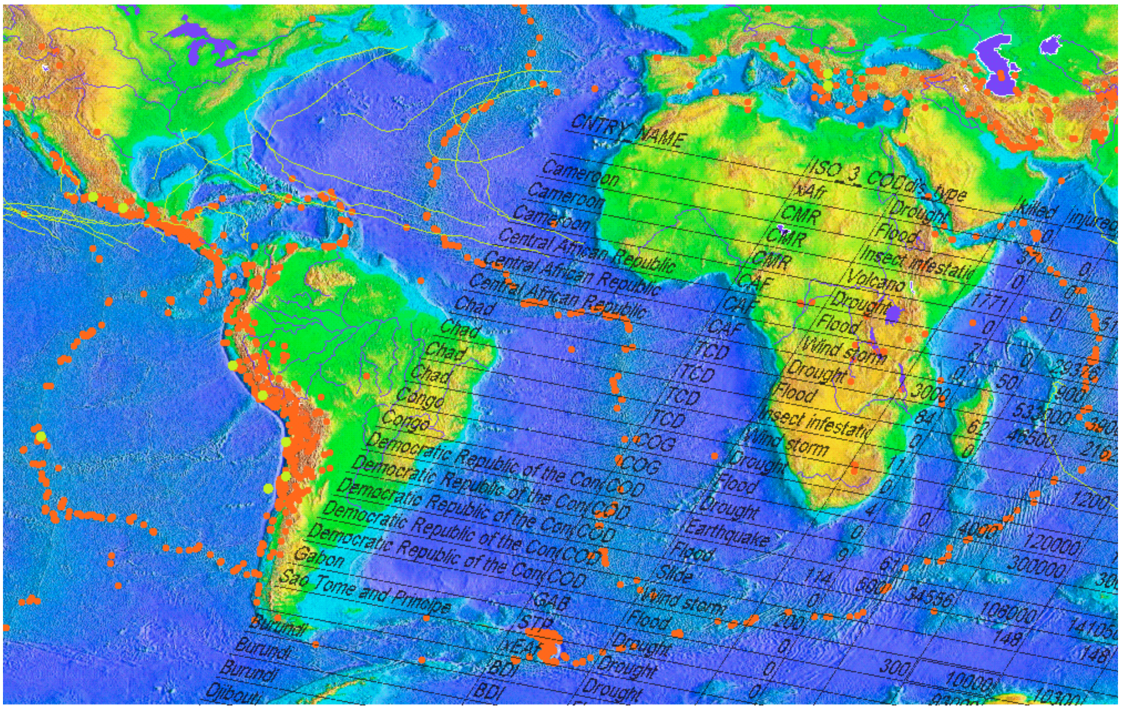
Global Risk And Vulnerability Index Trends per Year (GRAVITY)

Phase IV: Annex to WVR and Multi Risk Integration

A technical report for:

United Nations Development Programme
Bureau of Crisis Prevention & Recovery (UNDP/BCPR)



Hy Dao & Pascal Peduzzi
United Nations Environment Programme
Global Resource Information Database - Geneva
UNEP/DEWA/GRID-Geneva, May 2003

Contributors (by alphabetic order)

Main scientific collaborators:

Dr. Hy Dao, Christian Herold, Dr. Frédéric Mouton, Pascal Peduzzi.

Additional short term collaborators:

Stéphane Kluser, Antonio Martin-Diaz, Ola Nordbeck, Damien Rochette, Dr Thao Ton-That, Bernard Widmer.

Other acknowledgements

We would like also to thanks the following persons whose work has directly or indirectly contributed to the success of this research: Brad Lyon and his colleagues from IRI, Columbia University for his methodology on determining physical drought. Debarati Guha-Sapir and Regina Below from CRED for their kind and continuous support. Haris Sanahuja consultant at UNDP/BCPR for his support on socio-economical data sets and verification of the CRED drought/famine information. The GEO3 team from UNEP/GRID-Geneva for prioritising their tasks according to our needs in socio-economical data sets and our warm thanks to the World Meteorological Organisation (WMO) and its collaborators, Nanette Lombarda and their contacts Greg Holland, Bruce Harper for their very useful advices in tropical cyclones.

Disclaimer

Table of content

List of tables

List of Equations

List of Figures

Table of acronyms

AGR_{EMP}	Percentage of labour force in agricultural sector
BCPR	Bureau for Crisis Prevention and Recovery
CRED	Centre for Research on Epidemiology of Disasters
FAO	Food and Agriculture Organisation
GDP_{AGR}	Percentage of agriculture's dependency for GDP
GDP_{CAP}	Gross Domestic Product per capita
GEO	Global Environment Outlook
GIS	Geographical Information System
GLASOD	Human Induced Soil Degradation
GRAVITY	Global Risk And Vulnerability Index Trend per Year
HDI	Human Development Index
HPI	Human Poverty Index
IFRC	International Federation of the Red Cross
IRI	International Research Institute for Climate Prediction
ISDR	International Strategy of Disaster Reduction
PhExp	Physical Exposure (if not specified, for drought)
$U5_{MORT}$	Under five years old mortality rate
UNDP/BCPR	United Nation Development Programme, Bureau for Crisis Prevention and Recovery
UNEP/GRID	United Nation Environment Programme, Global Resource Information Database
WAT_{RUR}	Percentage of population having access to improved water supply in rural area
WAT_{TOT}	Percentage of population having access to improved water supply
WAT_{URB}	Percentage of population having access to improved water supply in urban area

Technical annex

Foreword

This technical annex describes the concepts, data and methods applied to achieve the Disaster Risk Index (DRI).

Results of the different analysis are presented in chapter IV.

A set of recommendations is also provided at the end of the annex on the appropriate use of the DRI as well as suggested future improvements to the methodology.

1 WORKING DEFINITIONS AND FORMULAE

1.1. Hazards, vulnerability and risk, definitions and concepts

The terminology used in this study is based on UN and other experts. The definitions of the concepts are provided in the following paragraphs:

- **Risk** : *"The term risk refers to the expected losses from a particular hazard to a specified element at risk in a particular future time period. Loss may be estimated in terms of human lives, or buildings destroyed or in financial terms"* [UNDRO 1979; in Burton et al. 1993, p.34].

  Specificity in this research : The term "risk" is used to describe potential human losses (casualties) resulting from expected future hazard.

- **Hazard** : *"The hazard can be defined as a potential threat to humans and their welfare*" [Smith, 1996]. The hazardous events varies in terms of magnitude as well as in "*frequency, duration, area extent, speed of onset, spatial dispersion, and temporal spacing"* [Burton et al. 1993, p.34].

  Specificity in this research : Only frequencies and area extent are considered in the model. The magnitude is taken into account indirectly when possible.

- **Natural Hazards** : *"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].

  Specificity in this research : Four types of natural threats are included in the model (floods, earthquakes, cyclones and droughts).

- **Physical Exposure** : **"***Elements at risk, an inventory of those people or artefacts which are exposed to the hazard"* [Coburn et al. 1991, p. 49].

  Specificity in this research : Computation of population exposed to a given hazard type. In this research the element at risk is the population.

- **Vulnerability** : *"Reflects the range of potentially damaging events and their statistical variability at a particular location*" [Smith, 1996]. *"The degree of loss to each element should a hazard of a given severity occur"* [Coburn et al. 1991, p. 49].

  Specificity in this research : The discrepancies of casualties induced by different vulnerabilities are used to identify socio-economical indicators reflecting such vulnerabilities

- **Disasters** : *"A sudden calamitous event producing great material damage, loss and distres".* [Webster's Dictionary. Found in : Carter 1991]

Specificity in this research : A disaster occurs when the high vulnerability of an exposed population intersects a hazardous event of a relatively strong magnitude.

By UN definition [UNDRO, 1979], the risk is resulting from three components:

*"**Hazard occurrence probability**, defined as the probability of occurrence of a specified natural hazard at a specified severity level in a specified future time period, **elements at risk**, an inventory of those people or artefacts which are exposed to the hazard and **vulnerability**, the degree of loss to each element should a hazard of a given severity occur"* [Coburn et al. 1991, p. 49].

1.2. Formula and method for estimating risk and vulnerability

The formula used for modelling risk combines the three components of the UNDRO definition [UNDRO 1979]: the risk is a function of hazard occurrence probability, element at risk (population) and vulnerability. The following hypothesis was made for modelling the risk: **the three factors explaining risk are multiplying each other**. This was introduced because, if the hazard is null, then the risk is null:

0 (hazard) x population x vulnerability = 0 (Risk)

The risk is also null if nobody lives in an area exposed to hazard (population = 0), same situation if the population is invulnerable, (vulnerability = 0, induce a risk = 0). Following the definition from UNDRO and the stated hypothesis the Equation 1 was derived:

Equation 1: Simplified equation of risk[1]

$$R = H \cdot Pop \cdot Vul$$

Where:

R is the risk, i.e. the expected human impacts (number of killed people).
H is the hazard, which depends on the frequency and strength of a given hazard,
Pop is the population living in a given exposed area.
Vul is the vulnerability and depends on socio-politico-economical context of this population.

From the previous discussion the physical exposure is defined as the *combination of both frequency and population exposed* (see p.1) to a given magnitude for a selected type of hazard. The hazard multiplied by the population can then be replaced by the physical exposure:

Equation 2: Simplification of risk evaluation using physical exposure

$$R = PhExp \cdot Vul$$

Where:

PhExp is the physical exposure i.e. the frequency and severity multiplied by exposed population

One way of estimating the risk is to look at impacts from previous hazardous events. The physical exposure can be obtained by modelling the area extent affected by one event. The frequency is computed by counting the number of events for the given area divided by the number of years of observation (in order to achieve an average frequency per year). Using the area affected, the number of exposed population can be extracted using a Geographical Information System (GIS), the population affected multiplied by the frequency provides the physical exposure. The identification of parameters leading to higher vulnerability can then be

---

[1] The model uses a logarithmic regression, the equation is similar but with exponent to each of the parameters.

carried out by replacing the risk in the equation by casualties reported in EM-Dat from CRED and running a statistical analysis for highlighting links between socio-economical parameters, physical exposure and observed casualties. The magnitude of the events is taken into account only by placing a threshold above which the event is included, except for earthquakes where the magnitude of the event is taken into account in the computation of the physical exposure. The magnitude is one field of new improvements needed, although subjected to some limitations as discussed later (see p.26).

The numbers of casualties can be aggregated at country level. 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 as provided by the Equation 3:

Equation 3: Estimation of the total risk

$$Risk_{Tot} = \sum (Risk_{Flood} + Risk_{Earthquake} + Risk_{Volcano} + Risk_{Cyclone} + ...Risk_{n})$$ [2]

Providing the total risk for a country induces the need to estimate the probability of occurrence and severity of each hazard, the number of persons 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.

2 CHOICE OF INDICATORS

2.1. *Spatial and temporal scales*

The risk analysis was performed on a country by country basis, i.e. the 249 countries defined in the GEO reports [UNEP 2002].

All the variables cover in principle the 21 year period ranging from 1980 to 2000. The starting date was set in 1980 because the access to information (especially on victims) was not considered as sufficiently homogenous and comparable before that year. The variables introduced in Equation 2, p.2 are aggregated figures (sum, averages) of the available data for that period, with the following major exceptions:

- Earthquake frequencies are calculated over a 36 year period, due to the longer return period of this type of disaster (1964 is the starting date for the first global coverage on earthquakes measurement).
- Cyclones frequencies are based on annual probabilities provided by CDIAC [Birdwell & Daniel 1991]
- HDI is available for the following years : 1980, 1985, 1990, 1995, 2000
- Population by grid cell (for physical exposure calculations) : 1990, 1995
- Corruption Perception Index (CPI) : 1995 to 2000

2.2. *Risk indicators*

The risk can be expressed in different ways (e.g. number of killed, percentage of killed, percentage of killed as compared to the exposed population) with their respective advantages and inconveniences (as depicted in Table 1).

---

[2] In the case of countries marginally affected by a hazard type, the risk was replaced by zero if the model could not be computed for this hazard.

Table 1. Advantages and inconveniences of respective risk indicators

Indicators for risk	Advantages	Inconveniences
Number of killed	Each human being has the same "weight"	10.000 persons killed split between ten small countries does not appeared in the same way as 10.000 killed in one country. Smaller countries are disadvantaged.
Killed / Population	It allows for comparisons between countries. Less populated countries have the same weight as more populated countries.	The "weight" of human being is not equal, e.g. one person killed in Honduras equal 160 killed in China.
Killed / Population exposed	Regional risk is highlighted even though the population affected is a smaller portion of the total national population.	This may highlight local problem that are not of national significance and give wrong priority for a selected country.

No scientific arguments can be used for selecting one indicator instead of another. At the end this is a political decision to select the indicators that best suit the purpose. The DRI is finally based on a combination of two indicators, namely the number of killed and killed per population (see p. 22). The third indicator is used as a proxy for observed vulnerability, but only for a selected hazard as exposed population to different hazard cannot be compared without standardisation.

2.3. Vulnerability indicators

The socio-economical parameters were chosen to reflect the level of quality of different constituents of a civil society such as (Table 2):

Table 2. Vulnerability indicators

Categories of vulnerability	Indicators	Drought	Flood Earthqu. Cyclones	Source[3]
Economic	Gross Domestic Product per inhabitant at purchasing power parity	X	X	WB
	Human Poverty Index (HPI)	X		UNDP
	Total dept service (% of the exports of goods and services),		X	WB
	Inflation, food prices (annual %),		X	WB
	Unemployment, total (% of total labour force)		X	ILO
Type of economical activities	%age of arable land		X	FAO
	%age of urban population		X	UNPOP
	%age of agriculture's dependency for GDP	X		WB
	%age of labour force in agricultural sector	X		FAO
Dependency and quality of the environment.	Forests and woodland (in %age of land area),		X	FAO
	%age of irrigated land		X	FAO
	Human Induced Soil Degradation (GLASOD)	X		UNEP
Demography	Population growth,		X	UNPOP
	Urban growth,		X	GRID[4]
	Population density,		X	GRID[5]
	Age dependency ratio,		X	WB
Health and sanitation	Average calorie supply per capita,		X	FAO
	%age of people with access to adequate sanitation,		X	WHO / UNICEF
	%age of people with access to safe water (total, urban, rural)	X	X	WHO / UNICEF
	Number of physicians (per 1000 inh.),		X	WB
	Number Hospital Beds		X	WB
	Life Expectancy at birth for both Sexes		X	UNPOP
	Under five years old mortality rate	X		UNPOP
Politic	Transparency's CPI (index of corruption)		X	TI
Early warning capacity	Number of Radios (per 1000 inh.)		X	WB
Education	Illiteracy Rate,		X	WB
	School enrolment,		X	UNESCO
	Secondary (% gross),		X	UNESCO
	Labour force with primary, secondary or tertiary education		X	WB
Development	Human Development Index (HDI)	X	X	UNDP

---

[3] FAOSTAT (Food and Agriculture Organisation, FAO) / GRID: UNEP/Global Resource Information Database / WB: World Development Indicators (World Bank) / TI: Transparency International / UNDP: Human Development Report (UNDP) / ILO: International Labour Office / UNPOP: UN Dep. Of Economic and Social Affairs/Population Division. Most of the data were reprocessed by the UNEP Global Environment Outlook team. Figures are available at the GEO Data Portal (UNEP), http://geodata.grid.unep.ch

[4] calculated from UNPOP data

[5] calculated from UNEP/GRID spatial modelling based on CIESIN population data.

The list of factors to be considered for the analysis was set on the basis of the following criteria:

- Relevance : select vulnerability factors (outputs orientated, resulting from the observed status of the population), not based on mitigation factors (inputs, action taken). Example : school enrolment rather than education budget.
- Data quality and availability : data should cover the 1980-2000 period and most of the 249 countries. Examples of rejected variables for the previous explained reasons: % of persons affected by AIDS, level of corruption, number of hospital beds per inhabitants.

2.4. *Data sources*

Table 3. Data sources for hazards

Hazard type	Data source
Earthquakes	Council of the National Seismic System (as of 2002), *Earthquake Catalog*, http://quake.geo.berkeley.edu/cnss/
Cyclones	Carbon Dioxide InformationAnalysis Center (1991), *A Global Geographic Information System Data Base of Storm Occurrences and Other Climatic Phenomena Affecting Coastal Zones*, http://cdiac.esd.ornl.gov/
Floods	U.S. Geological Survey (1997), *HYDRO1k Elevation Derivative Database*, http://edcdaac.usgs.gov/gtopo30/hydro/
Droughts (physical drought)	IRI/Columbia university, National Centers for Environmental Prediction Climate Prediction Center (as of 2002), *CPC Merged Analysis of Precipitation (CMAP) monthly gridded precipitation*, http://iridl.ldeo.columbia.edu/

Table 4. Data sources for victims, population and vulnerability factors

Theme	Data source
Victims (killed)	Université Catholique de Louvain (as of 2002), *EM-DAT: The OFDA/CRED International Disaster Database*, http://www.cred.be/ (for droughts, victims of famines were also included on a case by case basis by UNDP/BCPR)
Population (counts)	CIESIN, IFPRI, WRI (2000), *Gridded Population of the World (GPW), Version 2*, http://sedac.ciesin.org/plue/gpw/
	UNEP,CGIAR,NCGIA (1996), *Human Population and Administrative Boundaries Database for Asia*, http://www.grid.unep.ch/data/grid/human.php
Vulnerability factors	
Human Development Index (HDI)	UNDP (2002), *Human Development Indicators*, http://www.undp.org/
Corruption Perceptions Index (CPI)	Transparency international (2001), *Global Corruption Report 2001*, http://www.transparency.org/
Soil degradation (% of area affected)	ISRIC, UNEP (1990), *Global Assessment of Human Induced Soil Degradation (GLASOD)*, http://www.grid.unep.ch/data/grid/gnv18.php
Other socio-economic variables	UNEP/GRID (as of 2002), *GEO-3 Data portal*, http://geodata.grid.unep.ch/ (data compiled from World Bank, World Resources Institute, FAO databases)

3 COMPUTATION OF PHYSICAL EXPOSURE

3.1. General description

In broad term, the physical exposure was estimated by multiplying the hazard frequency by the population living in the exposed area. The frequency of hazard was derived for different strengths of events and the physical exposure was computed as in Equation 4:

Equation 4: Computation of physical exposure

$$\mathrm{PhExp}_{\mathrm{nat}} = \sum F_i \cdot Pop_i$$

Where:

$PhExp_{nat}$ is the physical exposure at national level
F_i is the annual frequency of a specific magnitude event in one spatial unit
Pop_i is the total population living in the spatial unit

For the case of earthquakes, the computation of a frequency could not be derived as the available information consisted on a 90% probability of an earthquake to be smaller than a given magnitude. To overcome this difficulty, the physical exposure computation was made by adding the population affected and then divided by the number of year as shown in Equation 5.

Equation 5: Physical exposure calculation without frequency

$$PhExp = \sum \frac{Pop_i}{Y_n}$$

Where:

Pop_i is the total population living in a particular buffer which radius from the epicentre varies according to the magnitude.
Y_n is the length of time in year for the cyclones (11)
PhExp is the total physical exposure of a country is the sum of all physical exposure of this country

Once the hazarduous event area was computed using UNEP/GRID-Geneva methods for earthquakes, floods and cyclones or using IRI's method for drought, then the affected population was computed for each affected area and then this number was aggregated at national level as needed in order to associate the victims from the last 21 years with the physical exposure and with socio-economical variables.

Depending on type of hazards and the quality of data, different methods were applied. Extraction of population was based on the CIESIN, IFPRI, WRI Gridded Population of the World (GPW, Version 2) at a resolution of 2.5' [6] (equivalent to 5 x 5 km at the equator). This layer was further completed by Human Population and Administrative Boundaries Database for Asia (UNEP) for Taiwan and CIESIN Global Population of the world version 2 (country level data) for ex-Yugoslavia. These datasets reflect the estimated population distribution for 1995. Since population growth is sometimes very high in the 1980-2000 period, a correction factor using country totals was applied in order to estimate current physical exposures for each year as follows :

Equation 6 : computation of current physical exposure

$$PhExp_i = \frac{Pop_i}{Pop_{1995}} \cdot PhExp_{1995}$$

Where:

$PhExp_i$ is the physical exposure of the current year
Pop_i is the population of the country at the current year
Pop_{1995} is the population of the country in 1995
$PhExp_{1995}$ is the physical exposure computed with population as in 1995

Due to the resolution of the data set, the population could not be extracted for some small islands. This has lead to the non-consideration of the small islands (even for large archipelagos). Refined study should be carried out in a further research (see recommendations in conclusion). Apart from these limitations, the extraction of the population living in exposed area is a simple task performed with a GIS.

The main difficulty consists in the evaluation of hazard area extent, frequency and intensity. At a global scale, data are not complete and generalisation is the rule. Help of specialists was asked in order to review the necessary simplifications. Out of the four hazards studied, only the case of floods requested the complete design of a global dataset built by

---

[6] GPW2 was preferred to the ONRL Landscan population dataset despite its 5 times lower spatial resolution (2.5' against 30") because the original information on administrative boundaries and population counts is almost two times more precise (127,093 administrative units against 69,350 units). Furthermore the Landscan dataset is the result of a complex model which is not explained thoroughly and which is based, among other variables, on environmental data (land-cover), making it difficult to use for further comparison with environmental factors (circularity).

linking CRED information with USGS watersheds. Drought maps were provided by the International Research Institute for Climate Prediction (IRI). For the other hazards, independent global datasets had already been updated, compiled or modelled by UNEP/GRID-Geneva and were used to extract the population. The Mollweide equal-area projection was used when calculations of areas were needed.

Although the quality can always be improved, the greatest care was taken and the level of accuracy achieved is believed to be relevant and appropriate for a global scale study.

3.2. *The case of cyclones*

The absence of complete coverage (i.e. India, Bangladesh and Pakistan data) prevents the use of the UNEP/GRID-Geneva PREVIEW Global Cyclones Asymmetric Windspeed Profile dateset. The data used to map cyclone hazard areas are produced by the Carbon Dioxide Information Analysis Centre [Birdwell & Daniel 1991]. These delineate annual probabilities of occurrence of tropical cyclones. The spatial unit is a 5 x 5 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 5. Wind speeds and appellations

Wind speeds	**Name of the phenomenon**
≥ 17 m/ s	Tropical storms
≥ 33 m/ s	Hurricanes, typhoons, tropical cyclones, severe cyclonic storms (depending on location[7])
≥ 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[7]: "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" [Landsea 2000].

The CDIAC is providing probability of occurrence for these three types of events. The average frequency (per year) was computed using Equation 7:

Equation 7: From probability to annual frequency for cyclones

$$E(x) = \lambda = -\ln(1 - P(x \geq 1))$$

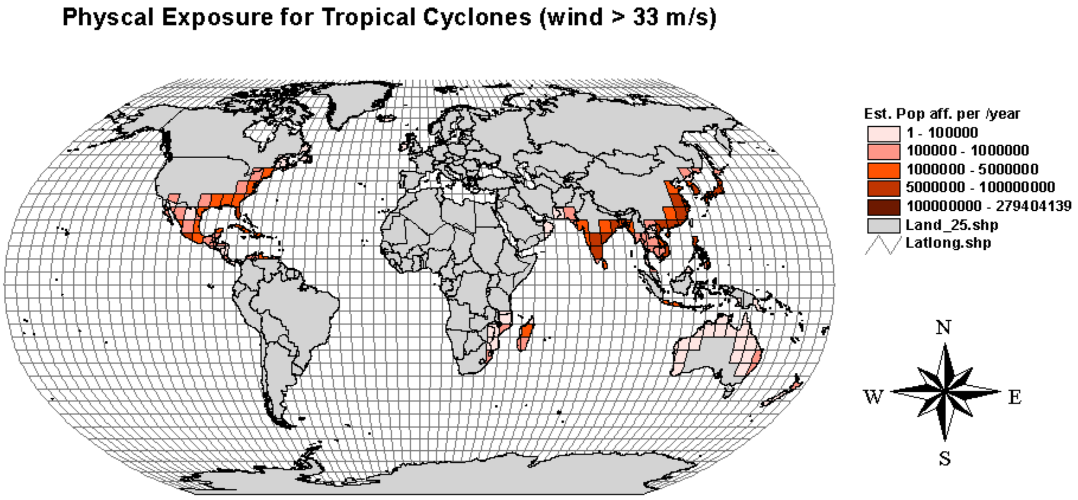
Where:

E(x) is the statistical expectation, i.e. the average number of events per year = λ
P(x) is the probability of occurrence

Hurricanes:	*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*

Source: NOAA/AOML, *FAQ: Hurricanes, Typhoons, and Tropical Cyclones*, http://www.aoml.noaa.gov/hrd/tcfaq/tcfaqA.html#A1

Figure 1: Example of physical Exposure for Tropical Cyclones



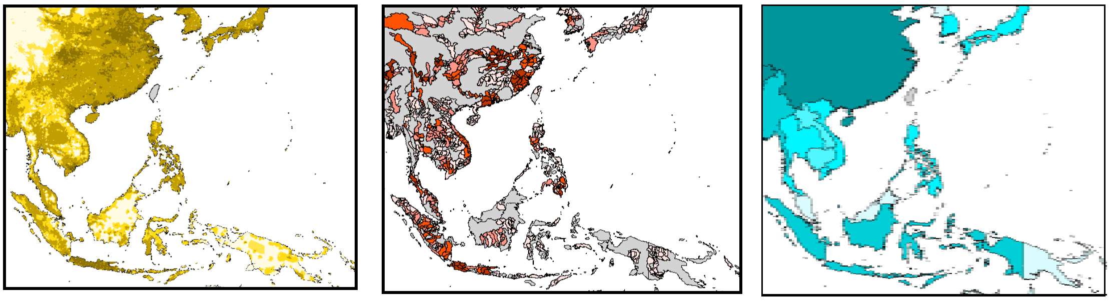
To obtain the physical exposure, a frequency per year is derived for each cell. Cells are divided to follow country borders, then population is extracted and multiplied by the frequency in order to obtain the average yearly physical exposure for each cell. This physical exposure is then summed by country for the three types of cyclones.

Physical exposure to tropical cyclones of each magnitude was calculated for each country using Equation 4 (p.7).

3.3. *The case of floods*

For the floods the method is slightly different. No global database on floods could be found except the one from Dartmouth Flood Observatory, but which was not covering the period of interest. Due to the lack of information on the duration and severity of floods, only one class of intensity could be made. Using the information in the column “comment” (e.g. name of town, river, valley,…) in the CRED database, a georeference of the floods was produced and a link between the watersheds and the events was made. Watersheds affected were mapped for the period 1980-2000. A frequency was derived for each watershed by dividing the total number of events by 21 years. The watersheds were then splitted to follow country borders, then population was extracted and multiplied by the frequency. The average yearly physical exposure was then summed at a country level using Equation 4 (p.7).

Figure 2: population, Frequency and Physical exposure for floods



Population

Frequency for each watershed

National physical exposure per year

3.4. *The case of earthquakes*

A choice was made to produce seismic hazard zones using the seismic catalogue of the CNSS (Council of the National Seismic System). The earthquakes records of the last 21 years (1980-2000) were grouped in five magnitude classes and a buffer with a radius length from the epicentres varying according to these classes. Choices of specific radius were made considering the table Table 6.

Table 6. Definition of radius for earthquakes based on bracketed duration (expressed in seconds).

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**

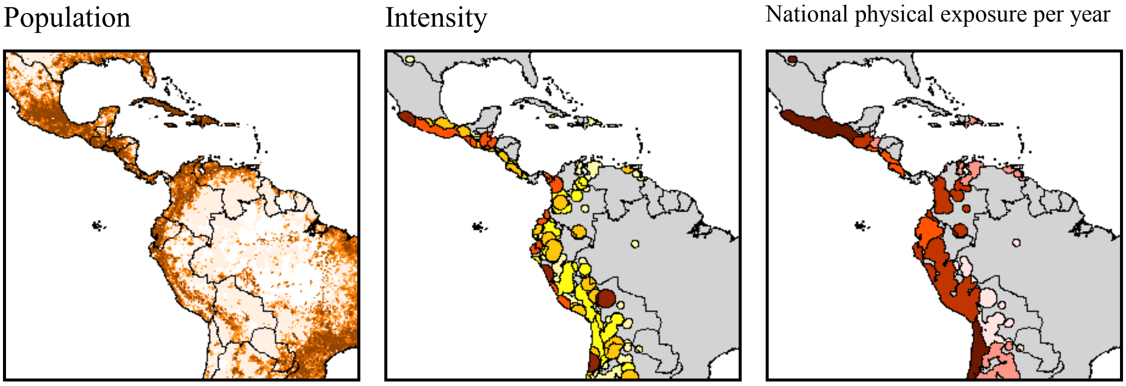
table source [Bolt et al. 1975] Acceleration > 0.05 g = ~ 0,49 m/s2, frequency > 2 Hz

These above numbers delineate the estimations of ground motions duration for specific acceleration and frequency ranges, according to magnitude and distance from epicentre [Bolt et al. 1975]. Bracketed duration is "the elapsed time (for a particular acceleration and frequency range) between the first and last acceleration excursions on the record greater than a given amplitude level (for example, 0.05 g)" [Bolt et al. 1975].

According to these figures, a specific buffer distance is defined for each class of magnitude to limit area affected by ground motions: 75 km for Magnitude ≤ 6.2, 125 km for M = 6.3 – 6.7, 150 km for M = 6.8 – 7.2, 175 km for M = 7.3 – 7.7, 200 km for M ≥ 7.8. This is a general approach that does not take into account any regional effects, for instance soil conditions or geotectonic characteristics.

Physical exposure to earthquakes was calculated for each country and each magnitude class using Equation 5 (p. 8), for details on how this could be improved, see recommendations (p. 27).

Figure 3: Population, intensity and physical exposure for earthquakes


Population

Intensity

National physical exposure per year

3.5. *The case of droughts*

Identification of drought

Explanations from the author : Brad Lyon, International Research Institute for Climate Prediction (IRI)

The data used in the analysis consists of gridded monthly precipitation for the globe for the period 1979-2001. This dataset is based on a blend of surface station observations and precipitation estimates based on satellite observations. Further details of the data are described at the end of this section.

The first step in assessing the exposure to meteorological drought was to compute, for each calendar month, the median precipitation for all grid points between the latitudes of 60S and 70N over the base period 1979-2001 (the 23-yr. period for which the data was available). Next, for each gridpoint, the percent of the long-term median precipitation was computed for every month over the period Jan 1980 to Dec 2000. For a given month, gridpoints with a long-term median precipitation of less than 0.25 mm/day were excluded from the analysis. Such low median precipitation amounts can occur either during the "*dry season*" at a given location or in desert regions and in both cases our definition of drought does not apply. Finally, a drought "event" was defined as having occurred when the percent of median precipitation was at or below a given threshold for at least 3 consecutive months. The different thresholds considered were 50%, 75% and 90% of the long-term median precipitation with the lowest percentage indicative of the most severe drought according to this method. The total number of events over the period 1980-2000 were thus determined for each gridpoint and the results plotted on global maps.

Computation of physical exposure

Using the IRI/Columbia University data set, the physical exposure was estimated by multiplying the frequency of hazard by the population living in the exposed area. The events were identified using different measurements based on severity and duration as described in Table 7. For each of the 6 following definitions, the frequency was then obtained by dividing the number of events by 21 years, thus providing an average frequency of events per year.

Table 7. Definition of drought

Duration	**Severity**
3 months	90% of median precipitation 1979-2001 (-10%)
3 months	75% of median precipitation 1979-2001 (-25%)
3 months	50% of median precipitation 1979-2001 (-50%)
6 months	90% of median precipitation 1979-2001 (-10%)
6 months	75% of median precipitation 1979-2001 (-25%)
6 months	50% of median precipitation 1979-2001 (-50%)

The physical exposure was computed as in Equation 4 (p.7) for each of the drought definitions. The statistical analysis selected the best fit which was achieved with droughts of three month duration and 50% decrease in precipitation.

4 STATISTICAL ANALYSIS : METHODS AND RESULTS

4.1. Defining a multiplicative model

The statistical analysis is based on two major hypotheses. Firstly, that the risk can be approached by the number of victims of past hazardous events. Secondly, that the equation of risk follows a multiplicative model as in Equation 8:

Equation 8. Estimation of killed

$$K = C \cdot (PhExp)^{\alpha} \cdot V_1^{\alpha_1} \cdot V_2^{\alpha_2} \ldots \cdot V_p^{\alpha_p}$$

Where:
K is the number of persons killed by a certain type of hazard.
C is the multiplicative constant.
PhExp is the physical Exposure: population living in exposed areas multiplied by the frequency of occurrence of the hazard.
V_i are the socio-economical parameters.
α_i are the exponent of Vi , which can be negative (for ratio)

Using the logarithmic properties, the equation could be written as follows:

Equation 9. Logarithm properties

$$\ln(K) = \ln(C) + \alpha \ln(PhExp) + \alpha_1 \ln(V_1) + \alpha_2 \ln(V_2) + \ldots + \alpha_p \ln(V_p)$$

This equation provides a linear relation between logarithmic sets of values. Significant socio-economical parameters V_i (with transformations when appropriate) and exponents α_i could be determined using linear regressions.

4.2. Detailed process

Data on victims

The number of killed was derived from the CRED database, and computed as the average number of killed per year over the 1980-2000 period.

Filtering the data

The statistical models for each disaster type were based on subsets of countries, from which were excluded:

- Countries with no physical exposure or no victims reported (zero or null values)
- Countries with dubious data on physical exposure (e.g. the case of Kazakhstan for floods) or socio-economic factors (100% access to water in North Korea)
- Countries with low physical exposure (smaller than 2 percent of the total population) because socio-economical variables are collected at national scale. Attempts delineate that the exposed population needs to be of some significance at national level to reflect a relationship in the model.
- Countries without all the selected socio-economic variables.
- Eccentric values, when exceptional events or other factors would clearly show abnormal level of victims like hurricane Mitch in Nicaragua and Honduras or droughts in Sudan and Mozambique (i.e. probably more related to political situation than from physical drought).

Transformation of variables

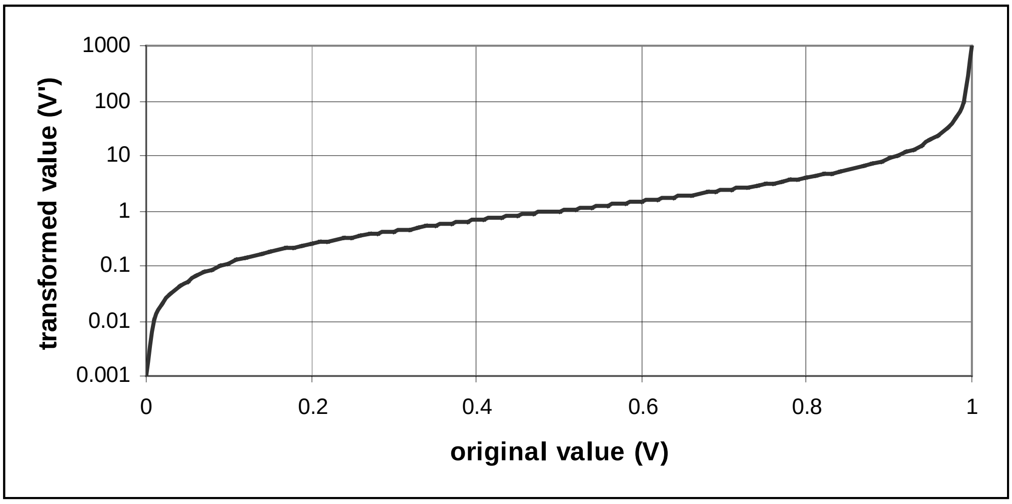
The average of socio-economical parameters was computed for the 21 year period. For some of the indicators the logarithm was computed directly, for other parameters expressed in percentage, a transformation was applied in order that all variables would range between -∞ and +∞. This appeared to be relevant as some of the transformed variables were proved to be significant in the final result. For others no logarithmic transformation was needed, for instance the population growth already behaves in a cumulative way.

Equation 10. Transformation for variables ranging between 0 and 1

$$V_i' = \frac{V_i}{(1 - V_i)}$$

Where:

V'_i is the transformed variable (ranging from -∞ to +∞)
V_i is the socio-economical variable (ranging from 0 to 1)



Choice between variables

One important condition, when computing regressions, is that the variables included in a model should be independent, i.e. the correlation between two sets of variables is low. This is clearly not the case of HDI and GDPcap purchasing power parity (further referred as GDPcap), which are highly correlated. GDPcap was more used than HDI uniquely because HDI was not available for several countries. In order to keep the sample as complete as possible a choice of available variables had to be made. This choice has been performed by the use of both matrix-plot and correlation-matrix (using low correlation, hence low p-value, as selection criteria).

The stepwise approach

For each type of hazard, numerous stepwise (back and forth steps) linear regressions were performed in order to highlight significant variables. The validation of regression was carried out using R^2, variance analysis and detailed residual analysis.

Once the model was derived, the link between estimated killed and number of killed observed was provided by both graphical plots and computation of Pearson correlation coefficients in order to ease the visualisation of the efficiency by the readers.

If intuitively one can understand that physical exposure is positively related with the number of victims and that GDPcap is inversely related with the number of victims (the lowest the GDP the highest the victims), this is less obvious for other variables such as percentage of arable land for example. This method allows the estimation of the α_i coefficients. Their signs provide information on if the variables are in a numerator or denominator position.

This model allows the identification of parameters leading to higher/lower risk, but should not be used as a predictive model, because small differences in logarithm scale induce large ones in the number of killed.

The results following this method are surprisingly high and relevant, especially considering the independence of the data sources (no auto-correlation suspected), the non consideration of the magnitude of hazard and the coarse resolution of the data at global scale.

4.3. Mapping Risk

A subjective – political – choice belonging to UNDP had to be made between the different risk indicators (i.e. killed, killed per million inhabitant,… , see p. 3). The UNDP aims to provide categories of countries taking into account both risk and disaster reduction measures.

Number of categories and method of classification

The number of classes was chosen taking into account by UNDP requirements, but also taking into consideration the levels of error and uncertainty of the data. **So far, the precision and quality of the data, as well as the sensitivity of the model do not allow the ranking of countries**. However, for the risk component, results indicate a possibility to provide five classes of countries. This number of categories minimise the error of misclassification and is simple enough to be incentive to persons not too familiar with statistics.

The number of classes and method for classifying the maps were chosen according to several criteria such as the optimum number of classes for visual representation or the number and level of errors between modelled and observed classes. According to these tests, with the aim to minimise internal class distances and maximise distances between classes, the number of five classes based on a Cluster analysis for grouping countries using both killed and killed per millions inhabitant was chosen.

4.4. Cyclones

Statistical model

The multiple regression was based on 32 countries and the best fit regression line follows the following model:

Equation 11 Multiple logarithmic regression model for cycloness

$$\ln(K) = 0.63\ln(PhExp) + 0.66\ln(\overline{Pal}) - 2.03\ln(\overline{HDI}) - 15.86$$

<u>Where:</u>
K is the number of killed from cyclones
PhExp is the physical exposure to cyclones

$\overline{Pal}$ *is the transformed value of percentage of arable land*

$\overline{HDI}$ *is the transformed value of the Human Development Index*

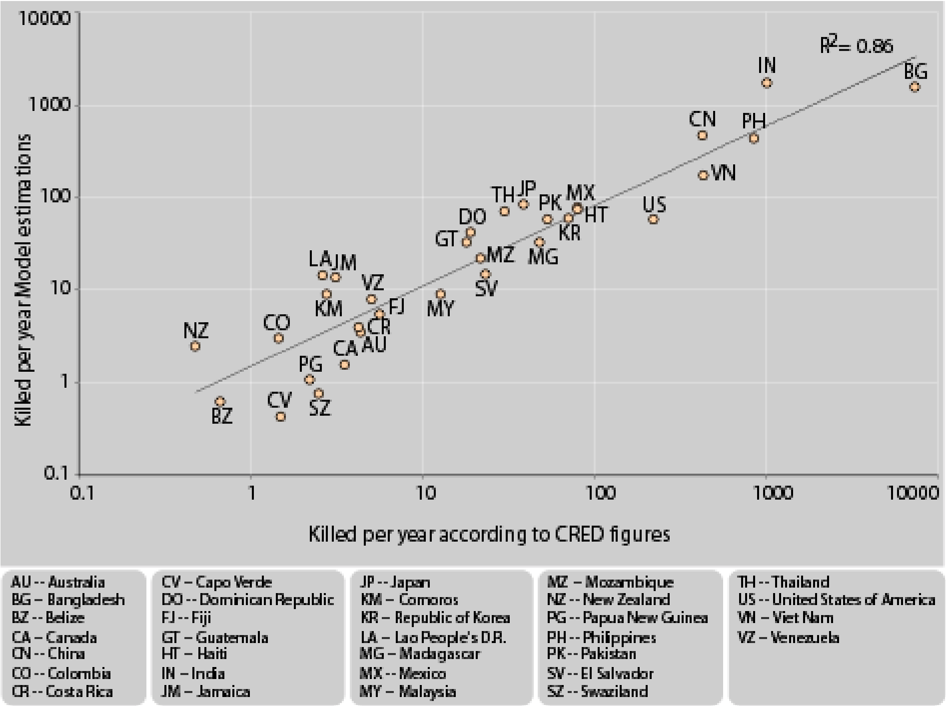
Table 8. Exponent and p-value for cyclones multiple regression

32 countries	B	p-value[8]
Intercept	-15.86	0.00000
ln(PhExp)	0.63	0.00000
$\ln(\overline{Pal})$	0.66	0.00013
$\ln(\overline{HDI})$	-2.03	0.00095

R= 0.93, R^2= 0.86, adjusted R^2= 0.85

The plot delineates a nice linear distribution of the data as seen in Figure 4:

Figure 4: Scatter plot of the observed number of people killed by windstorms (CRED figures) and the model predictions



---

[8] In broad terms, a p-value smaller than 0.05, shows the significance of the selected indicator, however this should not be used blindly.

The parameters highlighted show that besides the physical exposure, HDI and the percentage of arable land are selected indicators for vulnerability to cyclone hazards.

The percentage of arable land is probably an indirect way of measuring the dependency of a population from the agricultural activity. According to the analysis, a stronger dependence to agriculture is inducing a higher vulnerability. Although this was already mentioned by experts, it is now confirmed by statistical evidences. After a cyclone, economies relying on third sector are less affected than economy relying on agriculture, fields being devastated. The GDPcap is strongly correlated with the HDI or negatively with the percentage of urban growth. In most of the cases the variable GDPcap could be replaced by HDI as explained previously (see p.13). However, these results depict with confidence that poor countries and less developed in terms of HDI are more vulnerable to cyclones.

With a considerable part of variance explained by the regression ($R^2 = 0.863$) and a high degree of confidence in the selected variables (very small p-value) over a sample of 32 countries, the model achieved is solid.

In the model, the consequences of Mitch could easily be depicted. Indeed, Honduras and Nicaragua were far off the regression line (significantly underestimated). This is explained by the incredible difference of intensity of Mitch and other hurricanes. Mitch is a type of hazards on its own, the difference of intensity made this event impossible to compare with the other hurricanes. This is explaining the rejection of these two countries from the model.

4.5. *Floods*

Statistical model

The multiple regression was based on 90 countries and the best fit regression line follows the following model:

Equation 12. Multiple logarithmic regression model for floods

$$\ln(K) = 0.78\ln(PhExp) - 0.45\ln(GDP_{cap}) - 0.15\ln(D) - 5.22$$

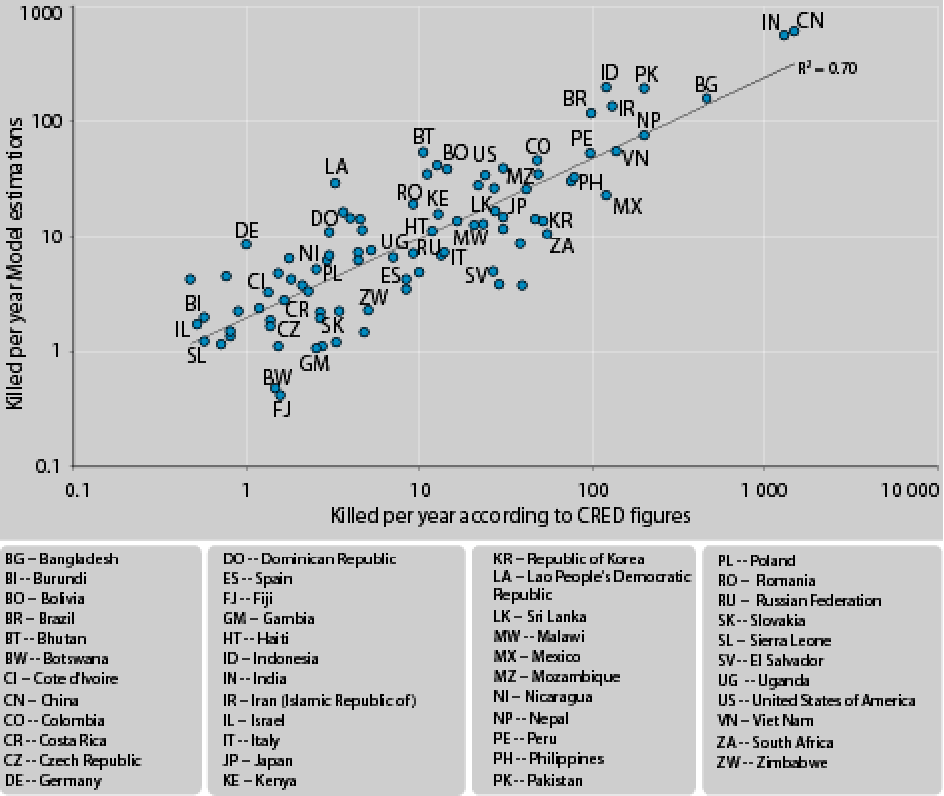
Where:
K is the number of killed from floods
PhExp is the physical exposure to floods
GDP_{cap} is the normalised Gross Domestic Product per capita (purchasing power parity)
D is the local population density (i.e. the population affected divided by the area affected)

Table 9: Exponent and p-value for flood indicators

90 countries	B	p-value (8)
Intercept	-5.22	0.00000
ln(PhExp)	0.78	0.00000
ln(GDPcap)	-0.45	0.00002
ln(Density)	-0.15	0.00321

R= 0.84, R^2= 0.70, adjusted R^2= 0.69

Figure 5. Scatter plot of the observed number of people killed by floods (CRED figures) and model predictions



Due to space constraints, only a selection of countries was included in the above scatter plot, a comprehensive list of countries affected by floods is provided below:

Albania, Algeria, Angola, Argentina, Australia, Austria, Azerbaijan, Bangladesh, Benin, Bhutan, Bolivia, Botswana, Brazil, Burkina Faso, Burundi, Cambodia, Cameroon, Canada, Chad, Chile, China, Colombia, Costa Rica, Cote d'Ivoire, Czech Republic, Dominican Republic, Ecuador, Egypt, El Salvador, Ethiopia, Fiji, France, Gambia, Georgia, Germany, Ghana, Greece, Guatemala, Haiti, Honduras, India, Indonesia, Iran (Islamic Republic of), Israel, Italy, Jamaica, Japan, Jordan, Kenya, Lao People's Democratic Republic, Malawi, Malaysia, Mali, Mexico, Moldova, Republic of Morocco, Mozambique, Nepal, Nicaragua, Niger, Nigeria, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Republic of Korea, Romania, Russian Federation, Rwanda, Saudi Arabia, Sierra Leone, Slovakia, South Africa, Spain, Sri Lanka, Thailand, Tunisia, Turkey, Uganda, Ukraine, United Kingdom of Great Britain and Northern Ireland, United Republic of Tanzania, United States of America, Viet Nam, Yemen, Zimbabwe,

The variables selected by the statistical analysis are physical exposure, GDPcap and local density of population. GDPcap being highly correlated with HDI, this later could have been chosen as well. The GDPcap was chosen due to slightly better correlation between the model and the observed killed, as well as because of lower p-value. Regression analysis supposes the introduction of non-correlated parameters, thus preventing the use of all these variables.

Without surprise, the regression proves that highly exposed and poorer populations are more subject to suffer casualties from floods. More surprisingly, it shows that countries with low population density are more vulnerable than countries with high population density. The part of explained variance ($R^2 = 0.70$) associated with significant p-value (between 10^{-23} and $2 \cdot 10^{-3}$) on 90 countries is confirming a solid confidence in the selection of the variables (see Table 9).

4.6. *Earthquakes*

Statistical model

The multiple regression was based on 48 countries and the best fit regression line follows the following model:

Equation 13. Multiple logarithmic regression model for earthquakes

$$\ln(K) = 1.26\ln(PhExp) + 12,27 \cdot U_g - 16.22$$

<u>Where:</u>

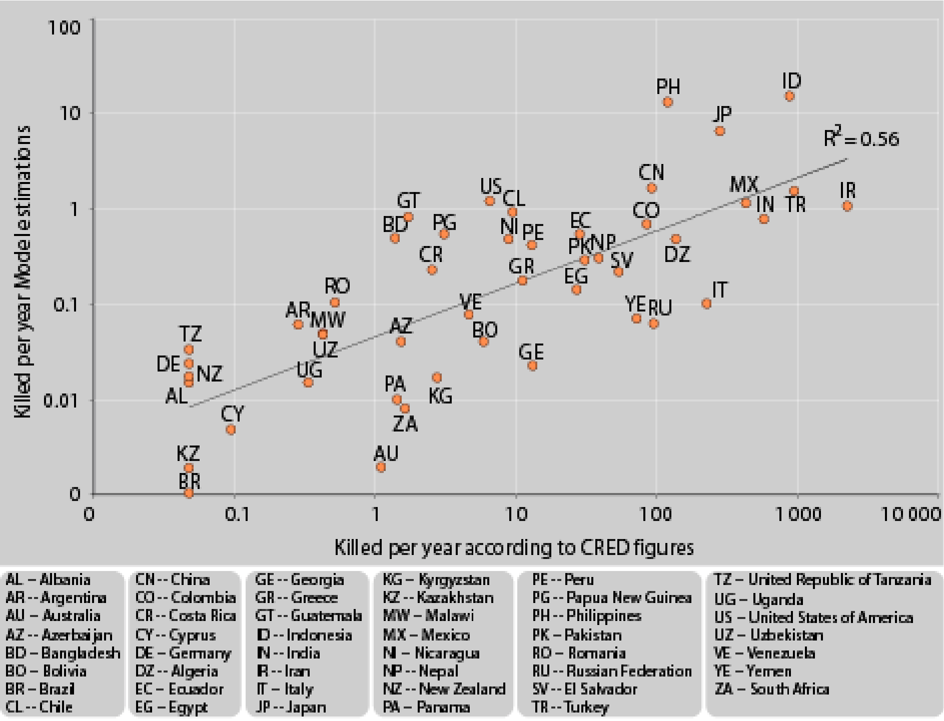
K is the number of killed from earthquakes
PhExp is the physical exposure to earthquakes
U_g *is the rate of urban growth (rates do not request transformation as it is already a cumulative value)*

Table 10: Exponent and p-value for earthquake multiple regression

48 countries	B	p- value (8)
Intercept	-16.22	0.000000
PhExp	1.26	0.000000
Ug	12.27	0.047686

R= 0.75, R^2= 0.56, adjusted R^2= 0.54

Figure 6: Scatter plot of the observed number of people killed by earthquakes (CRED figures) and the model predictions



The variables retained by the regression include the physical exposure and the rate of urban growth. The part of explained variance is smaller than for flood or cyclones (R2=0.544), however considering the small length of time taken into account (21 years as compared to earthquakes long return period), the analysis delineates a reasonably good relation. The physical exposure is of similar relevance than for previous cases, relevant p-value. The urban growth was expected to be selected as indicators. A high rate of population moving into a city is usually synonym of low quality urban planning and building standard. The urban growth is also highly negatively correlated with GDP and HDI. Thus, similar correlation (but slightly inferior) could have been derived using HDI or GDP.

4.7. Drought

Statistical model

The regression analysis was performed using the 6 different exposure datasets derived from IRI drought maps (see p. 12). In general, the models based on 3 month thresholds give better results. The dataset based on a drought threshold set at 3 months at 50% below the median precipitation 1979-2001 was finally selected as the exposure data.

The multiple regression was based on 15 countries and the best fit regression line follows the following Equation 14:

Equation 14: Multiple logarithmic regression model for drought

$$\ln(K) = 1.26\ln(PhExp3\_50) - 7.58\ln(WAT_{TOT}) + 14.4$$

Where:

K is the number of killed from droughts

PhExp3_50 is the number of people exposed per year to droughts ; a drought is defined as a period of at least three months less or equal to 50% of the average precipitation level (IRI, CIESIN/IFPRI/WRI)

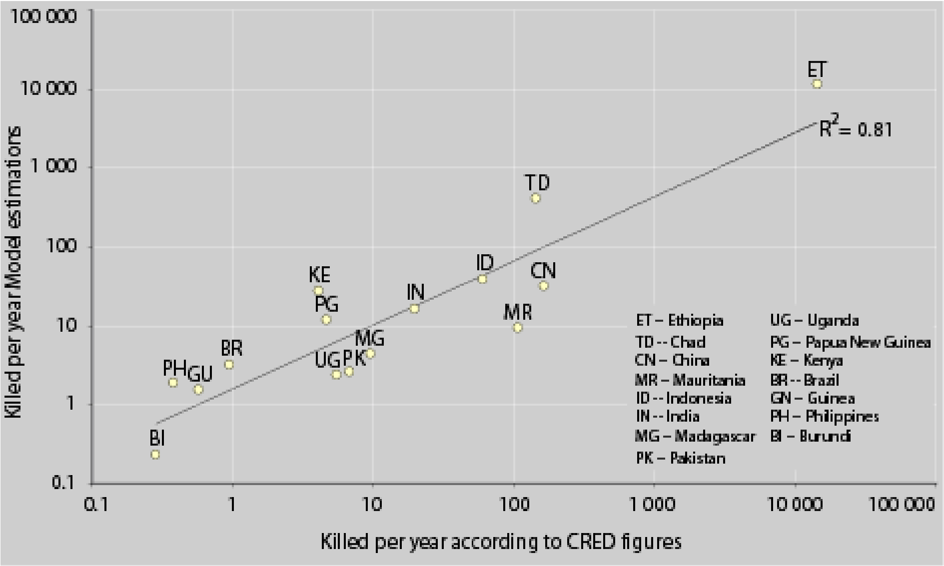
WAT_{TOT} *is the percentage of population with access to improved water supply (WHO/UNICEF)*

Table 11: Exponent and p-value for drought multiple regression

Predictor	Coef	SE Coef	T	p-value[9]
Constant	14,390	3,411	4,22	0.001
Phexp3_5	1,2622	0,2268	5,57	0,000
$WAT_{TOT}{}^{(\ln)}$	-7,578	1,077	-7,03	0,000

S = 1,345   R-Sq = 0.812 R-Sq(adj) = 0.78

Figure 7: Scatter plot of the observed number of people killed by droughts (CRED figures) and the model predictions



Rejected countries : Swaziland and Somalia (WAT_{TOT} value inexistent), North Korea (reported WAT_{TOT} of 100% is highly doubtful), Sudan and Mozambique (eccentric values, suggesting other explanation for casualties)

---

[9] In broad terms, a p-value smaller than 0.05, shows the significance of the selected indicator, however this should not be used blindly.

The small p-values observed are suggesting a relevant selection of the indicators among the list of available datasets. It is to be noted that the high coefficient for WAT_{TOT} (-7.578) **denotes a strong sensibility to the quality of the data**. This implies that even a change of 1% in the percentage of total access to water will induce significant change in the results, especially for small values (where small changes have bigger influence in proportion).

If the model allows the selection of socio-economic parameters indicating the vulnerability of the population, this model cannot be used for predictive purpose. Some inconsistencies were depicted in the data that require verification.

The two indicators selected through the statistical analysis are not surprising. Physical exposure summarise the frequency of hazard and the element at risk (here the population) while the percentage of population with access to improved water supply is an obvious indicator of vulnerability to drought. This could obviously be derived through common sense (population with good water supply less suffer from drought is not surprising). However, the fact that it was selected and that a strong correlation could be established ($R^2 = 0.81$) between independent datasets such as level of precipitation, population, casualties from drought and access to water, assess the solidity of the method as well as the reliability of these datasets for such scale.

Given the level of precision of the data and the compulsory simplification of the drought model, such match is much higher than originally expected.

The Figure 7 shows the distribution (on a logarithmic scale) of expected casualties from drought and as predicted from the model. A clear regression can be drawn. It is true that if Ethiopia is removed the correlation will fall to a mere (R2 = 0.6), however the off set and the slope of the regression line do not change significantly thus assessing the robustness of the model.

As far as 1.26 is close to 1, the number of killed people grows proportionally to physical exposure. Also the number of killed people is decreasing as the percentage of population when improved water supply is growing. This latter variable should be seen as an indicator of the level of development of the country as it was correlated to other development variables, such as under five mortality rate ($U5_{MORT}$, Pearson correlation r = -0.64) and Human Development Index (HDI, r = 0.65).

There were some concerns about some features reported in CRED. Some countries with large physical exposure did not reported any killed (United States of America, Viet Nam, Nigeria, Mexico, Bangladesh, Iran, Iraq, Colombia, Thailand, Sri Lanka, Jordan, Ecuador, etc.). This could be for different reason: either the vulnerability is null (or extremely low) e.g. USA, Australia or the number of reported killed from food insecurity is placed under conflict (e.g. Iraq, Angola,…) for other countries further inquiries might be necessary.

5 MULTIPLE RISK INTEGRATION

5.1. *Methods*

How to compare countries and disasters

One of the main difficulties of this research might consist in the comparison between disaster types. Indeed, how to compare an earthquake with a drought? Rapid on-set, not predictable, with sudden and significant effect for the earthquake, slow on-set, fairly predictable, with fuzzy boundaries difficult to determine in both spatial and temporal way for a drought inducing long term effect. These considerations prevent the use of the proxy for multiple hazard vulnerability as the physical exposure cannot be compared.

The idea was then to compare what is comparable: the casualties. A model for multiple risks integration was made by adding expected casualties. In order to reduce the number of countries with no data, preventing them to be modelled, the value “no data” for countries without significant exposure was replaced by zero risk of casualties. Thus reducing the number of country removed from the multiple risk model, because some data are missing for modelling a type of hazard that do not affect these countries. By performing these boolean conditions it was possible to incorporate 210 countries (out of 249) into the DRI.

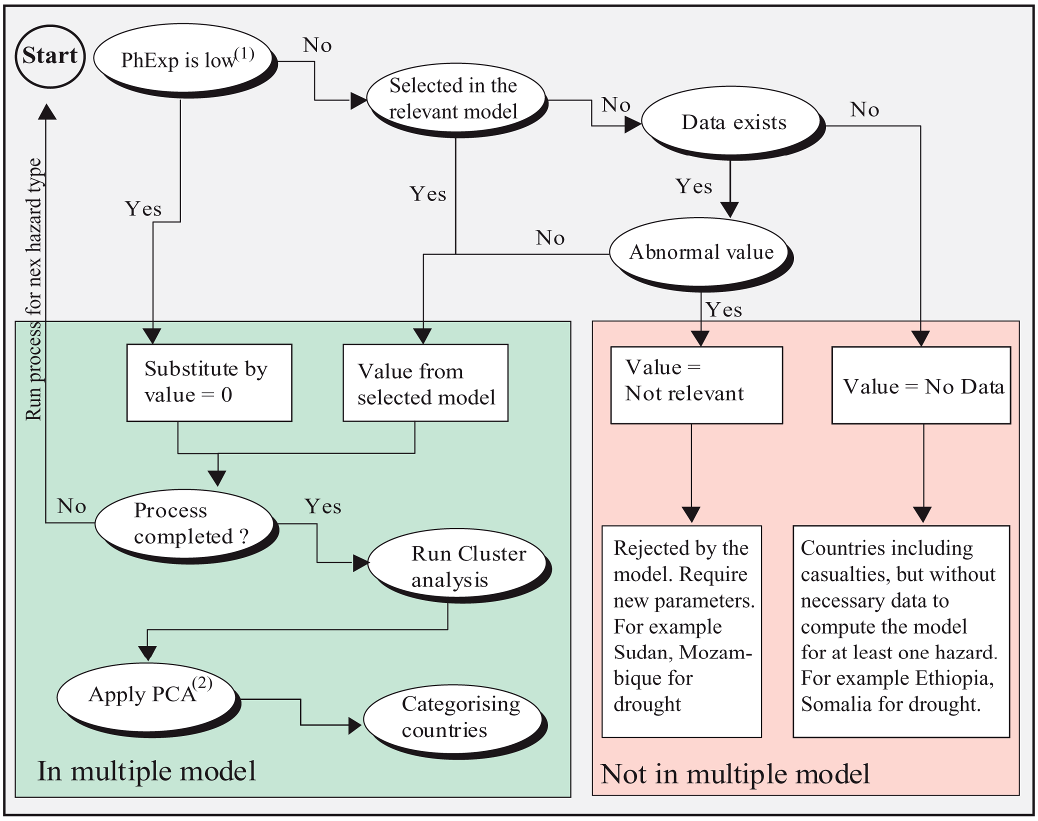
A country was considered as not or marginally affected if the two following conditions where gathered: a physical exposure smaller than 2% of the national population AND an affected population smaller than 1000 per year.

The remaining 39 countries are cases with exposure and for a large majority with recorded casualties (for 37 of them). This list of countries with casualties is therefore important to complement the model. It identifies countries where improvement in data collection is needed. Seven different cases were identified, differentiating countries marginally affected by a specific hazard, countries affected but without data, countries with situation that cannot be explained by the model (e.g. Mozambique, Sudan for drought, where conflict is playing a more significant role than physical drought on food security).

Once the different cases were identified, it was possible to run a boolean process in order to allocate the relevant values depending on the cases. The Figure 8 illustrates the different steps for incorporating the values into a multiple risk index. Once the values for the countries were computed, three different products were available:

- A table of values for the countries that include the data for relevant hazards or countries without data but marginally affected (210 countries)
- A list of countries with missing data (countries with reported casualties but without appropriate data).
- A list of countries where the model could not be applied (indicators do not capture the situation in these countries, case of countries not explained by the model, rejected during the analysis because the indicators are not relevant to the situation).

Figure 8 Process for multiple risk integration



(1) Physical Exposure is considered as marginal if smaller than 1000 per year

(2) PCA: Principal Component Analysis, used to combine killed per year and killed per population in one component

The total risk computation

The multiple risk, in the case where the condition allows its computation, is computed using the succession of formulae as described in Equation 15:

Equation 15. Computation of the multiple Risk by summing the casualties as modelled for risk for cyclone, flood, earthquake & drought

$$K_{cyclone}(PhExp_{Cyclones}^{0.63} \cdot \overline{Pal}^{0.66} \cdot \overline{HDI}^{-2.03} \cdot e^{-15.86}) +$$

$$K_{floods}(PhExp_{floods}^{0.78} \cdot GDP_{cap}^{-0.45} \cdot D^{-0.15} \cdot e^{-5.22}) +$$

$$K_{earthquakes}(PhExp_{earthquakes}^{1.26} \cdot U_g^{12.27} \cdot e^{-16.27}) +$$

$$K_{drought}(PhExp3\_50^{1.26} \cdot WAT_{TOT}^{-7.58} \cdot e^{14.4})$$

Where:

e is the Euler constant (=2.718...)
PhExp is the physical exposure of selected hazard
HDI is the Human Development Index
GDP_{cap} is the Gross Domestic Product per capita at purchasing power parity
D is the local density (density of population in the flooded area)
U_g is the Urban growth (computed over 3 years period)
Wat_{tot} is the access to safe drinking water.

Between each addition, the whole process described in Figure 8 needs to be run in order to identify the cases where the value is replaced by zero, calculated from the selected hazard model or placed in the “not-relevant” or “no data” lists see Table 13 and Table 14.

To classify the countries into the five categories, a cluster analysis minimising the intra-class distance and maximising the inter-classes (K-means clustering method) was performed. The choice to define 5 classes is discussed in page 15.

In order to take both risk indicators (killed and killed per inhabitant) into account a Principal Component Analysis (PCA) was performed to combine the two informations into one. Then a distinction was made between countries smaller than 30’000 km^2 and with population density higher than 100 inhabitants per km^2 following UNDP/BCPR requirements.

5.2. Results

List of countries with observed casualties and risk modelled

Modelled countries without reported casualties

The Disaster Risk Index (DRI) was computed for 210 countries, this includes 14 countries where no reported casualties were reported in the last two decades (from CRED) as seen in list in Table 12.

Table 12. List of the 14 countries with DRI computed but without observed casualties

Barbados, Croatia, Eritrea, Gabon, Guyana, Iceland, Luxembourg, Namibia, Slovenia, Sweden, Syrian Arab Republic, The former Yugoslav Republic of Macedonia, Turkmenistan, Zambia.

No data, abnormal values and specific cases

Following the Principal Component Analysis transformation, inferior and superior thresholds could be identified. This was performed on both observed and modelled casualties. In the cases of 14 countries the model was produced even though no recorded casualties were recorded by CRED in the last two decades. On the other hand 37 countries where casualties were recorded, could not be modelled either because of lack of data or because they did not fit with the model (specific socio-economical contexts).

Table 13. List of the 37 countries with recorded casualties in CRED but not modelled

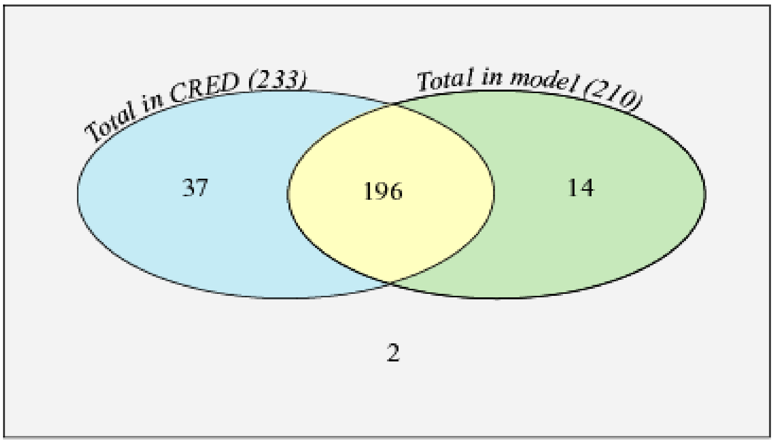
Afghanistan, Azerbaijan, Cuba, Democratic People's Republic of Korea, Democratic Republic of the Congo, Djibouti, Dominica, France, Greece, Liberia, Malaysia, Montserrat, Myanmar, New Caledonia, Portugal, Solomon Islands, Somalia, Spain, Sudan, Swaziland, Taiwan, Tajikistan, Vanuatu, Yugoslavia, Antigua and Barbuda, Armenia, Guadeloupe, Guam, Israel, Martinique, Micronesia (Federated States of), Netherlands Antilles, Puerto Rico, Reunion, Saint Kitts and Nevis, Saint Lucia, United States Virgin Islands.

Table 14. List of the two countries absent of both CRED and Model

Anguilla and Bosnia-Herzegovina

The Table 15 indicates when the comparisons were possible (196) out of 249.

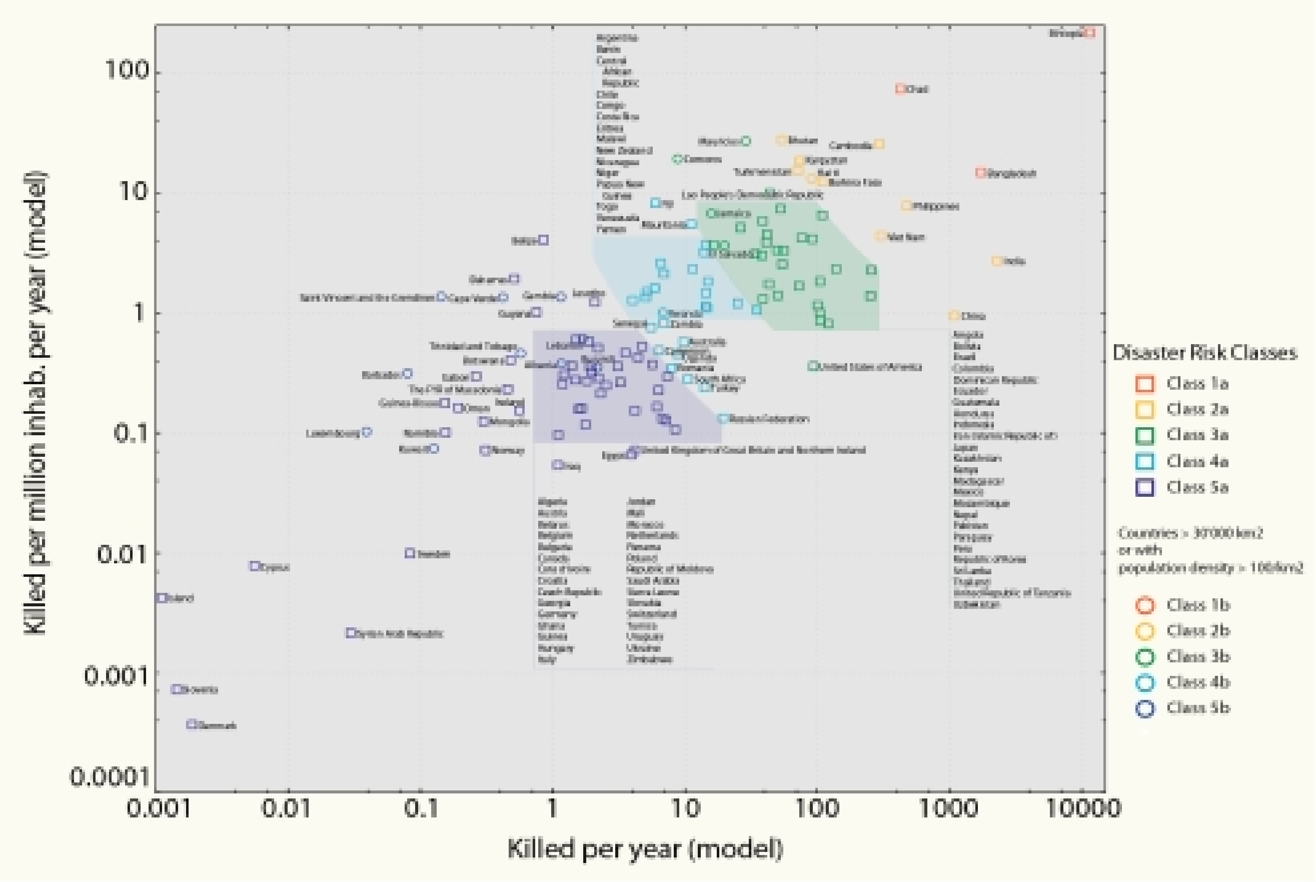
Table 15. Result from the classification



Modeled risk

The Figure 9 depicts the distribution of killed and killed per population for the model with categories in different colours. One can notice the diagonal cut between the categories, meaning that both killed and killed by population have their respective role. A small country with lower number of casualties but high in proportion of its country being in the same categories as large countries with higher number of casualties but with lower percentage.

Figure 9 Scatter plot of expected casualties from multiple model (windstorm, drought, earthquake, flood)



6 FINAL COMMENTS AND RECOMMENDATIONS

6.1. General comments

The purposes of the DRI was to identify whether global data sets could be used for identifying population living in exposed areas and demonstrate the link between socio-economical parameters and vulnerability. The level of correlation achieved delineates that both physical exposure and variables tested are significant and could be used for categories of risk identification. The correlation found is even much higher than initially thought. This is particularly true for climatic events. Smaller correlation was achieved with earthquakes, this is believed to be due to the long returning period, which is less compatible with the 21 year period observed.

Except for drought, in the three other hazards the role of physical exposure appeared to be the most significant, however socio-economical parameters such as GDPcap, HDI, urban growth, percentage of arable land and local population density, were also selected depending on the type of hazards. The sign of the exponents was always following what the common sense and specialists would have recommended with the notable exception of local density for flood. In the case of drought, the socio-economical context is playing the stronger role and is highly sensitive to the data quality (in this case access to safe water).

All in all, the method used in this statistical analysis proved to be appropriate and allows the identification of the parameters leading to a higher risk and vulnerability. **Such a model should not be used as a predictive model**. Firstly because the level of data precision can not reach such precision. Secondly, because a significant discrepancy of losses between two events in the same country can be found. This shows the variability due to micro spatio-temporal context. The risk maps provided in this research are not to be confused with danger maps. At a local scale predictive model can and should be made allowing better urban planning and improved evaluation of risk. Maps at global scale are only produced in the aim of identifying the countries with the highest needs corresponding to the request from UNDP.

Extraordinary events – also called century disasters – do not follow the normal trend. Hurricane Mitch (Central America, 1998), or the flood causing the landslide in Caracas (Venezuela, 1999), earthquake in Armenia (1988), were clearly off the regression line. This is due to the abnormal intensity of such events which do not correspond with the average intensity. These events are (hopefully) too rare to be approached by a two decade period. Incorporating the intensity can only be done on an event per event approach. When entering an average intensity, the numerous low intensity events are biasing the average and finally the intensity was rejected as explicative variable.

6.2. *Recommendations*

Use of the model

The very high sensitivity to the data of the drought model confirms that the equation should not be used for predictive applications.

Socio-economical variables

Results delineate that global data sets can still be improved both in terms of precision and completeness, however they are already allowing the comparison of countries. Other indicators such as corruption index (transparency) or political indicators would be interesting to test in the model, when all the countries will be available. Efforts on compilation are still needed. Tremendous amount of work was involved (by UNEP/GRID-Geneva GEO team) to verify and complete the data.

Floods

The geophysical data can also be improved. The watersheds for flood physical exposure is based on a 1 km cell resolution for elevation. A new global data set on elevation from radar measures taken from the NASA shuttle is expected in 2004. It consists in a 30m resolution grid for the US and 90m resolution for the global coverage. This would allow the refining of the estimated area flooded. This would be especially welcome for the central Asian countries where the watersheds taken were of very poor resolution. Collaboration with Dartmouth Flood Observatory would be an asset.

Earthquakes

If information on soil (quaternary rocks) and faults orientations can be implemented, it would then be possible to compute intensity using modified Mercali scale and thus with much higher precision on the area affected. Alternatively a method for deriving frequency based on the Global Seismic Hazard Map from the GSHAP [Giardini 1999] could be also used.

Cyclones

Once the data from North Indian Ocean are available, a vector approach should be applied using the PreView Global Cyclone Asymetric Windspeed Profile model developed by UNEP/GRID-Geneva. This method computes the areas affected based on central pressure and sustainable winds.

Drought

It might be interesting to test other precipitation data sets with higher spatial resolution, although the resolution did not seem to be causing so much problem. The use of geo-climatic zones might be useful in order to take into account the usual climate of a specific area. Indeed a drop of 50% precipitation might not have the same consequence on a humid climate as compared to a semi-arid area. The use of the Global Humidity Index (from UNEP/GRID UEA/CRU) might help in differentiate these zones. Measuring food insecurity (using e.g. information on conflict and political status) would be also a significant improvement as compared to physical drought.

The case of small islands and archipelagos

Small islands and archipelagos are causing problems. In some cases they were too small to be considered by the GIS automated algorithms. This was typically the case for the population. The raster information layer for the population can not be used to extract the population of small islands. For single island countries, the problem might be overcome by using the population of the country, but for the others this was not possible. Indeed, when superimposing cyclone tracks on top of archipelago, the population is needed for each island. A manual correction is needed, but could not be performed due to the time frame of the study. The compilation of socio-economical parameters was also not complete for the islands. This could probably be improved by collaborating with SOPAC (Fiji).

For all these reasons, the case of small islands and archipelagos would need a separate study and intuitively, the vulnerability for isolated countries might be different than other connected countries.

The issue of indicators

To what extend the casualties are proportional to the significance of total losses (including losses of livelihood)? In the case of earthquakes where no early warning exists, this might be a good proxy, although it will highly depend if the earthquake epicentre is located in rural or urban areas. For flood, however, the casualties are usually much smaller in relation to losses of houses, infrastructures and crops. The ideal would be to have access to records of livelihood losses in order to calibrate the severity of a hazard type as compared to another (while considering the magnitude). Without such data no scientific method can be implemented. Several options can be explored:

1. Organising a workshop with experts and asking them for a comparative severity of hazard.
2. Ask relief and aid organisations what is their average budget for recovery and mitigation for each type of hazard.

This is a significant issue that could not be taken into account during this first research, but needs to be incorporated in the future to improve the index quality. This would also allow to normalise physical exposure. So far it is impossible to state how bad it is to be exposed to an earthquake as compared to a flood, except eventually in terms of surviving chances.

Extending to other hazards

Volcanoes

An attempt was made to model volcanic eruption. It revealed the impossibility of modelling physical exposure and vulnerability to volcanic eruption at global scale. If danger maps can be derived for a single volcano at local scale, the variability of the volcanic manifestations is far too complex to be generalised, it ranges from lahars (linked with precipitation level, seismicity, topography, soils characteristics,…), tephra falls (depending from wind direction and strenght), phreatomagmatic eruption,… . However, numerous experts are working on surveying these activities and each volcano is well described.

1. Cautions must be taken on the use of small areas (smaller than 100 pixels) for the calculation of population exposures, due to the precision of population data.
2. The risk results in the inversion of the frequency role: the less frequent the more dangerous. Indeed the larger impacts are resulting from volcanoes with low historical activity. The habitants of these areas are living in a false confidence.

> Indeed, the volcano *Ruiz* (Colombia, 1985) which melting of its summit icecap provoked the South America's deadliest eruption (22’800 killed) had a large eruption in 1595 and smaller in 1828, 1829 and then has only been reported smoking in 1831, 1833 [Herd, 1982]. Same situation with *Pinatubo* (Philippines, 1991) “*Prior to 1991 Pinatubo volcano was a relatively unknown, heavily forested lava dome complex with no records of historical eruptions*.” [Global Volcanism Program, Smithsonian Institution, 2001]. And similar with the lake Monoun (Cameroun, 1986) “*No previous eruptions are known from Lake Monoun” [*Le Marechal, 1975a]. These disasters demonstrated that low frequency does not mean low risk, this is completely different as for the other hazards.

Data requested are probably existing. Finer resolution for elevation is a must, for showing shape and relief of volcanoes, computing slopes and lahars danger. Remote sensing analysis for local assessment of danger and population distribution would also be requested. Numerous maps of volcanoes can be found at http://www.nmnh.si.edu/gvp/volcano/index.htm.

Tsunamis and Landslides

Some countries are not well represented by the model, because they are affected by hazards which were not of global significance. This is the case of Papua New Guinea and Ecuador, which are affected by tsunamis (respectively 67.8 and 14.3% of national casualties); landslides are also causing significant impact in Indonesia (13,88%), Peru (33%) and Ecuador (10.2%). As a result, the global risk is under evaluated for these countries.

Epidemics

This is more a health angle and should probably be taken care of by the World Health Organisation (WHO). However, the appropriate sanitation, access to safe water, number of physicians per inhabitants and other health infrastructure are also significant parameters of development. Data on epidemics are now starting to be available. Epidemics is representing a significant amount of casualties and AIDS is definitely impacting developing societies especially (but not only) in Africa.

Conflicts

The case of conflicts although much more politically difficult to approach is probably also highly correlated to human vulnerability. Results from a statistical analysis would be extremely interesting.

Last word

These results delineate the relation between level of development and low casualties from these four types of hazards. Stating that there is a relation can be understood both ways: lower development may lead to higher casualties, but high hazard occurrence may also lead to lower economical development as it destroys infrastructures and crops as well as it scares the investors away. If higher impacts from natural hazards in developing countries were depicted, the message should not be perceived as “developed countries should be taken as models”. Other figures such as death from suicides, drug abuses or excess of fat food, are also leading to numerous casualties and are highly and positively correlated with HDI!

This research underlines the usefulness of continuing the improvement of data collection for a better identification of populations at risk. This is, however, not a final result as such. Final results will be achieved when proper risk reduction measures will be implemented leading to an observed decrease of casualties.

7 REFERENCES

Articles and books

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).

Burton, I., Kates, R.W. and White, G.F. 1993: *The Environment as Hazard, Second Edition.* New York/London: Guilford Press, 290 pp. [Pp. 31-47]

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

Coburn, A.W., Spence, R.J.S. and Pomonis, A. 1991: *Vulnerability and Risk Assessment.* UNDP Disaster Management Training Program 57 pp.

Demuth, S., Stahl, K., 2001, *Assessment of the Regional Impact of Droughts in Europe (ARIDE), Final Report*, Institute of Hydrology, University of Freiburg, Freiburg, Germany.

Giardini, D., 1999, *Annali di Geofisica, the global seismic hazard assessment program (GSHAP) 1992-1999*, Instituto Nazionale di Geofisica, Volume 42, N. 6, December 1999, Roma, Italy.

Herd, D.G., 1982, *Glacial and volcanic geology of the Ruiz-Tolima Volcanic Complex*, Cordillera Central, Colombia: Publicaciones Geológicas Especiales del INGEOMINAS, no. 8, 48 p.

Holland, G. J., 1980, An analytic model of the wind and pressure profiles in hurricanes. *Monthly Weather Review*, 108, 1212 1218.

Le Marechal, A., 1975a, *Carte geologique de l'ouest du Cameroun et de l'Adamaoua*, 1:1,000,000; ORSTOM.

Newhall, C. G., Self, S., 1982, The volcanic exposivity index (VEI): an estimate of explosive magniude for historical volcanism*. Jour. Geophys. Res.* (Oceans & Atmospheres), 87:1231-8.

Peduzzi, P., Dao, H., Herold, C., Frédéric Mouton (2003), *Global Risk And Vulnerability Index Trends per Year (GRAVITY), Phase IIIa: Drought analysis*, scientific report UNDP/BCPR, Geneva, Switzerland.

Peduzzi, P., Dao, H., Herold, C., (2002), *Global Risk And Vulnerability Index Trends per Year (GRAVITY), Phase II: Development, analysis and results*, scientific report UNDP/BCPR, Geneva, Switzerland.

Peduzzi, P., Dao, H., Herold, C., and Rochette, D., (2001), *Feasibility Study Report On Global Risk And Vulnerability Index Trends per Year (GRAVITY)*, scientific report UNDP/BCPR, Geneva, Switzerland.

Schloemer, R. W., 1954, *Analysis and synthesis of hurican wind patterns over Lake Okehoee, Fl.* Hydromet Rep. 31, 49 pp. [Govt. Printing Office, No. C30.70:31].

Simkin, T. and Siebert, L.1994, *Volcanoes of the World*, (Geoscience Press, Washington DC).

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

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

UNDRO (United Nations Disater Relief Coordinator), 1979: Natural Disasters and Vulnerability Analysis in *Report of Expert Group Meeting* (9-12 July 1979). Geneva: UNDRO. 49 pp.

UNEP, 2002, GEO: Global Environment Outlook 3. Past, present and future perspectives, UNEP.

References on Internet

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

Babin, S. and Sterner, R., 2001, Atlantic Hurricane Track Maps & Images, http://fermi.jhuapl.edu/hurr/index.html

Birdwell K.R., Daniels, R.C., 1991, A Global Geographic Information System Data Base of Storm Occurrences and Other Climatic Phenomena Affecting Coastal Zones (1991) http://cdiac.esd.ornl.gov/ndps/ndp035.html

CIESIN, IFPRI, WRI, 2000, Gridded Population of the World (GPW), Version 2, http://sedac.ciesin.org/plue/gpw/

Council of the National Seismic System (as of 2002), Earthquake Catalog, http://quake.geo.berkeley.edu/cnss/

Deichmann, Uwe, 1996, GNV197 - Human Population and Administrative Boundaries Database for Asia, UNEP/GRID-Geneva, http://www.grid.unep.ch/data/grid/gnv197.php

Donlin C. and Fitzgibbon, T., 2001, Geopubs - Online Geologic Publications of the Western United States, USGS: http://geopubs.wr.usgs.gov/docs/wrgis/fact

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

Global Volcanism Program, National Museum of Natural History, E-421, Smithsonian Institution, Washington DC 20560-0119, http://www.nmnh.si.edu/gvp/index.htm

Global Volcanism Program, Volcanic Activity Reports, Pinatubo, Index and All Reports, National Museum of Natural History, Smithsonian Institution, Washington: http://www.nmnh.si.edu/gvp/volcano/region07/luzon/pinatubo/var.htm#1605

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

IRI/Columbia university, National Centers for Environmental Prediction Climate Prediction Center (as of 2002), CPC Merged Analysis of Precipitation (CMAP) monthly gridded precipitation, http://iridl.ldeo.columbia.edu/

ISRIC, UNEP, 1990, Global Assessment of Human Induced Soil Degradation (GLASOD), http://www.grid.unep.ch/data/grid/gnv18.php

Landsea Christopher W., 2000,NOAA/AOML, FAQ: Hurricanes, Typhoons, and Tropical Cyclones. http://www.aoml.noaa.gov/hrd/tcfaq/tcfaqA.html#A1

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

Peduzzi, P., 2000, Insight of Common Key Indicators for Global Vulnerability Mapping, presentation for the Expert Meeting on Vulnerability and Risk Analysis and Indexing, Geneva 11-12 September 2000, UNEP/DEWA/GRID-Geneva: http://www.grid.unep.ch/activities/earlywarning/preview/appl/reports/reports.htm

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

Topinka, L., 2001, Cascades Volcano Observatory, USGS, Vancouver, Washington, USA: http://vulcan.wr.usgs.gov/Volcanoes/

Transparency international, 2001, Global Corruption Report 2001, http://www.transparency.org/

U.S. Geological Survey, 1997, HYDRO1k Elevation Derivative Database, http://edcdaac.usgs.gov/gtopo30/hydro/

UNDP, 2002, Human Development Indicators, http://www.undp.org/

UNEP,CGIAR,NCGIA, 1996, Human Population and Administrative Boundaries Database for Asia, http://www.grid.unep.ch/data/grid/human.php

UNEP/GRID (as of 2002), GEO-3 Data portal, Http://geodata.grid.unep.ch/

VHP., 2000, Volcano Hazards Program, Strategy for reducing volcanic risk, U.S. Department of the Interior, U.S. Geological Survey, Menlo Park, California, USA: http://volcanoes.usgs.gov/About/What/Assess/

William P. Leeman W.P., 1999, Volcanism & Volcanic Hazards Summary of basic terms and concepts, Rice University: http://www.ruf.rice.edu/~leeman/volcanic_hazards.html