazard being basically the yearly probability for a population of a given country to undergo a dangerous event[8].

In these terms, an interpretation of risk is the "expected" number of victims in a year *for a given country*.

Computation of risk

Estimation of vulnerability at country level is based on national figures using the estimated functions. Vulnerability to a given hazard for a given country is estimated as follows:

$$\hat{Y}_{2000} = \hat{\beta}_1 \cdot x_{1,2000} + \hat{\beta}_2 \cdot x_{2,2000} + \hat{\beta}_3 \cdot x_{3,1998} + \ldots + \hat{\beta}_7 \cdot x_{7,2000}$$

Historical data was used in vulnerability functions estimations, but it is obviously current values of vulnerability factors (not available for variable x_3) that are used for calculation of estimated vulnerability.

Based on estimated current vulnerabilities to each one of the 4 hazards[9] *h* considered for every country, risk for country *c* is calculated as follows:

$$R_c = \sum_{h=1}^{4} \hat{Y}_{hc} \, H_{hc}$$

Where H_{hc} is hazard of type *h* for country *c* and $\hat{Y}_{hc}$ is estimated vulnerability of country *c* to hazard *h*.

Risk tables

The following tables show countries' estimated specific risk. For every considered hazard the 25 countries with highest value of estimated risk are presented.

---

[8] In the case of estimation of vulnerability with national figures. This interpretation does not apply to local approach used for earthquakes.

[9] Global risk for earthquake hazard was not calculated, because it involves local population densities. Dealing with figures at a finer scale than country level implies use of GIS tools.

Table 19. Flood risk

Country	Probability	Vulnerability	Estimated risk	Estimated risk (ranked)	Realised risk (ranked
China	4.810	976	4'694	1	1
Bangladesh	2.048	1'753	3'590	2	8
Indonesia	2.667	988	2'636	3	2
India	4.000	629	2'514	4	4
Philippines	1.810	830	1'503	5	18
Malawi	0.524	2'515	1'317	6	27
Nepal	1.048	976	1'023	7	9
United Republic of Tanzania	0.714	1'389	992	8	43
Haiti	0.810	1'047	848	9	30
Ethiopia	0.952	888	845	10	38
Thailand	1.333	565	753	11	14
Republic of Korea	0.762	947	721	12	21
Sri Lanka	1.095	656	718	13	22
Brazil	2.286	282	644	14	11
Nigeria	0.667	920	613	15	53
Pakistan	1.000	550	550	16	13
Sudan	0.571	877	501	17	5
Turkey	0.667	708	472	18	41
Iran (Islamic Republic of)	1.857	248	460	19	15
Honduras	0.524	763	400	20	44
United States of America	3.524	102	358	21	33
Japan	0.667	510	340	22	24
Benin	0.476	674	321	23	62
Ecuador	0.571	554	316	24	26
Mozambique	0.333	874	291	25	31

Table 20. Volcano risk

Country	Probability	Vulnerability	Estimated risk	Estimated risk (ranked)	Realised risk (ranked
Indonesia	1.190	347	412.8	1	4
Chile	0.333	803	267.6	2	11
Mexico	0.333	691	230.3	3	5
Philippines	0.381	561	213.8	4	3
Guatemala	0.429	376	161.1	5	9
Colombia	0.190	706	134.5	6	1
Nicaragua	0.190	564	107.4	7	7
Cameroon	0.143	457	65.3	8	2
Iceland	0.143	440	62.9	9	13
Italy	0.143	318	45.5	10	13
Costa Rica	0.095	410	39.1	11	13
Papua New Guinea	0.238	145	34.5	12	10
Trinidad and Tobago	0.048	668	31.8	13	13
Ecuador	0.048	602	28.7	14	13
New Zealand	0.048	565	26.9	15	13
Cape Verde	0.048	545	26.0	16	11
United States of America	0.048	259	12.3	17	8
Comoros	0.048	252	12.0	18	13
Japan	0.476	1	0.5	19	6

Table 21. Wind storm risk

Country	Probability	Vulnerability	Estimated risk	Estimated risk (ranked)	Realised risk (ranked
Philippines	6.000	419	2’512	1	3
India	2.905	651	1’891	2	4
China	6.000	286	1’717	3	1
Viet Nam	2.429	652	1’584	4	2
United States of America	12.667	40	511	5	5
Mauritius	0.524	752	394	6	10
Japan	1.810	150	271	7	7
Mozambique	0.333	761	254	8	16
Thailand	0.714	350	250	9	15
Republic of Korea	0.905	261	236	10	11
Switzerland	0.571	321	184	11	30
Costa Rica	0.238	644	153	12	6
Jordan	0.143	813	116	13	27
El Salvador	0.190	607	116	14	23
Canada	0.714	143	102	15	12
South Africa	0.381	261	99	16	13
Israel	0.143	641	92	17	50
Netherlands	0.190	401	76	18	42
United Kingdom of Great Britain and Northern Ireland	1.048	67	71	19	24
Uganda	0.095	708	67	20	52
New Zealand	0.286	207	59	21	49
Indonesia	0.238	205	49	22	26
Austria	0.238	171	41	23	48
Egypt	0.143	276	39	24	30
Senegal	0.095	382	36	25	28

Missing countries (no data on vulnerability)

Country	probability	realised risk
Bangladesh	3.524	17,998
Honduras	0.286	860
Nicaragua	0.333	160
Madagascar	0.714	114
Pakistan	0.571	106
Haiti	0.238	83

Validity of the results

Although this report focuses on methodology development, validity of our results is necessary to highlight possible limitations in interpretations.

Previsions versus observations

In order to evaluate quality of the results, *risk predicted by the model* is compared to *realised risk*.

Since predicted risk can be interpreted as the expected number of victims in a year for a given country, it is common sense to consider that observed average yearly number of victims by country constitutes a measure of realised risk. Computation of Realised Risk (RR) relative to hazard h in country *c* is based on the *N* disasters recorded since 1980[10]. RR_{hc} is calculated as follows:

$$RR_{hc} = \frac{1}{21}\sum_{i=1}^{N} victims_i$$

The following figures show comparison of countries rankings resulting from the model and rankings resulting from Realised Risk (CRED data). Ranking plots are computed for Wind storms and Floods, because of the lack of data. Given the fact that Realised Risk calculation is necessarily based on disaster observations (is not the case for calculation of estimated risk which is based on observation of vulnerability factors) only a few countries have sufficient data allowing RR calculation.

---

[10] Truncation of the dataset was necessary in order to avoid bias introduced by limited access to information before 1980. See global report for details.

Figure 22. Compared rankings for Flood Risk

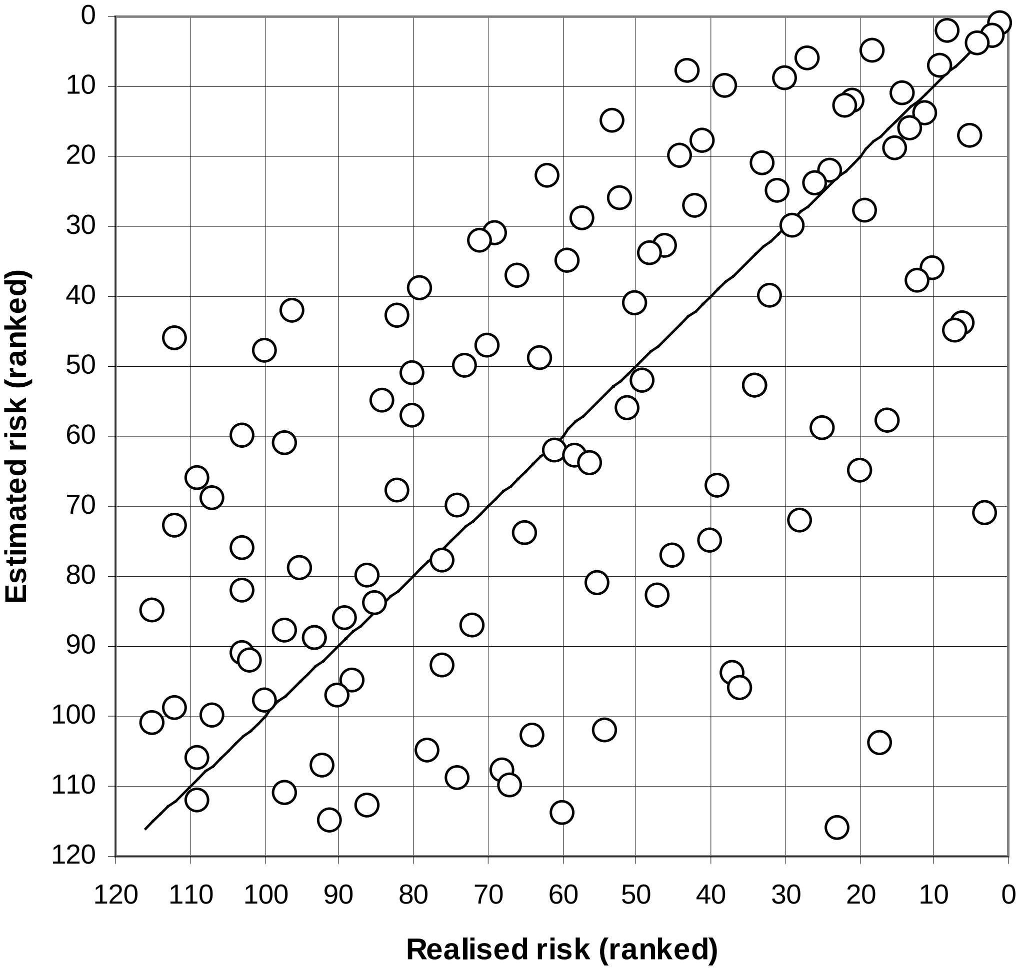


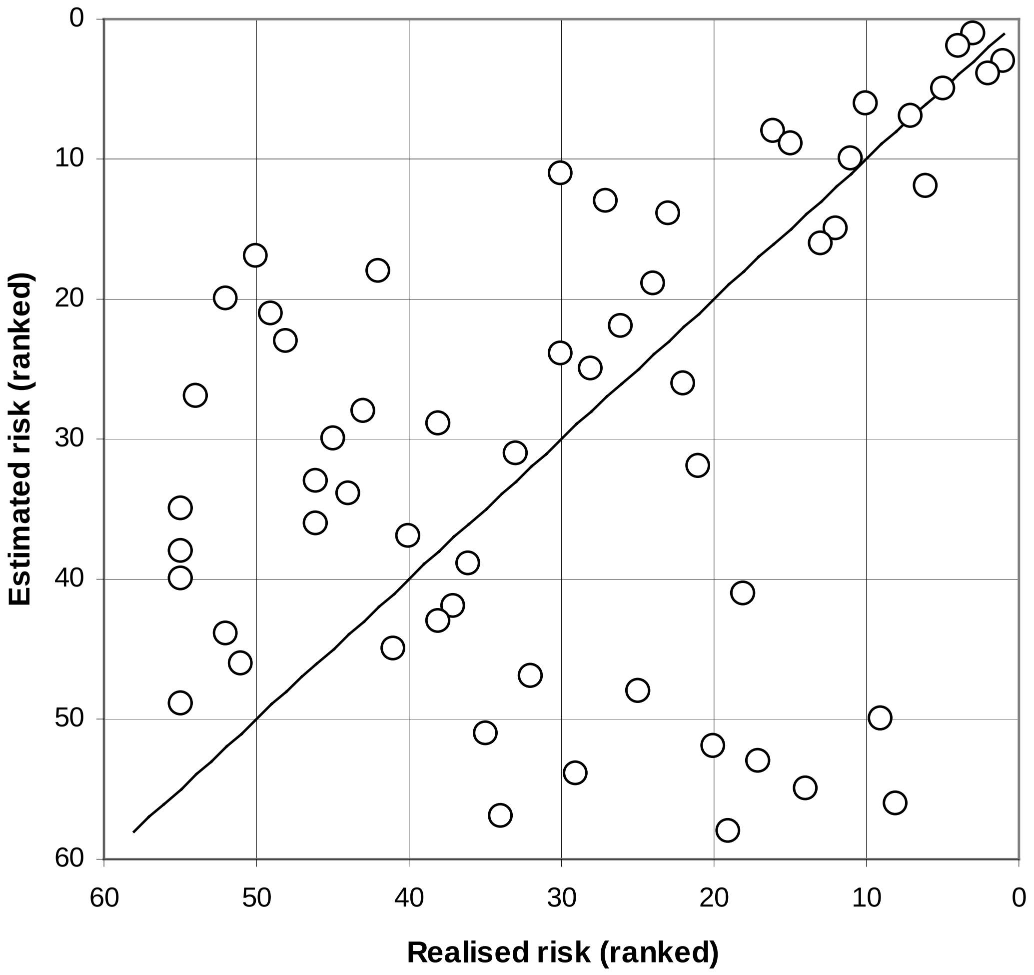
Figure 22 shows a plot of both rankings. If both methods resulted in the same ranking, dots would fit the diagonal dashed line. On the contrary if there was no relation between both rankings, it would result in a "cloud" of observations with no linear structure at all.

In the case of flood risk, a strong dependency between both rankings can be observed. Correlation coefficient is rather high:

$$\rho_{flood}=0.61$$

That tends to show that estimated risk is fairly close to realised risk. In other terms, predictions of the model are not too far from what can be observed with a simple descriptive approach. However, extreme differences in ranking should be analysed with more detail in order to highlight possible elements that would need to be included in the model.

Figure 23. Compared rankings for Wind storm Risk



Regarding wind storm risk, Figure 23 also shows a relation between estimated and observed rankings, although correlation coefficient is lower:

$$\rho_{\text{wind}}=0.37$$

The relation appears to be non-linear, which would be probably better estimated with non parametric methods (see conclusions). The estimated risk for a group of countries including Mexico, Australia, Brazil, France, Russia is much lower than what is shown in the CRED data (lower right part of the diagram).

General remarks

Regarding general methodology here are a few facts influencing validity of results:

i. As mentioned above, *disasters classification* in CRED database is based on primary cause of disaster. It was assumed in the methodology that disaster type data was accurate. Results entirely rely on that assumption. Maybe that fact can explain the poor statistical adequacy for particular hazard types. A major improvement would be to study a possible re-classification of disasters or an attempt to collect alternative data.

ii. In the implemented conceptual model, the number of victims is considered as an indicator of realised damages, which is most conventional. Although this may very well be relevant for particular disaster types, is seems necessary to define specific *realised damages* for every disaster type.

iii. For the time being, lack of data on disaster *intensities* limitates the implementation of conceptual model. Statistical adequacy in vulnerability modelling is certainly partly limited by the fact that intensities can not be taken into account as they should.

iv. For all hazards, excepting earthquakes, it was not possible to include local variables in the model. In other terms, *local phenomenons (disasters)* were modelised with national figures. Although this approach may be reasonable for large scale disasters such as droughts, it is probably not valid for local disasters. This is the major aspect on which validity of our methodology depends. It calls for important methodology development.

Considering now statistical aspects of the method, some comments on estimation methods must be made:

i. In vulnerability functions estimations, the set of explained variables is truncated (the number of victims is always positive). Therefore, OLS are not the optimal estimation method. Truncated regression would be more appropriate.

ii. Given the fact that Normality hypothesis is probably not satisfied, OLS are again not optimal. Maximum Likelihood with non-Normal residuals can be an alternative method.

5. GENERAL CONCLUSION AND RECOMMENDATIONS

5.1. Spatial Analysis

The actual analysis covers the four disasters (wind storms, floods, earthquakes and volcanoes), responsible together for 50 % of the victims according to CRED database. A GRAVITY index without drought (causing 46.5% of the victims worlwide) would be too weak and unfair for African countries. If data can be found on drought, the coverage will reach 96.5% of the victims produced by natural disasters. This would be solid enough for a GRAVITY index. Consolidation of the model should then be a priority before adding new disasters, although some specific areas may be taken into account as explained below.

Improvement of actual parameters

Earthquakes

An interesting research to approach the vulnerability would consist on comparing the physical measures (PGA from GSHAP) with intensity (index measured mainly using impacts on human infrastructures). This would enable to link strength and human impact and thus with the combination with victims, it should highlight vulnerability toward earthquake.

Cyclones

Using tropical cyclones tracks provided by Unisys and Australian Severe Weather, it should be possible to extract affected population (using a “buffer” around the track). This could be used to improve the precision of the ratio killed per affected population and thus providing improved information to introduce in the model. The task is highly time consuming as each tropical cyclone’s track has to be transformed from points to a line. This was partly done by a previous research at UNEP/GRID-Geneva and the information is compile for the year 1998 and 1999. Six remaining years would need to be transformed.

Floods

The database provided by Dartmouth Flood Observatory is on MapInfo format. During the conversion to ArcInfo, the information contains in the shapefiles was lost. It would be extremely useful to get in touch with this organisation and to ask them is the information that they have collected could be introduced in the present statistical model. This could be used to improve the computation of frequency of floods for each watershed, as well as deriving the expected duration of the event. The time dimension needs to be approached in a better way. So far the probability of occurrence was based on an average per country and some approximations were derived, however this could be refine quite easily with their inputs.

Tsunamis

Although tsunamis are not a principal cause of victims worldwide (0.32%) this disaster type is responsible of 67.8% of victims in Papoua New Guinea and 14.3% in Equator. These country would not be ranked accordingly if tsunamis are not taken into account. As already discussed previously in point 3.1, the population affected by tsunamis could not be derived as it was not possible with the actual data to extrapolate from the point coordinates to an area. Precise data on depth and height of the coast lines are requested to derive the areas potentially affected. The collection of the data does not need to be global, if the extraction could be performed for Papoua New Guinea (maybe Irian Jaya-Indonesia) and Equator, it would already cover more than 80% or the victims as recorded in the CRED database during the last 21 years.

CRED database

Some validation of the work carried out at global scale could be done using “La Red” which contains much more detailed information.

A extremely time consuming task but tremendously useful would be to georeference the events of the last two decades with a more precise location, for flood, the water basin central location would be enough, for earthquakes, the epicentre, for cyclones this might be done using the name as connector. This would allow the determination of the total affected population and the computation of the ratio killed/total affected, which could be directly measuring the vulnerability. It may be too expensive to introduce this information for the past two decades, however if CRED is going to be used as a base for the GRAVITY index, it would be a great asset. At least this information might be introduced for the next events.

Expending the model to other types of disasters

Droughts

Droughts are causing 48% of the victims, mostly in Africa but also in West Asia. Not taking into account such huge cause of death would be prejudicial.

Extensive work has already be done by FAO for these countries and probably by other organisations. As already mentioned before, drought is mainly killing through food shortage, however food shortage is a mixture between natural disasters and human made disasters (wars, political problems,…) this lead to a very complex problem for modelling. Services like Global Information Early Warning Service (GIEWS) have approach the phenomenon in a very sophisticated way. If a recommendation can be done, it is suggested that people are contacted from this service and a methodology should be produce to see how their data can be included in the present model. This should be compatible as GIEWS is using UNEP/GRID data sets for the background information.

Several index for observing droughts exist such as the Palmer index. This model uses both Normalised Difference Vegetation Index (NDVI) derived from satellite sensors – such as the National Oceanographic and Atmospheric Administration (NOAA) Advance Very High Resolution Radiometer (AVHRR) – and data on temperature to identify vegetation stress, as a proxy for drought.

Landslides

Due to the facts that landslides are:

- Very local and request extensively detailed data
- Often provoked by earthquakes or floods were better global proxy exist
- Causing – according to CRED – 1.5 % of the victims worldwide.

It is suggested that it should not be a priority task to identified them in a global way. The most affected regions of the most affected countries could be modelled first, such as Peru, Indonesia and/or Ecuador, where landslides are responsible for respectively 33.0 %, 13,9 % and 10.2 % of the victims.

Rank (killed)	Countries	number of victims	% of victims for all disasters
1	China	2884	6%
2	India	2649	3.46%
3	Indonesia	1360	13.88%
4	Peru	1286	**33%**
5	Colombia	1142	4.2%
6	Philippines	1113	4.59%
8	Ecuador	714	**10.2%**

Extreme temperatures

Various data produced by scientific organisation depicts different temperature parameters, such as annual mean or zonal anomalies. However, given our purpose, one of the most interesting seems to be a global digital data depicting monthly temperature anomalies, available at GRID-Geneva. These are monthly maps showing areas of estimated temperature anomalies, available from January 1985 until December 1991.

The source of the data is the Climate System Monitoring (CSM), Bulletin of the WMO/World Climate Data Program (WCDP), and have been digitised at GRID-Geneva.

This data could be an interesting starting point to a spatial analysis of the phenomenon. For instance, if several of these areas could be linked with specific events of CRED's database, it could eventually help to define which socio-economic parameters enhance people vulnerability to extreme temperatures in a particular region.

Table 22. Extreme temperatures (victims and % of victims for 1980-2000)

Rank (killed)	Countries	number of victims	% of victims for all disasters
1	India	6457	8.4
2	USA	3280	37.6
3	China	3037	6.4
4	Greece	1084	67.9
5	Bangladesh	923	0.5
6	Mexico	844	5.9

Extreme temperatures are responsible of 1,5% of the victims worldwide, it should not be a priority to introduce it in a global way. However, as seen on the above table, this phenomenon is responsible of 67.9% and 37.6% of all the victims from natural disasters for Greece and USA respectively. If models are derived, it is suggested that they concentrate on these countries.

5.2. Statistical analysis

Results showed presence of a *strong relationship* between damages caused by disasters and socio-economic context. Although results obtained here are still not robust enough to be properly exploited, the methodology developed can be considered as a first approach of the question and calls for major developments. First, *data quality* can be dramatically improved (even considering only vulnerability factors data) in order to better reflect the complex phenomenon to be model. Then, based on the same approach, it would certainly be valuable to explore *alternative tracks in relationship modelling*. A linear model was used because its simplicity, but there are so many alternatives. A further research should concentrates on all these possibilities.

A major advantage of the present explanatory method on simple descriptive methods is that it enables estimation of risk for every country. Even if there is no historical data on disasters at all for a given country, its *vulnerability can be modelled* with its socio-economic context. Modelling vulnerability would even allow predictions based on forecasted socio-economic context. Such models are certainly crucial in a possible temporal analysis of vulnerability.

To summarise, this first approach of the question highlighted promising tracks that call for further exploration. A major improvement could be achieved by deriving the intensity even in a broad way. This would be possible for some earthquakes, floods from other data and tropical cyclones. The new data sets will include less year of records but with more detailed information.

An immediate development could be the complementary use of non-parametric methods in vulnerability functions estimation. Instead of assuming different shapes for possible non-linear relationships, these methods not only fit curves to data but also let data define by itself shape of the relationship.

5.3. Further development

During the elaboration of this research both statistician and spatial analyst were developing and testing new methods, producing new outputs. This present study had highlight what should now really be carried out, the following tasks (for UNDP/ERD and/or GRAVITY team if contracts are renewed) are as follows:

- Contact FAO for incorporating the drought into the model.
- Define a risk indicator (or a composite risk indicator).
- Elaborate procedures for measuring mitigation at national level.
- Get appropriate tracks of windstorms to study the effects on affected population.
- Compare intensity and magnitude for earthquakes to derive vulnerability.
- Remove the regularly active volcanoes as they are mostly causing victims when a volcanic eruption is unexpected.
- Define appropriate weights according to these new inputs using proposed statistical analysis method.
- Produce a first version of the GRAVITY index for categorising countries.
- Case studies for tsunamis, landslides and extreme temperatures with some of the countries listed in p. 32. If successful these development could then be incorporated in a later version of the GRAVITY index.
- A special cases study should be considered for the islands. It was discovered that statistics results were very different for the islands and some of the extraction of data caused problem.
- Validation could be performed in two ways: firstly by testing the model in different regions/countries. Secondly, by using more detailed databases such as “La Red” for example.
- Finally, the **main improvement** that would refine the model in a significant way, would be to extract the georeference from the comments column in CRED database. This is a significant task as it is usually provide in terms of city names. However this would allowed to extracted the physical exposure (in the same way as it was performed for the 250 earthquakes) and by dividing the number of killed by the physical exposure, the vulnerability of countries could then be calculated.

5.3. Final remark

The results found in this research both in spatial and statistical analysis are promising. Given the resolution of the data provided, given the compulsory short-cuts taken, the necessary approximations for bridging the gap between the need for specific data and the lack of them, the results are better that initially expected. Not in terms of expected losses, which are quite often underestimated but in terms of rank, where the relation in rank between the expected and the realised risk is very similar.

Even in the cases where the rank is different, this does not necessary means that the model is irrelevant. Could the difference in ranking be related with the mitigation level ? In which case the graph of the Figure 22 and the Figure 23 would be extremely useful to identify the countries that should improve their level of mitigation. Or is it connected with the differences of intensities of natural disasters? In such case new detailed data sets should be found or even elaborated. A detailed analysis of these differences in ranking would be extremely useful. It should be undertaken by comparing the different indicators of the countries as well with a examination of their geographic location. Incorporation intensities and mitigation are the next aims to reach.

REFERENCES

Books and articles

Anand, S. and Sen, A., 2000, The Income Component of the Human Development Index, *Journal of Human Development*, Vol. 1, No. 1

Anderson, M. and Woodrow, P., 1989, *Rising from the Ashes: Development Strategies in Times of Disaster*. (Westview Press, Boulder).

Blaikie, P. et al., 1996, *At Risk: Natural Hazards, Peoples Vulnerability and Disasters* (Routledge).

Blong, R.J., 1984, *Volcanic Hazards, A Sourcebook on the Effects of Eruptions* (Academic Press Australia).

Bolt, B.A., Horn, W.L., Macdonald, G.A., Scott, R.F., 1975, *Geological Hazards* (Springer-Verlag Berlin-Heidelberg-New York).

Carter, N., 1991, *Disaster Management, a disaster Manager's Handbook* (Asian Development Bank, Manilla).

Simkin, T., Siebert, L., 1994, *Volcanoes of the World* (Geosciences Press in Assoc. with Smithsonian Institution, Global Volcanism Program).

Smith, K., 1996, *Environmental Hazards, Assessing risk and reducing disaster* (Routldlege: London and New York).

Tazieff, H., Sabroux, J.-C., 1983, *Forecasting Volcanic Events*. (Elsevier Science Publishers B.V.).

Tobin, G.A., Montz, B.E., 1997, *Natural Hazards, Explanation and Integration* (The Guildford Press, New York, London).

Wilches Chaux, G., 1989, *Desastres, Ecologismo y Formacion Profesional* (SENA. Colombia).

References on internet

Anderson, E., Brakenridge, G.R., 2001. NASA-supported Dartmouth Flood Observatory. http://www.dartmouth.edu/artsci/geog/floods/index.html.

Christopher W., 2000. FAQ: Hurricanes, Typhoons, and Tropical Cyclones, http://weather.about.com/science/weather/gi/dynamic/offsite.htm?site=http%3A%2F%2Fwww.aoml.noaa.gov%2Fhrd%2Ftcfaq%2FtcfaqA.html

OFDA/CRED, 2001. EM-DAT: The OFDA/CRED International Disaster Database, http://www.cred.be/emdat

Giardini, D., Grünthal, G., Shedlock K., Zhang, P., 2000. Global Seismic Hazard Assessment Program, http://www.seismo.ethz.ch/GSHAP/

International Decade for Natural Disaster Reduction, http://www.unisdr.org/unisdr/indexidndr.html

Lockridge, P., 2001. World-Wide Tsunamis 2000 B.C. – 1996, http://www.ngdc.noaa.gov/seg/hazard/tsuintro.shtml.

Peduzzi, P., 2001. Project of Risk Evaluation, Vulnerability Indexing and Early Warning (PREVIEW), http://www.grid.unep.ch/activities/earlywarning/preview/index.htm

Schwarzer, S., 2001. GEO-3 Data Portal Website, http://geo3.grid.unep.ch/

Table 23. Sources of the data sets downloaded from the Internet

Organisations	URL
Australian Severe Weather	http://australiansevereweather.simplenet.com
Carbon Dioxide InformationAnalysis Center	http://cdiac.esd.ornl.gov/
CNSS, Council of the National Seismic System	http://quake.geo.berkeley.edu/cnss/
Dartmouth Flood Observatory	http://www.dartmouth.edu/artsci/geog/floods/index.html
Global Seismic Hazard Assessment Program.	http://www.seismo.ethz.ch/GSHAP/
National GeophysicalData Center	http://www.ngdc.noaa.gov/seg/hazard/volcano.shtml
National GeophysicalData Center	http://www.ngdc.noaa.gov/seg/hazard/tsu.shtml
OFDA/CRED	www.cred.be/emdat
UNEP/GRID-Geneva: PREVIEW	http://www.grid.unep.ch/activities/earlywarning/preview/
UNEP/GRID-Geneva: GEO 3	http://geo3.grid.unep.ch/
Unisys Weather	http://weather.unisys.com
University of Dartmouth, Dartmouth Flood Observatory	http://www.dartmouth.edu/

APPENDIX I: RISK INDICATORS

Earthquakes ranking (1980-2000)

Countries	Killed	Area (1000km^2)	K/1000 km2	Killed (Rank)	K/area (Rank)	Composite ranking
Turkey	19950	779.35989	25.60	2	4	1
Iran (Islamic Republic of)	41230	1624.41612	25.38	1	5	1
Afghanistan	9333	642.13128	14.53	4	8	3
Ecuador	4306	256.99065	16.76	7	6	4
Japan	6017	372.64068	16.15	6	7	4
Taiwan	2277	36.44902	62.47	12	2	6
El Salvador	2020	20.66171	97.77	14	1	7
Mexico	8973	1961.91289	4.57	5	12	8
India	11800	3160.76002	3.73	3	14	8
Italy	2666	300.00510	8.89	11	9	10
Colombia	3035	1140.66974	2.66	9	16	11
Indonesia	4055	1908.81036	2.12	8	17	11
Philippines	1773	295.85468	5.99	16	10	13
Nepal	809	147.41610	5.49	17	11	14
Algeria	2838	2320.93274	1.22	10	18	14
Georgia	278	69.83599	3.98	23	13	16
Greece	411	131.64719	3.12	22	15	17
Myanmar	741	669.97553	1.11	18	20	18
Papua New Guinea	508	466.20641	1.09	21	21	19
Guinea	275	246.05304	1.12	24	19	20
Pakistan	674	877.55405	0.77	19	26	21
China	2077	9368.27096	0.22	13	33	22
Egypt	571	1000.25955	0.57	20	27	23
Wallis and Futuna	5	0.16699	29.94	49	3	24
Nicaragua	123	128.66866	0.96	29	23	25
Russian Federation	2001	16850.31016	0.12	15	37	25
Costa Rica	53	51.27370	1.03	33	22	27
Yugoslavia	80	101.69964	0.79	31	25	28
Chile	199	745.61128	0.27	26	32	29
Peru	272	1296.13190	0.21	25	35	30

Tsunamis (1980 – 2000)

Countries	Killed	Killed/1000km2	K/area (Rank)	Killed (Rank)	Composite ranking
Papua New Guinea	2182	4.6803	1	1	1
Ecuador	1000	3.8912	2	2	2
India	400	0.1266	3	3	3
China	149	0.0159	5	4	4
Peru	12	0.0093	6	5	5
Philippines	10	0.0338	4	7	5
Indonesia	11	0.0058	8	6	7
Bangladesh	1	0.0072	7	9	8
Colombia	3	0.0026	9	8	9
Democratic People’s Republic of Korea	0	0	10	10	10
Thailand	0	0	11	11	11

Windstorms ranking (values for 1980-2000)

Countries	Killed	Killed (Rank)	Countries	K/1000 km2	K/area (Rank)	Countries	Composite Ranking
Bangladesh	156944	1	Bangladesh	1128	1	Bangladesh	1
India	25081	2	American Samoa	108	2	Philippines	2
Philippines	18723	3	Cook Islands	102	3	Honduras	3
Viet Nam	9949	4	Montserrat	95	4	Viet Nam	4
China	9745	5	Saint Lucia	85	5	Haiti	5
Honduras	6025	6	Philippines	63	6	Nicaragua	6
USA	4718	7	Haiti	62	7	India	7
Nicaragua	2847	8	Honduras	53	8	Republic of Korea	8
Haiti	1706	9	Comoros	34	9	Taiwan	9
Mexico	1696	10	US Virgin Islands	31	10	El Salvador	10
Madagascar	1329	11	Viet Nam	30	11	Dominican Rep.	11
Republic of Korea	1240	12	Reunion	27	12	China	12
Pakistan	1106	13	Nicaragua	22	13	Madagascar	13
Japan	816	14	Saint Kitts	20	14	Japan	14
Thailand	675	15	Taiwan	13	15	Fiji	15
Taiwan	476	16	Antigua and Barb.	13	16	Pakistan	16
Mozambique	464	17	Republic of Korea	13	17	Reunion	16
Dominican Rep.	344	18	El Salvador	12	18	Saint Lucia	16
Malaysia	270	19	Tonga	10	19	Thailand	19
Guatemala	263	20	Cape Verde	9	20	Guatemala	19
El Salvador	257	21	India	8	21	Mexico	21
Great Britain and Northern Irland	246	22	Vanuatu	7	22	Comoros	22
France	237	23	Samoa	7	23	Costa Rica	23
Iran (Islamic Republic of)	234	24	Dominican Republic	7	24	Vanuatu	24
Dem. Rep. Congo	200	25	Martinique	7	25	USA	25
Senegal	187	26	Fiji	7	26	Solomon Islands	26
Russian Federation	180	27	Wallis and Futuna	6	27	Switzerland	27
Costa Rica	154	28	Jamaica	6	28	Malaysia	28
Chile	147	29	Guadeloupe	5	29	Mozambique	29
South Africa	129	30	Dominica	4	30	Senegal	30

Sources: CRED, analysis UNEP/GRID-Geneva

Volcanoes (1980 - 2000)

Coutries	Killed	K/1000km2	Killed (Rank)	K/area (Rank)	Composite Ranking
Colombia	21810	19.120346	1	2	1
Cameroon	1771	3.7938021	2	3	2
Philippines	719	2.4302472	3	4	3
Indonesia	422	0.2210801	4	5	4
Montserrat	32	277.03229	8	1	4
Japan	82	0.2200511	6	6	6
Mexico	120	0.0611648	5	8	7
Guatemala	17	0.1554252	9	7	8
USA	60	0.0063467	7	12	9
Papua New Guinea	9	0.0193048	10	9	9
Chile	6	0.0080471	11	11	11
Nicaragua	2	0.0155438	12	10	11

Drought (1980 – 2000)

Countries	Killed	K/1000 km2	K/area (Rank)	Killed (Rank)	Composite ranking
Ethiopia	300367	265.291	1	1	1
Sudan	150000	60.0948	3	2	2
Mozambique	105050	133.1442	2	3	2
Chad	3000	2.3506	5	4	4
Swaziland	500	29.1504	4	8	5
Indonesia	1266	0.6632	7	6	6
Somalia	621	0.9718	6	7	6
China	2000	0.2135	10	5	8
Uganda	115	0.473	8	11	9
Pakistan	143	0.163	12	10	10
India	410	0.1297	14	9	11
Papua New Guin.	98	0.2102	11	12	11
Kenya	85	0.1454	13	13	13
Burundi	6	0.2198	9	17	13
Guinea	12	0.0488	15	15	15
Brazil	20	0.0024	17	14	16
Philippines	8	0.027	16	16	17

Floods (1980-2000)

Countries	Killed	K/1000 km^2	Killed (Rank)	K/area (Rank)	Composite Ranking
Venezuela	30172	32.9569	1	4	1
Bangladesh	11968	86.0141	4	1	1
Nepal	5597	37.9674	6	3	3
Afghanistan	8831	13.7526	5	7	4
India	29704	9.3977	2	14	5
Viet Nam	3407	10.4315	8	10	6
Republic of Korea	1414	14.4351	12	6	6
Tajikistan	1435	10.0854	11	11	8
Puerto Rico	568	61.2344	20	2	8
Malawi	1190	10.0025	14	12	10
El Salvador	562	27.2001	21	5	10
Sri Lanka	702	10.5452	18	9	12
Pakistan	4555	5.1906	7	23	13
Philippines	1879	6.3511	10	20	13
China	27336	2.9179	3	30	15
Jamaica	118	10.6561	26	8	16
Somalia	2465	3.8573	9	27	17
Cambodia	1019	5.5773	15	21	17
Guatemala	803	7.3416	17	19	17
Haiti	250	9.1436	23	15	20
Yemen	1308	2.9637	13	29	21
Djibouti	180	8.3647	25	17	21
Seychelles	5	9.7295	29	13	21
Dem. People Rep. of Korea	607	4.9689	19	24	24
Ecuador	975	3.7939	16	28	25
Cape Verde	32	9.0462	28	16	25
Bhutan	222	5.5616	24	22	27
Honduras	491	4.3522	22	26	28
Saint Vincent	3	7.6492	30	18	28
Gambia	53	4.9276	27	25	30

APPENDIX 2: REGRESSION OUTPUTS

Linear regression models were computed with Time Series Processor (TSP) Version 4.5.

Principal elements of regression outputs are presented here for each vulnerability estimation.

```
           Flood

           =====

Number of observations = 1821     Log likelihood = -18805.4


                         Standard

Parameter  Estimate        Error       t-statistic   P-value

POPD_      1.24845       1.05098       1.18789       [.235]

URBANG3_   13653.0       4731.83       2.88536       [.004]


       Mean of dep. var. = 593.455              R-squared =
.356827E-02

  Std. dev. of dep. var. = 7405.93      Adjusted R-squared =
.302048E-02

Sum of squared residuals = .994747E+11        LM het. test = 1.00175

   Variance of residuals = .546865E+08       Durbin-Watson = 1.99586

Std. error of regression = 7395.03


           Tsunami

           =======

Number of observations = 24       Log likelihood = -201.073


                         Standard

Parameter  Estimate        Error       t-statistic   P-value

POPG3_     16563.9       8420.12       1.96719       [.049]

URBAN_     -7.30077      8.75308       -.834080      [.404]

URBANG3_   -8107.12      10820.6       -.749232      [.454]


       Mean of dep. var. = 437.583              R-squared = .111077

  Std. dev. of dep. var. = 1140.36      Adjusted R-squared = .026417

Sum of squared residuals = .265957E+08        LM het. test = 1.59935

   Variance of residuals = .126646E+07       Durbin-Watson = 2.50808

Std. error of regression = 1125.37
```

```
           Wind storm
           ==========
Number of observations = 1289     Log likelihood = -11547.0

                         Standard
Parameter  Estimate        Error       t-statistic   P-value
CORUP_     67.4019       37.4691       1.79887       [.072]
POPD_      1.19115       .442659       2.69091       [.007]
POPG3_     6443.85       1693.39       3.80531       [.000]
URBAN_     -9.49129      3.76167       -2.52316      [.012]

       Mean of dep. var. = 271.103               R-squared = .016541
  Std. dev. of dep. var. = 1896.27      Adjusted R-squared = .014245
Sum of squared residuals = .455684E+10        LM het. test = 4.28972
   Variance of residuals = .354618E+07       Durbin-Watson = 1.99204
Std. error of regression = 1883.13
```

```
           Volcano
           =======
Number of observations = 118      Log likelihood = -1088.64

                         Standard
Parameter  Estimate        Error       t-statistic   P-value
URBAN_     10.4354       5.69951       1.83093       [.067]
GDPCAP_    -.026139      .024012       -1.08856      [.276]

       Mean of dep. var. = 333.136               R-squared = .010789
  Std. dev. of dep. var. = 2481.18      Adjusted R-squared =
.226098E-02
Sum of squared residuals = .712563E+09        LM het. test = 2.14088
   Variance of residuals = .614279E+07       Durbin-Watson = 2.00474
Std. error of regression = 2478.46
```

```
           Low intensity Earthquakes
           =========================
Number of observations = 27       Log likelihood = -193.189

                         Standard
Parameter  Estimate        Error       t-statistic   P-value
CORUP_     103.150       70.6708       1.45958       [.144]
POPD_      3.18731       1.07368       2.96858       [.003]
HDI_       -1620.32      596.191       -2.71779      [.007]
GDPCAP_    -.032471      .011132       -2.91692      [.004]
URBAN_     16.4797       5.53252       2.97869       [.003]

       Mean of dep. var. = 184.074              R-squared = .375040
  Std. dev. of dep. var. = 399.324      Adjusted R-squared = .261411
Sum of squared residuals = .259291E+07        LM het. test = 1.24285
   Variance of residuals = 117860.           Durbin-Watson = 1.32959
Std. error of regression = 343.307



           High intensity Earthquakes
           ==========================
Number of observations = 162      Log likelihood = -1760.63

                         Standard
Parameter  Estimate        Error       t-statistic   P-value
POPDL_     8.10555       5.45480       1.48595       [.137]
GDPCAP_    -.090231      .132954       -.678665      [.497]
URBAN_     36.9286       24.0420       1.53600       [.125]

       Mean of dep. var. = 2573.20              R-squared = .015509
  Std. dev. of dep. var. = 12836.8      Adjusted R-squared =
.312571E-02
Sum of squared residuals = .261193E+11        LM het. test = .113867
   Variance of residuals = .164272E+09       Durbin-Watson = 2.08613
    Std. error of regression = 12816.9
```