**Natural Hazards
and Earth
System Sciences**

Assessing global exposure and vulnerability towards natural hazards: the Disaster Risk Index

P. Peduzzi[1,2], H. Dao[3], C. Herold[1], and F. Mouton[4]

[1]UNEP/GRID-Europe, Geneva, Switzerland
[2]Institute of Geomatics and Analysis of Risk (IGAR), University of Lausanne, Lausanne, Switzerland
[3]Department of Geography, University of Geneva, Geneva, Switzerland
[4]UFR de Mathématiques, Université J. Fourier, Grenoble, France

Abstract. This paper presents a model of factors influencing levels of human losses from natural hazards at the global scale, for the period 1980–2000. This model was designed for the United Nations Development Programme as a building stone of the Disaster Risk Index (DRI), which aims at monitoring the evolution of risk. Assessing what countries are most at risk requires considering various types of hazards, such as droughts, floods, cyclones and earthquakes. Before assessing risk, these four hazards were modelled using GIS and overlaid with a model of population distribution in order to extract human exposure. Human vulnerability was measured by crossing exposure with selected socio-economic parameters. The model evaluates to what extent observed past losses are related to population exposure and vulnerability. Results reveal that human vulnerability is mostly linked with country development level and environmental quality. A classification of countries is provided, as well as recommendations on data improvement for future use of the model.

1 Introduction

According to available global statistics, least developed countries represent 11% of the population exposed to hazards but account for 53% of casualties (Peduzzi et al., 2002). On the other hand, the most developed countries represent 15% of human exposure to hazards, but account only for 1.8% of all victims. Obviously, similar exposures with contrasting levels of development lead to drastically different tolls of casualties. These are general figures, however, in order to better understand what development parameters are associated with risk, each exposure to specific hazard types should be analysed separately.

This paper presents the methodology and the results of the Disaster Risk Index (DRI), the central component of the report "Reducing Disaster Risk" by the United Nations Development Programme (UNDP/BCPR, 2004). The mandate from UNDP was to analyse potential links between vulnerability to natural hazards and levels of development. The DRI is the first model providing a statistical evidence of such links at the global scale. By setting reference risk values for the period 1980–2000, this model will be the basis for comparisons with subsequent calculations of the DRI in the 21st century.

Since the publication of this report, several other global and regional efforts have been published. The World Bank/University of Columbia published a report (Dilley et al., 2005) including numerous hazard- exposure- and risk maps, also using similar datasets. This study placed more emphasis on the effect of multiple hazards exposure. Above-mentioned studies did not try to model and address vulnerability by grouping past losses per exposed by countries and territories (thereafter referred to as countries+) of similar levels of economic development. At the other extreme, a report also published by the Inter-American Development Bank (Cardona, 2005) in 2005 proposed different sets of complex indicators, e.g. they compared the likely economic loss attributed to a major disaster in a given time period with the economic coping capacity of the country, resulting in an indicator known as the Disaster Deficit Index (DDI). The DDI can therefore be considered as an indicator of a country's economic vulnerability to disaster. Unfortunately, at present the indicator has only been applied in Latin America and the Caribbean, and therefore it is impossible to identify global trends. There is a need for a global index for comparing countries, including an identification of human vulnerability,

Correspondence to: P. Peduzzi
(pascal.peduzzi@grid.unep.ch)

which can be used by aid organisations and governments. Our first version of the DRI was published as on-going work (UNDP/BCPR, 2004), it included several gaps and recommendations that we try to address in this present paper.

There are different challenges when comparing risk levels for different countries, e.g. how to compare large countries with small ones, or how to compare countries affected by earthquakes and those affected by droughts? Because of the specific nature of each hazard type (rapidity of onset, spatial extent and destruction potential), exposures to different hazard types cannot be compared. Being affected by drought differs drastically from being exposed to earthquakes. In the first case, infrastructures generally do not suffer, the impact is slow and gradual, but the duration is long, while the inverse is true for earthquakes. Complexity is higher than considered here as primary hazards often unfold into different secondary hazards (e.g. tropical cyclones triggers storm surges leading to coastal flooding, tempestuous rains and winds leading to landslides). However, this is a level of simplification that has to be accepted once dealing with global risk assessments. To overcome part of the difficulties associated with different types of exposures, the model is based on hazard-specific risk models (cyclones, droughts, earthquakes and floods), which are further combined in a multiple DRI allowing a classification of countries+.

The model is built on both available and newly created global datasets. Exposure, vulnerability and risk have been estimated by means of statistical and Geographical Information Systems (GIS) methodologies which are presented in this article.

1.1 Defining and measuring risk

In this research, the term risk follows the definition by the Office of the United Nations Disaster Relief Co-ordinator (UNDRO) and "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" (Cardona, 2005; Burton, 1978).

There are different sorts of losses from natural hazards: human, economic, cultural, etc. However, this study concentrates on life losses for two main reasons. First, the number of killed people is the most reliable and least subjective figure that can be found in the Emergency Disasters Data Base (EM-DAT, Centre for Research on the Epidemiology of Disasters, http://www.em-dat.net/), the only publicly available global database on human impacts from hazardous events. By comparison, the definition and estimation of other variables like "homeless", "affected" and "total affected" are not reliable and depend largely on gross evaluations. Another reason was the difficulty of using economic data since EM-DAT only records events with estimated losses above 100 000 US$, ignoring many smaller events affecting developing countries. Yet, the socio-economic impacts of a

100 000 US$ loss is not the same when considering countries like USA and Bangladesh, thus inducing a statistical bias since most of the records including economical losses concern developed countries. It is to be noted that EM-DAT includes only medium to large-scale disaster events. Thus, disasters with less than 10 killed are not included.

However, considering the number of killed people does not solve the problem of comparison between countries+. If the raw number of killed is taken, more populated countries+ will always be on the top of the list (e.g. China, India), whereas several countries+ having in total an equivalent population would not be well represented. If the percentage of population killed is used, then the reverse problem appears: small islands and less populated countries+ will always be ranked first and the equity of one person killed is no longer ensured. In order to enable relevant comparisons between hazards and countries+, a risk indicator was computed combining both the total number and the percentage of killed people (Dao and Peduzzi, 2003).

1.2 Choice of hazard types and time period

The study focussed on droughts, earthquakes, tropical cyclones and floods, the four hazards accounting for 94% of casualties reported for the period 1980–2006 in the EM-DAT database. The period 1980–2006 was chosen for its homogenous level of information quality and completeness.

For identifying the period for which the access to information is comparable worldwide, a ratio between physically recorded earthquakes (magnitude > 5.5) on land and reported earthquakes in EM-DAT was used. The choice of earthquakes as a benchmark was made because this hazard is not suspected of being influenced by climate patterns. The ratio of reported versus physically recorded events is rather low until 1979 (average of 11%) and suddenly increases in 1979 (average of 26%) and then remains steady around this value. From the previous observation, a time span of twenty-one years was chosen (1980–2000), completed by a period used for comparison (2001–2006). For hydro-meteorological hazards, the same time span was chosen; however, for tectonic events, frequencies were computed over a longer period (1965–2006).

An even longer time period would have been relevant for hazards like earthquakes, but reports on casualties were probably not as homogenous before the 1980's and the problem of finding the corresponding vulnerability variables would have arisen. Given the significance of the four major hazards, modelling others would not have had a significant effect on the final classification of countries+, except for few specific countries+ affected mostly by tsunamis, landslides, volcanic eruptions or extreme temperatures (e.g. Equator, Papua New Guinea). The EM-Dat dataset was split into two parts: 1980–2000 was used for the calibration of vulnerability to produce a model. The records 2001–2006 were used for comparison with recent events.

2 Modelling risk

By UN definition (Cardona, 2005), the risk of losses is a function of three components: hazard, element at risk and vulnerability. In the case of risk of human losses, the element at risk is the exposed population. The hazard occurrence refers to the frequency of returning period at a given magnitude, whereas the vulnerability is "the degree of loss to each element should a hazard of a given severity occur" (Blaikie et al., 1994).

A hypothesis was made that risk follows a multiplicative formula as described in the simplified Eq. (1).

$$R = H_{fr} \, \text{Pop Vul} \tag{1}$$

where:
R = number of expected human impacts [killed/year].
H_{fr} = frequency of a given hazard [event/year].
Pop = population living in a given exposed area [exposed population/event].
Vul = vulnerability depending on socio-politico-economical context of this population [non-dimensional number between 0–1].

According to this formula, if there is no hazard, then the risk is null (the same if population or vulnerability is null).

2.1 Identifying physical exposure

The combination of both yearly average frequency of hazards and exposed populations provides the physical exposure and can be computed, depending on cases, using Eqs. (2) or (3).

$$\text{PhExp} = \sum_i^n F \, \text{Pop}_i \tag{2}$$

where:
PhExp = yearly average physical exposure for the spatial unit [exposed population/year].
F annual frequency of a given magnitude event [event/year].
Pop_i = total population living in the spatial unit for each event "i" [exposed population/event].
n = number of events considered

$$\text{PhExp} = \sum \frac{\text{Pop}_i}{Y_n} \tag{3}$$

where:
PhExp = yearly average physical exposure for the spatial unit [exposed population/year].
Pop_i = population living in affected area for each event "i" [exposed population/event].
Y_n = length of time [year].

The frequency and geographical extent of each hazard were modelled and further used for extracting the exposed population (Fig. 1). Equation (1) for risk was then transformed into Eq. (4) for computing the physical exposure:

$$R = \text{PhExp Vul} \tag{4}$$

2.2 Approaching human vulnerability

2.2.1 The use of indicators

The last component, vulnerability, is less easily apprehended. It is a concept to be quantified using indicators. A selection of 32 socio-economical and environmental variables (Supplementary material A: http://www.nat-hazards-earth-syst-sci.net/9/1149/2009/nhess-9-1149-2009-supplement.pdf) was introduced in a database for further statistical analysis.

A correlation study (matrix-plot and correlation-matrix) was performed to ensure that the variables were independent before applying the regression analysis. This was for instance not the case for the highly correlated Human Development Index (HDI) and Gross Domestic Product per capita (at Purchasing Power Parity). In order to keep a valid sample size, a preference was given to variables with the lowest number of missing values.

2.2.2 Parametric model used

A generalisation of the multiplicative approach (Eq. 4) was defined with the following parametric model (Eq. 5):

$$K = C \, (\text{PhExp})^\alpha \, V_1^{\alpha_1} \, V_2^{\alpha_2} \ldots V_p^{\alpha_p} \tag{5}$$

where:
K = number of persons killed by a certain type of hazard.
C = multiplicative constant.
PhExp = physical exposure : population living in exposed areas multiplied by the frequency of occurrence of the hazard.
Vi = socio-economical variables.
αi = exponent of $V i$, which can be negative (for ratio).

Taking the logarithms in Eq. (5) gives Eq. (6):

$$\ln(K) = \ln(C) + \alpha \, \ln(\text{PhExp}) + \alpha_1 \, \ln(V_1) + \alpha_2 \, \ln(V_2) \\ + \ldots + \alpha_p \, \ln(V_p) \tag{6}$$

Significant socio-economical variables V_i (transformed when appropriate, see below) and exponents α_i were determined by means of linear regressions that were carried out for each hazard. The variable K to be estimated was the number of killed people as reported by EM-DAT.

2.2.3 Transformation of variables

Since socio-economical indicators fluctuate across time, a weighted average was computed for each variable:

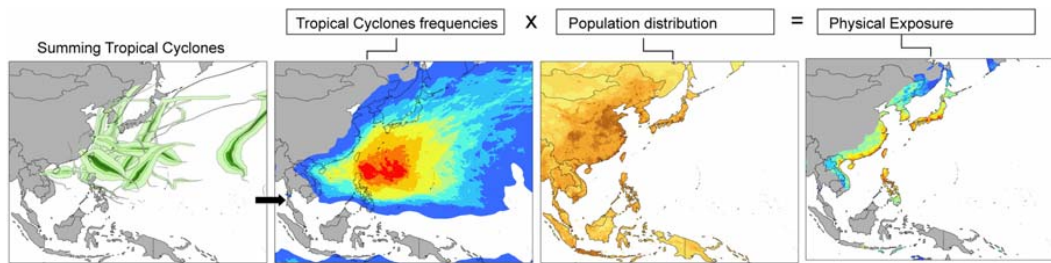$$V' = (V_{1980} K_{1980} V_{1981} K_{1981} + \ldots + V_{2000} K_{2000}) / K_{\text{tot}} \tag{7}$$

where:
V' = weighted average of a given variable.
V_i = value of the variable for the year i.
K_i = number of recorded casualties for the year i.
K_{tot} = total number of recorded casualties for all years.

Fig. 1. Example of physical exposure extraction (tropical cyclones). Once the spatial extents of individual cyclones are modelled, each cell is used to count the average cyclone frequency over the available period. The average frequency is then multiplied by the population identified in each cell (population distribution) in order to obtain the physical exposure. This cell-by-cell physical exposure is further aggregated (summed) at national level.

The result of Eq. (7) is an averaged value that is obtained from yearly values weighted according to the number of casualties in each year. For example, this process avoids taking the Gross Domestic Product of a selected year if the bulk of the victims occurred 10 years before or after (see example in Table 1).

Since the population is also changing through time, this affects the computation of the physical exposure (PhExp). The same formula was applied to the physical exposure.

For the variables with unlimited positive values (e.g. population) the logarithms were computed directly, but for others expressed in percent, a logistic transformation was applied, $V* = V'/(1-V')$, so that their logarithms range between $-\infty$ and $+\infty$. This appeared to be relevant as some of the transformed variables proved to be significant in the final results. For others, no logarithm was needed: for instance the urban growth U_g already behaves in a cumulative way.

3 Calibration of the risk model hazard per hazard

In the regression analysis, physical exposure (PhExp) was considered as an explanatory variable and proved to be statistically significant in all cases detailed below, thus validating the methodology developed for obtaining PhExp.

3.1 Tropical cyclones

Exposed populations to each cyclone were estimated by computing buffers along the cyclone track, where windspeed is greater than a certain threshold [42.5 m/s]. These buffers had to be generated for the study by modifying a wind profile model initially developed by Greg Holland (Holland, 1980). The modification adds the movement of cyclone's centre, leading to asymmetric buffers (Mouton et al., 2002). A global dataset was produced using tracks of tropical cyclones available on the internet from different meteorological centres. Information on latitude/longitude, date, hour, windspeed and central pressure are usually included, although each centre has its own way and units for measuring cyclone characteristics. The PreView Global Cyclones Asymmetric Windspeed Profile dataset developed for this study provides

Table 1. Cyclone casualties and HDI in El Salvador (1980–2000).

Year	K	V (HDI)	$V*K$
1988	28	0.781	21.879
1996	3	0.810	2.429
1996	51	0.810	41.300
1998	8	0.815	6.523
Total	90		72.132
		$V = 72.132/90 = 0.801$	

users with a standardised version, with units converted into the metric system. Using the areas derived from the asymmetric windspeed profiles, it was possible to extract the physical exposure using Eq. (3).

The variables highlighted by the statistical analysis are PhExp, the GDPcap and the percentage of country+ area dedicated to cropland. According to the analysis, the number of killed people is growing with PhExp and decreasing with the GDPcap. The percentage of cropland can also be understood as a proxy of the type of population/habitat, i.e. rural, scarcely distributed population being more vulnerable than urban population. This statistical result is in line with what was expected by consulted experts (IWTC-V, 2002). After a tropical cyclone, an economy relying on the tertiary sector is less affected than one relying on agriculture, the fields having been devastated. These results confirm that poor populations are more vulnerable to tropical cyclones. With a considerable part of variance explained by the regression ($R^2=0.81$), a high degree of confidence in the selected variables (p-values < 0.05) over a sample of 34 countries+ and a residual analysis showing no particularity or abnormality, the model achieved is robust. Notice that although the consequences of hurricane Mitch (in 1998) could easily be depicted, Honduras and Nicaragua were far off the regression line (significantly underestimated) and were not used for the model. This is explained by the incredible difference of intensity between Mitch and other hurricanes. Cuba's success

in risk reduction (i.e. by evacuating population exposed to a coming cyclone) is confirmed by the analysis: observed casualties are so much lower than the expected values, that this country was identified as an outlier. The partial correlation analysis highlights that the physical exposure explains the major part of the casualties, followed by GDPcap and then percentage of country dedicated to cropland.

The plot of observed versus expected values delineates a linear distribution as seen in Fig. 2. The model is the following (see Table 2):

$$\ln(K) = 0.621 \ln(\text{PhExp}_{Cy}) - 0.534 \ln(\text{GDPcap})$$
$$+0.347 \ln(\text{CROPpca}) - 0.487 \tag{8}$$

where:
K = number of estimated killed.
PhExp = physical exposure to tropical cyclones.
GDPcap = transformed value GDP purchasing power parity per capita.
CROPpca = transformed value of the percentage of the country dedicated to Crop land.

3.2 Droughts

Drought is a complex process to model as it is not clear when a drought starts both in spatial and temporal terms. The same deficit in precipitation may not induce similar impacts depending on types of soil, vegetation and agriculture as well as on differences in irrigation infrastructures. Moreover, casualties are not directly induced by physical drought but rather by food insecurity which is not purely a natural hazard as it includes human induced causes (such as conflicts, poor governance, etc.). However, a global approach on risk to human development would not be achieved without drought, as most of Africa is affected mainly by food insecurity.

A first attempt to identify physical drought was developed by Brad Lyon and his team (Dilley et al., 2005) from the International Research Institute for Climate Prediction (IRI), who produced several methods with different thresholds on duration (3 and 6 months) and shortage of precipitation (50%, 25% and 10%) at 2.5° resolution. For this research, their method was re-applied to a 0.5° resolution raster dataset from the Climatic Research Unit (University of East Anglia, Norwich). This proved to significantly improve results as compared to the original IRI data at 2.5° resolution. Secondly, physical exposure was computed on a cell-by-cell basis using Eq. (2) and was further aggregated at the national level. During this research, a calibration using reported casualties identified the best global match with the thresholds set at 50% of precipitation shortage during a period of 3 months.

The indicators identified by the statistical analysis are PhExp, GDPcap and the percentage of arable land. This latter variable was computed in order to take into account the percentage of arable land excluding deserts.

3.2.1 Computing the percentage of arable land (for droughts)

The original figure for percentage of arable land came from the FAO database. It was modified in order to take into account the percentage of arable land excluding deserts.

$$\text{mAL\_pc} = \frac{\text{ALA}}{(\text{TA} - \text{DA})} \tag{9}$$

where:
mAL_pc = modified percentage of arable land.
ALA = arable land area (in km^2).
TA = total area (in km^2).
DA = desert area (in km^2).

The desert areas were identified using the Global Land Cover 2000 dataset. This is to avoid the case of countries+ largely covered by deserts where populations are concentrated on a small portion of the territory.

According to the analysis, the number of killed people from physical drought grows with PhExp, decreases as the GDPcap grows and decreases if the percentage of arable land grows. The interesting point is that, as opposed to the other hazards, the main contribution to casualties is poverty (low GDP) followed by physical exposure and percentage of arable land.

A country with a large portion of arable land is less likely to be totally affected by a drought and might still be able to provide enough food for its inhabitants.

Again, the part of variance explained by the regression ($R^2 = 0.70$) is important and p-values are smaller than 0.05. The plot can be seen on Fig. 2. The residual analysis shows no abnormality or particular structure (see Table 3), which validates the regression:

$$\ln(K) = 1.373 \ln(\text{PhExpDr}) - 1.322 \ln(\text{mAL\_pc})$$
$$-4.535 \ln(\text{GDPcap}) + 10.536 \tag{10}$$

where:
K = number of estimated killed.
PhExpDr = physical exposure to drought.
GDPcap = GDP purchasing power parity per capita.
mAL_pc = modified percentage of arable land.

It is worth noticing that Sudan and Swaziland were rejected by the model. The food shortage in Sudan is more likely due to conflict rather than climatic conditions. The case of Swaziland is a problem of country size. The raster layer at 0.5° resolution was not precise enough to provide an accurate model for this country.

3.3 Earthquakes

Earthquakes affecting seismic hazard zones were modelled using the seismic catalogue of the CNSS (Council of the National Seismic System). Hypocentres records of the last 40 years (1965–2004) with magnitude equal to, or higher
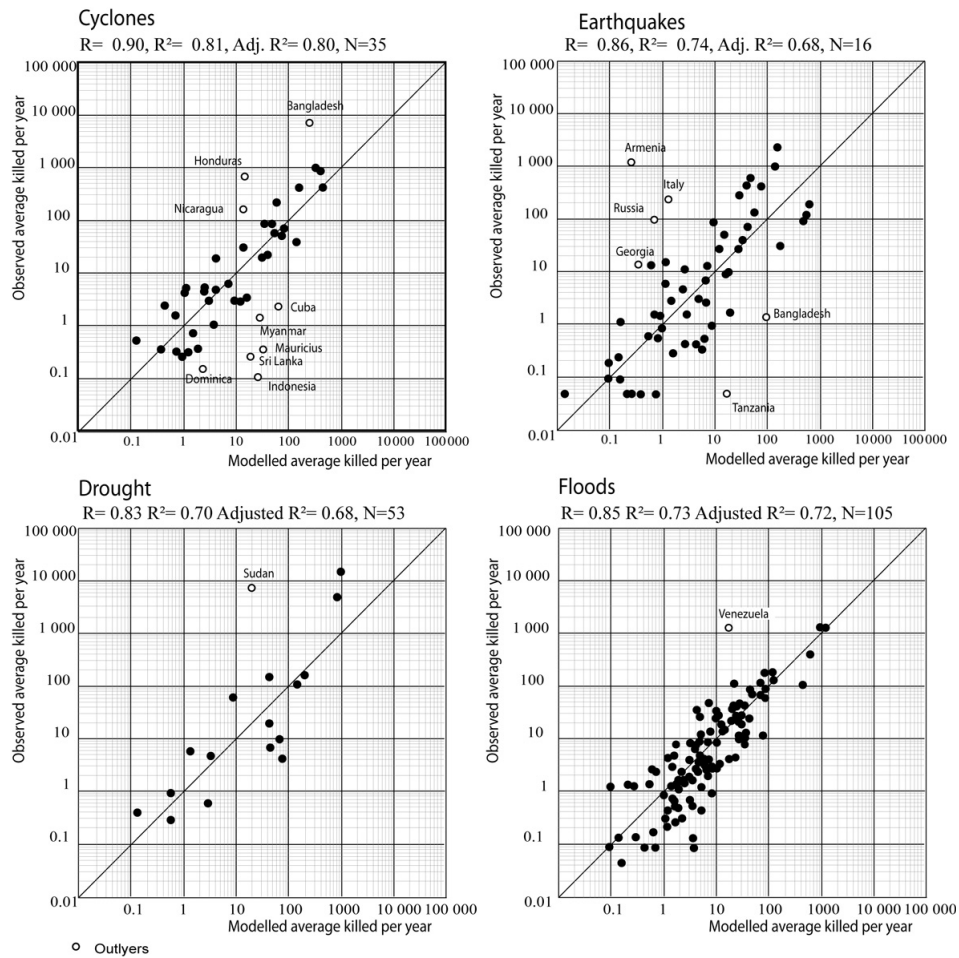
Fig. 2. Regressions between observed and modelled casualties (log/log scale). Observed casualties are the number of people killed per year during the period 1980–2000, according to the EM-DAT database (CRED). Modelled casualties are derived from the statistical model based on socio-economical indicators of vulnerability and physical exposures for each hazard.

Table 2. Model for tropical cyclones.

Variables	Coefficients	St. Err.	t Stat	P-value
Intercept	−0.487	1.897	−0.257	0.799278
GDPcap	−0.534	0.197	−2.719	0.010767
CropPC	0.347	0.152	2.283	0.029714
PhExp	0.621	0.067	9.301	0.000000

Table 3. Model for drought.

Variables	Coefficients	St. Err.	t Stat	P-value
Intercept	10.536	6.637	1.588	0.138375
GDPcap	−4.535	1.087	−4.172	0.001294
PhExp	1.373	0.408	3.365	0.005620
mAL_pc	−1.322	0.478	−2.764	0.017148

than, 5.5 on the Richter scale were used to generate circular buffers of Modified Mercalli Intensity (IMM). The radius of each buffer was based on intensity derived from depth of hypocentre and magnitude based on Kawasumi Eqs. (16).

3.3.1 Kawasumi equation (earthquakes)

$$IJMA = -0.3 + 2M - 4.6\log(d) - 0.0018d \quad \text{when } d < 100\,\text{km}$$

$$IJMA = -4.0 + 2M - 2.0\log(X) - 0.0167X \quad \text{when } d > 100\,\text{km}$$

where:
IJMA = intensity of Japan Meteorological Agency.
M = magnitude.
d = distance from epicentre (km).
X = distance from hypocentre (km).

For each buffer, the whole range of intensity (1–12) was taken into account. This is a general approach that does not take into account any regional effects, for instance soil conditions or geotectonic characteristics (JSSMFE, 1993). Physical exposure to earthquakes was then calculated and aggregated at country levels.

Physical exposure to earthquakes was then calculated and aggregated at country levels using Eq. (3). The variables retained by the regression are PhExp, U_g (rate of urban growth) and percentage of forest cover.

A high exposure and a high urban population growth being positively correlated to high risk of casualties, whereas a high forest coverage was correlated with less risk of casualties. This can be interpreted as high rates of population influx to cities as a synonymous of low quality urban planning and building standards. Or newcomers are living in areas previously empty because of the risk from earthquakes (unstable land, slopes, etc.). The percentage of forest, although with a low significance in the model, can be understood as the consequence of deforestation on slopes, thus leading to higher risk of landslides in earthquake prone areas.

The model is the following (see details in Table 4):

$$\ln(K) = 1.097 \ln(\text{PhExp40}_{\text{Eq.}}) + 25.696\, U_g$$
$$-0.425 \ln(\text{WoodPC}) - 17.344 \tag{11}$$

where:
$K =$ number of estimated killed.
PhExp40 = average population exposed to earthquakes (1964–2004).
$U_g =$ percentage of urban growth (computed using a three-year moving average).
WoodPC = percentage of country forest coverage.

The part of explained variance is smaller than for droughts or cyclones ($R^2 = 0.74$); however, considering the small length of time taken into account (36 years as compared to earthquakes long return period), the analysis delineates a reasonably good relation. Physical exposure is as relevant as in previous cases.

3.4 Floods

Although floods can be modelled using GIS, they request highly detailed data and complex procedures. For this study, a more generalised model was achieved by using the "comments" information in the EM-DAT database (Burton et al., 1978). The locations found in EM-DAT were used to select the watersheds in the HYDRO1k Elevation Derivative Database (US Geological Survey, http://edc.usgs.gov/products/elevation/gtopo30/hydro/) as approximations of the flooded areas as explained in previous study (Burton et al., 1978). This constitutes a global flood dataset taking into account 21 years covering 82.2% of the events (the one that had the necessary information) and 85.4% of the victims. Once the watersheds were identified, a computation of physical exposure was performed using the Eq. (3).

Table 4. Model for earthquakes.

Variables	Coefficients	St. Err.	t Stat	P-value
Intercept	−17.344	1.934	−8.970	0.000000
WoodPC	−0.425	0.135	−3.141	0.002856
Ug	25.696	4.342	5.918	0.000000
PhExp	1.097	0.126	8.714	0.000000

The variables identified by the statistical analysis are PhExp and GDPcap. Once again, GDPcap being highly correlated with HDI, this latter could have been chosen as well. The GDPcap was chosen due to a slightly better correlation between the model and observed casualties, and also due to lower p-values. Not surprisingly, the regression proves that highly exposed and poorer populations are more subject to suffer casualties from floods. The part of explained variance ($R^2 = 0.73$) associated with significant p-value on 90 countries+, as well as correct residual analysis, confirma solid confidence in the selection of the variables.

The model is the following (see details in Table 5):

$$\ln(K) = 0.905 \ln(\text{PhExp}_{Fl}) - 0.697 \ln(\text{GDP}_{\text{cap}})4.799 \tag{12}$$

where:
$K =$ number of estimated killed.
GDPcap = normalised Gross Domestic Product per capita (purchasing power parity).
PhExpFl = average number of persons living in watersheds affected by floods.

4 Multiple risk and categories

Multiple risk figures were computed by summing up modelled human losses from droughts, earthquakes, floods and tropical windstorms. For 16 out of 38 countries+ with missing data (i.e. either socio-economic parameters or exposure data), an estimated risk value of 0 was assigned because the exposure was considered to be negligible (less than 1000 people or 2% of the total population of the country exposed).

The DRI was computed for each country by taking into account both the absolute (killed per year) and the relative multiple risk figures (killed per year as percentage of the total country population). First of all, the log value of the two variables were normalised into 0–1 scales using the following thresholds : 0.5–500 killed per year and 0.1–10 killed per million per year (Fig. 3).

Then, the two normalised absolute and relative variables were averaged and classified using an equal-interval classification scheme (see Table 6), which was also applied to the observed data from EM-DAT for further comparison.
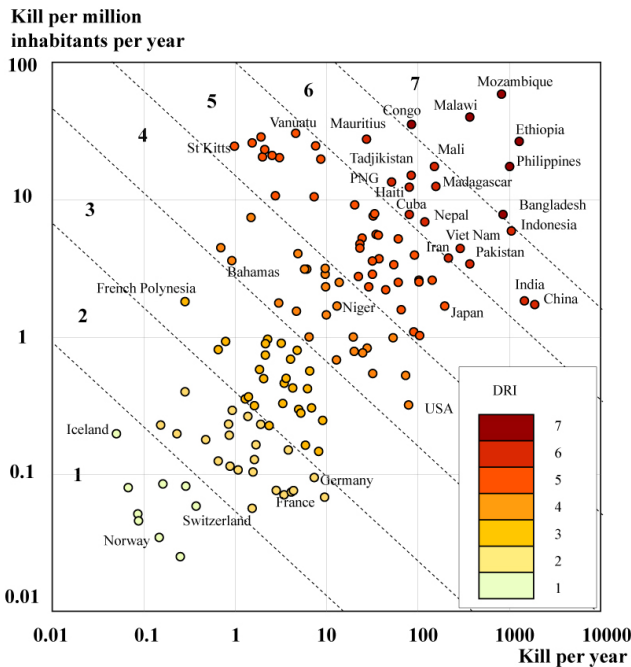
Fig. 3. Two dimensional classification in categories of risk.

5 Discussion

5.1 Identifying human vulnerability

Although a significant database was generated on vulnerability parameters (32 indicators) only five of them were finally retained by the multiple regression analysis (i.e. GDP purchasing power parity per capita, modified percentage of arable land, percentage of urban growth, percentage of country forest coverage, transformed value of the percentage of the country dedicated to Crop land). The selection was made by statistical tests (hence without subjectivity). Poverty (low GDPcap) is the most selected indicator, many other indicators are strongly correlated with poverty (such as Human Development Index (HDI), Urban Growth, number of physicians per inhabitants, etc.). Given that we cannot place two indicators that are strongly correlated in the same model, the selected indicators are those that provided the best R2, the smallest p-value and also the best countries+ and time coverage. However, GDPcap can most often be replaced by HDI or other correlated indicators, also with less precision in the model or with less countries+ covered).

5.2 Geographical distribution

The DRI could be computed for 215 countries+ (86% of the 249 countries+, representing 96% of the world population and 79% of the killed from EM-DAT). The main countries+ not included in the multiple model were: North Korea, Afghanistan, Somalia, Taiwan, Puerto Rico (missing socio-

Table 5. Model for floods.

Variables	Coefficients	St. Err.	t Stat	P-value
Intercpt	−4.799	1.055	−4.551	0.000015
GDPcap	−0.697	0.102	−6.812	0.000000
PhExp	0.905	0.057	15.824	0.000000

Table 6. DRI classes.

DRI value	DRI class
$-\propto$	0 (no killed)
$]-\propto, 0]$	1
$]0.0, 0.2]$	2
$]0.2, 0.4]$	3
$]0.4, 0.6]$	4
$]0.6, 0.8]$	5
$]0.8, 1.0]$	6
>1.0	7

economic data), Swaziland, Tanzania (bad exposure data). These seven countries+ account for 99.7% of the missing killed from EM-DAT that could not be modelled. Other missing countries+ include several small island territories (see Table 7).

Without much surprise the top countries at risk in terms of killed per year are the most populated countries (China, India, Indonesia, Bangladesh), whereas small islands states (Vanuatu, Dominica, Mauritius, Antigua and Barbuda, St Kitts and Nevis, Solomon Islands, Grenada, etc.) come first in terms of killed per million inhabitants per year. Once the two indicators are combined to obtain the DRI, six of the top 10 countries are in Africa, the other countries+ being located in Asia. Islands states rank high in the DRI (Fig. 4). Some countries could not be modelled due to lack of data (see Table 7). All the DRI values are provided in Supplementary material B: http://www.nat-hazards-earth-syst-sci.net/9/1149/2009/nhess-9-1149-2009-supplement.pdf.

For nine countries (in bold in Fig. 5) the weights have little influence on the DRI since they are ranked high (in the top 25 countries+) in both indicators.

5.3 Unexpected DRI values

Although more than 90% of the modelled classes have a difference of less than classes (50% with no difference, 30% with a difference of one class, and 9.3% with a difference of 2). There are some unexpected values. When comparing DRI classes based on the models with those derived from observed data (EM-DAT), 9 countries show a difference greater than 3 classes (see Table 8). This does not necessarily reflect limits of the models. Underestimated values rather highlight
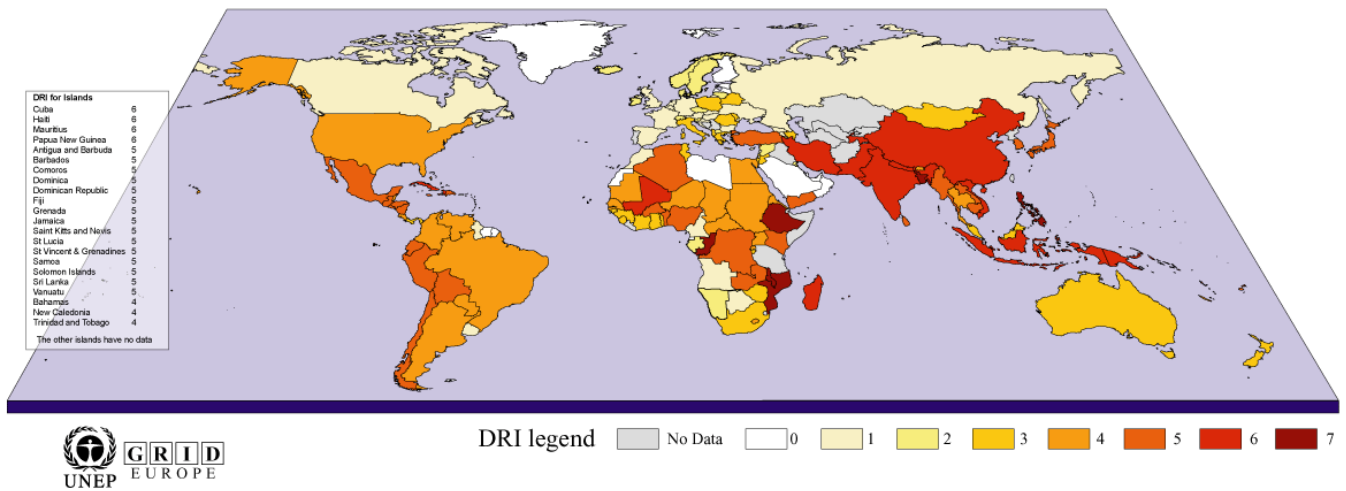
The Disaster Risk Index (DRI)

DRI legend: No Data | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7

Fig. 4. Spatial distribution of DRI classes.

Table 7. List of countries+ that could not be modelled due to lack of data.

independent states	Micronesia, Tonga
territories	American Samoa, Anguilla, Bermuda, British Virgin Islands, Cook Islands, Guadeloupe, Guam, Martinique, Montserrat, Netherlands Antilles, Niue, Puerto Rico, Reunion, Turks and Caicos Islands, United States, Virgin Islands, Wallis and Futuna

Fig. 5. Top 25 countries+ according to DRI (in bold: countries+ with both indicators in the top 25).

countries affected by extraordinary events (such as cyclone Mitch in 1998, earthquake in Armenia in 1988, floods in Venezuela in 1999). Overestimated values concern countries that are either drought prone areas or cyclone prone islands: in these cases, there were problems when computing physical exposure and/or in the classification of the victims in EM-DAT (e.g. the dubious value of 0 reported killed from droughts in Mali). Spatial comparisons between modelled and observed values can be seen in Fig. 6.

5.4 Comparison of the model with 2001–2006 observed casualties

The DRI cannot be used for estimating future number of casualties, e.g. it underestimated Pakistan (2005 earthquake) and Iran (Bam, 2003) or Haiti (hurricane Jeanne, 2004). However, these three countries were correctly classified (based on 1980–2000 data) as countries facing very high risk (class 6). The comparison of modelled and observed
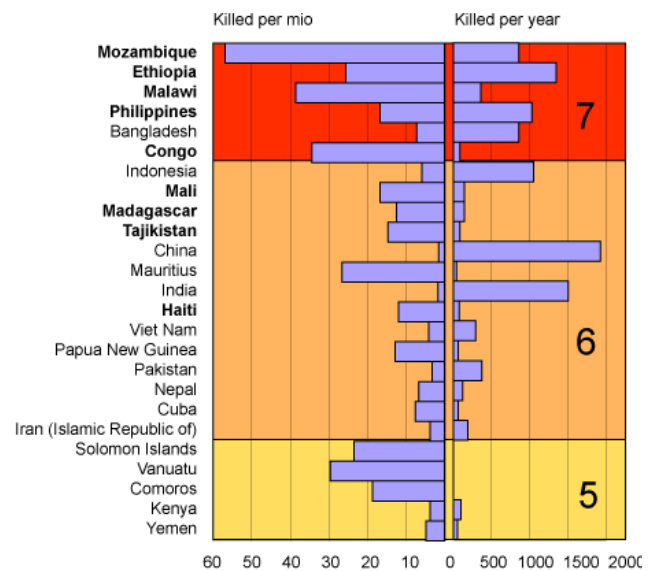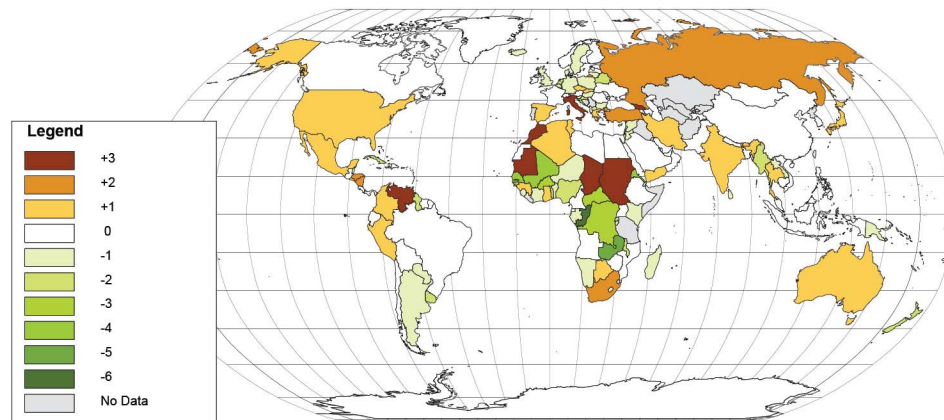
categories of risk shows that 71% of the countries have 1 class or less difference. Six countries show a difference of 4 classes: Morocco (underestimated, earthquakes in 2004), Congo, Papua New Guinea, Mali (overestimated for drought), Mauritius, Laos (overestimated for cyclones).

Fig. 6. Maps of differences between DRI modelled and DRI observed.

Table 8. Major class differences between observed data and model.

	Country	DRI (model) (model)	DRI (with CRED data)	Diffe-rence	Main cause of under/over-estimation
Underestimation	Armenia	2	7	−5	Earthquakes
	Chad	4	7	−3	Droughts
	Mauritania	4	7	−3	Droughts
	Venezuela	4	7	−3	Floods
	Sudan	4	7	−3	Droughts
	Italy	3	6	−3	Earthquakes
	Morocco	2	5	−3	Floods
	Georgia	2	5	−3	Earthquakes
Overestimation	Mali	6	2	+4	Droughts
	Mauritius	6	2	+4	Cyclones
	Eritrea	4	0	+4	Droughts
	Senegal	4	0	+4	Floods
	Grenada	5	0	+5	Cyclones
	Zambia	5	0	+5	Droughts
	Barbados	5	0	+5	Cyclones
	Congo	7	1	+6	Droughts

6   Conclusion: a flexible classification system to use with care

The DRI takes into account both an absolute and a relative risk indicator, allowing to consider populated and small countries+ in different manners. The somewhat arbitrary choice was to place a similar weight on killed per year and killed per inhabitant; any other combination would be possible according to the users' choice since the two indicators are provided along with the DRI.

For other countries+, the flexibility of the DRI classification system allows the users to specify whether more weight should be given to killed per year or to killed per inhabi-

tant. Typically, the users interested in Small Island Developing States (SIDS) will obviously choose the second solution. This model also remains open to adding future components to the DRI, such as the economic losses once the input data are improved.

Correlations between modelled and observed risks for each hazard type were surprisingly high and relevant, given the heterogeneity of the data sources and the coarse resolution of the data at the global scale. They successfully demonstrate a correlation between high levels of development and low numbers of casualties from these four types of hazards. This correlation can be understood both ways: low development may lead to high casualties, and high disaster occurrence may also lead to low economic development as it destroys infrastructure and crops as well as deterring investors.

As EM-DAT does not include small-scale disasters, the models calibrated on past events cannot address these kinds of events, which are more frequent and may cumulate, to be of concern for the developing process in poor countries. Further studies might be carried out on more detailed databases (e.g. based on DesInventar) to identify patterns of small disaster hazards.

There is a gap between the resolution of the hazard and exposure (5×5 km cells) and the vulnerability parameters (country level). Some indicators are now being generated at sub-national level, but so far only for a limited number of countries+ and indicators. Given that the DRI is provided at a national level, this is less of an issue. However, in large countries+ with significant discrepancies, (e.g. China, India) more disaggregated figures should be used in the future.

Even though the model was designed for understanding past casualties (1980–2000), by using modelled losses based on 2005 physical exposure and vulnerability parameters (as opposed to recorded losses), the DRI offers a picture of risks (due to natural hazards) at a specific time. This provides a risk level that is comparable; it doesn't depend on quality of country reporting and demographic changes are taken into

account. However, this model should not be used in a predictive way to estimate potential casualties that are usually also highly dependent on proximal parameters such as the time of the day (working hours, people asleep, etc.). The DRI does not differentiate risk from rare severe events and equivalent risk resulting from less severe, yet more frequent events. To overcome this issue, future modifications to the model should focus on less aggregated analyses and use event-based analysis to include hazard severity. Additionally, models should be made at sub-national levels with much higher resolution data, leading to danger maps for planning prevention and relief. In some cases, early warning systems and prevention measures can be implemented, while in others, the country is affected too frequently for coping with the new event : the high recurrence makes each new event worse than the previous ones, leading to what is called the ratchet effect.

The DRI is being used by UNDP\BCPR to identify countries+ in highest need for prevention and development. This study sets the basis of risk status for the period 1980–2000. Risk (from natural hazards) is likely to increase in the coming decades, since higher exposure to hazardous events will occur following population increases. However, exposure is not the only risk component that is likely to increase: depletion of natural resources and increasing gaps between rich and poor populations, political unrest and AIDS will probably have an impact on human vulnerability. Hydro-meteorological hazard frequencies and magnitudes might also change in the near future due to climate change and/or environmental degradation. To better understand where risk might increase in the future and to prepare for these future risk patterns, further refined analysis should be carried out on the three risk components. DRI can be used to prioritize where such detailed studies should be carried out and where improvements on data collection are needed; this is, however, not a final goal per se. Rather, final results will be achieved when proper risk reduction measures are implemented leading to an observed decrease in casualties.

Hazard and exposure data

Hazard and exposure data created for this study can be freely accessed at: http://www.grid.unep.ch/activities/earlywarning/preview/data/.

They can be visualised using the PREVIEW – Global Risk Data Platform: http://preview.grid.unep.ch/.

Edited by: T. Glade
Reviewed by: two anonymous referees

References

Blaikie, P., Cannon, T., Davis, I., and Wisner, B.: At Risk: Natural Hazards, Peoples Vulnerability, and Disasters, Routledge, London, 1994.

Burton, I., Kates, R. W., and White, G. F.: The environment as hazard, Oxford University Press, New York, 240 pp., 1978.

Cardona, O. D.: Indicators for Disaster Risk and Risk Management. Program for Latin America and the Caribbean: Summary Report, Manizales, Columbia: Instituto de Estudios Ambientales, Universidad Nacional de Columbia, 2005.

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

Dao, H. and Peduzzi, P.: Global evaluation of human risk and vulnerability to natural hazards, Enviroinfo 2004, Sh@ring, Editions du Tricorne, Genève, I, 435–446, 2004.

Dao, H. and Peduzzi, P.: Global Risk And Vulnerability Index Trends per Year (GRAVITY), Phase IV: Technical annex and multiple risk integration, UNDP/BCPR, Geneva, Tech. Rep., 31 pp., 2003.

Dilley, M., Chen, R. S., Deichmann, U., Lerner-Lam, A. L., and Arnold, M.: Natural Disaster Hotspots, a global risk analysis, the World Bank, p. 112, p. 132, 2005.

Fifth International Workshops on Tropical Cyclones – IWTC-V, Cairns, December 2002.

Holland, G. J.: An analytic model of the wind and pressure profiles in hurricanes, Mon. Weather Rev., 108, 1212–1218, 1980.

Kawasumi, H.: Measures of earthquake danger and expectancy of maximum intensity throughout Japan as Inferred from the seismic activity, Bull. Earthq. Res. Inst. Univ., Tokyo, 29, 469-482, 1951.

Mouton, F., Nordbeck, O., and Peduzzi, P.: Cyclone Database Manager, A tool for converting point data from cyclone observations into tracks and wind speed profiles in a GIS, UNEP/GRID-Europe, Geneva, Tech. Rep., 67 pp., 2002.

Peduzzi, P., Dao, H., and Herold, C.: Mapping Disastrous Natural Hazards Using Global Datasets, Nat. Hazards, 35(2), 265–289, 2005.

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

UNDP/BCPR: A Global Report: Reducing Disaster Risk, A Challenge for Development, New York, 146 pp., 2004.

UNDRO: "Natural Disasters and Vulnerability Analysis": Report of Expert Group Meeting, 9–12 July 1979.