Mapping Disastrous Natural Hazards Using Global Datasets

Peduzzi, Pascal; Herold, Christian; Dao, Quoc-Hy

Natural Hazards (2005) 35: 265–289
DOI 10.1007/s11069-004-5703-8

Mapping Disastrous Natural Hazards Using Global Datasets

P. PEDUZZI★, H. DAO and C. HEROLD
UNEP/GRID-Geneva, 11 ch. des Anémones, 1219 Châtelaine-Geneva, Switzerland

(Received: 30 September 2003; accepted: 1 November 2004)

Abstract. The increased interest for categorising countries at risk calls for an improved methodology allowing comparison of natural hazard impacts at a global level. A disaster is the intersection between a hazardous event, the elements at risk (population, infrastructures) and their vulnerability. In order to associate reported impacts with affected elements and socio-economic or geophysical contextual parameters, geographical location and extent of hazards is needed. The scope of this paper is to present improved automated procedures for a rapid mapping of large disastrous hazard events (floods, earthquakes, cyclones and volcanoes) using Geographical Information Systems (GIS) and available global datasets. Up to 82% of the events and 88% of the reported victims could be geo-referenced and the results highlight both the potentialities and limitations of the methods applied.

Key words: natural hazards, geo-referencing, GIS, CRED, EM-DAT, risk

Abbreviation: CNSS – Council of the National Seismic System; CRED – Centre for Research on the Epidemiology of Disasters; DEM – Digital Elevation Model; DFO – Dartmouth Flood Observatory (University of Dartmouth, USA); EM-DAT – Emergency Events Database (OFDA/CRED); FAO – Food and Agriculture Organization; FEWS – Famine Early Warning System Network; GIS – Geographic Information System; NGDC – National Geophysical Data Center; OFDA – Office of US Foreign Disaster Assistance; PREVIEW – Project of Risk Evaluation, Vulnerability, Information & Early Warning (UNEP/GRID-Geneva); UNEP/GRID-Geneva – United Nations Environment Programme/Global Resource Information Database – Geneva; VEI – Volcanic Explosivity Index; WRI – World Resources Institute

1. Introduction

1.1. CONTEXT

A number of initiatives from international organisations attempt to categories countries according to their levels of risk from natural hazards, either for advocacy purposes or for the prioritisation of budgets dedicated to risk reduction or to relief aid. Two recent examples of such studies are the annual *World Disasters Report* (IFRC, 2003) and *Reducing Disaster Risk: A*

★ Author for correspondence: E-mail: pascal.peduzzi@grid.unep.ch

Challenge for Development (UNDP, 2004). A key component for the preparation of such publications is the use of international databases with global coverage and comparable methods for the estimation of figures. The two above reports are based on the only publicly available global database on impacts from natural hazards: EM-DAT from the Centre of Research on Epidemiology of Disasters (CRED).[1]

1.2. RISK, HAZARD, VULNERABILITY

Besides reliable and comparable data, clear definitions of concepts are first needed. "*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). The risk results from three factors:

- "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).

Databases such as EM-DAT provide figures on past human and financial losses (the realised risk). Statistics delineate great discrepancies between the number of casualties resulting from hazardous events of comparable scales depending on their locations (Blaikie *et al.*, 1996; Smith, 1996; Tobin and Montz, 1997). These differences can be explained by the vulnerability of the elements at risk, e.g. the human vulnerability to a natural hazard, the population capacity to cope with the event.

An estimation of human loss can be performed by computing the ratio of killed per population affected (Peduzzi *et al.*, 2001, 2002; UNDP, 2004). For this purpose the areas affected by individual hazards need to be identified in order to evaluate the population affected. Therefore, knowledge on its location constitutes a compulsory step for understanding why a hazardous event is turning into a disaster. Comparing discrepancies between resulting losses in different countries calls for a standardised method for the computation of the areas and populations exposed to natural hazards.

[1] EM-DAT: The OFDA/CRED International Disaster Database, Université Catholique de Louvain, Brussels, Belgium (http://www.em-dat.net/).

1.3. GEO-REFERENCING

Global databases on losses, such as EM-DAT, do not provide a proper geo-reference of the reported events. Figures are generally given on a country by country basis; place names and, less frequently, a unique pair of coordinates provide a more precise indication of the location, but say nothing about the event extent nor on the exposed elements (population, land, etc.).

Methods for global mapping of disastrous events at a sub-national level were developed by Verelst (1999) and Food and Agriculture Organization, FAO (2001).[2] These methods, based on the names of first administrative units, were applied to respond to a certain level of decision for intervention and preparedness. They are useful for decision-makers since the legal responsibility for intervention lies with administrative entities.

However these methods are inappropriate for comparing situations across countries and regions because the first administrative level, as well as other levels, varies significantly in size and population. Furthermore, earthquakes, floods, cyclones and all natural hazards do not follow administrative boundaries, especially when these administrations cover large areas. The identification of the areas physically affected by such hazards requires extensive modelling. If sophisticated GIS models have been developed for this purpose, the main difficulty while producing a global survey is the access to data with a global coverage offering similar levels of quality for comparison purposes. One cannot use detailed data available in some locations and rough estimates for others.

1.4. OBJECTIVES

The issue of missing relevant spatial information at the global scale was a major concern during previous research at United Nations Environment Programme/Global Resource Information Database-Geneva (UNEP/GRID-Geneva) (Peduzzi *et al.*, 2001, 2002; UNDP, 2004). The aim of the present research is to improve the identification of affected areas and populations by moving from an approach based on administrative boundaries to an identification of affected areas by geo-physical models. Although EM-DAT was not initially designed to be geo-referenced, the information included such items as coordinates of earthquakes, names of cyclones, names of volcanoes or comments, which can be used to produce links with geo-spatial databases and maps, in order to identify the spatial extent of each reported event. Subsequently, this study should highlight the information to be systematically introduced into impact databases for a simplified linkage of reported losses with geo-physical databases. This article does not address the precise location of events for prevention and mitigation purposes, but

[2] Some geo-spatial datasets created with this method are available at http://www.em-dat.net/.

responds to the needs of decision-makers at the global level who want to further understand and identify vulnerable populations.

Part 2 of the article presents the best available datasets with a global coverage for identifying areas affected by floods, earthquakes, volcanoes and tropical cyclones. It is followed by the presentation of procedures that can be applied for linking the hazards with their impacts as reported in the EM-DAT database with the help of Geographical Information Systems (GIS). The strengths and weaknesses of the methods are presented but a full description of the geo-spatial modelling process is not the focus of the present paper. The article concludes with a discussion on potential improvements and applications of the EM-DAT.

2. Availability and Level of Geo-reference of Global Datasets

2.1. DATA ON LOSSES: THE CRED EM-DAT

Three databases currently record impacts from natural hazards world-wide. Two of them are maintained by private re-insurance companies (namely Sigma from Swiss Re and NatCat from Munich RE) but these are not in the public domain, whereas the Emergency Events Database (EM-DAT) is. It was created in 1988 and maintained ever since by the Centre for Research on the Epidemiology of Disasters (CRED) at the University of Louvain (Belgium). It is freely downloadable on the web[3] and regularly improved and updated. Other publicly available databases exist but without global coverage, such as for example DesInventar maintained by La RED, which covers so far 16 countries of Latin America.

CRED plays a key role in global risk assessment studies. Data is compiled from verifiable sources, including UN agencies, non-governmental organisations, insurance companies, research institutes and press agencies.[4] EM-DAT contains data on about 13,000 disasters, retrieving available information back to 1900 and continues to collect any new event that includes at least one of the following characteristics: more than 10 casualties, 100 affected, a call for international assistance and/or a declaration of a state of emergency.

The natural disasters reported by CRED are droughts, earthquakes, epidemics, extreme temperatures, famines, insect infestations, floods, slides, volcanoes, wave/surges (tsunamis), wild fires and windstorms (including cyclones). Human impacts are reported as follows: killed, injured, homeless, as well as an evaluation of affected people. The financial losses are introduced when available in equivalent US$ and/or in local currency. A comparative

[3] See at http://www.em-dat.net/.
[4] See detailed list at http://www.em-dat.net/.

study, of EM-DAT, Sigma, and NatCat carried out on four countries between the three databases (Guha-Sapir and Below, 2002), shows that the data collected vary significantly (up to 37% difference for casualties, 66% for those affected and 35% for damage). This highlights the difficulties of achieving a precise evaluation of areas affected by a hazardous event and the need for an independent method to assess the population affected.

The information included in EM-DAT that can be used for geo-referencing varies from general comments on place names to precise latitude/longitude coordinates (e.g. for earthquakes). The columns providing direct or indirect information on the location are the following:

– Country (name of the country)
– DisName (name of the disaster, useful for cyclone and hurricanes)
– Year, month, day (for location in time)
– Location (usually name of provinces)
– Latitude, longitude
– Comments

The present study focuses on four natural hazards (cyclones, floods, earthquakes and volcanic eruptions) during the 21 year period from 1980 to 2000. This period was chosen after establishing that access to information on casualties appears to decrease logarithmically before 1980.

Geographical coordinates (latitude and longitude) are only provided for 41% of earthquakes and 50% of volcanoes (see Table I).

2.2. DATA ON HAZARDS

A number of institutions world-wide are concerned with the collection of data on hazards at a global scale. In the context of the present study, the specific characteristics needed are as follows:

– time, location and extent of the disasters;
– severity (i.e. any information permitting the distinction between major and minor events);
– disaster names (cyclones and volcanoes).

Table I. EM-DAT disaster records with information on location (1980–2000).

Disaster type	No. of events	Killed	Country names	Place names	Event name	Coordinates
Earthquake	617	158,551	100%	44%	–	41%
Flood	1628	170,010	100%	37%	–	0%
Volcano	88	25,977	100%	84%	97%	50%
Cyclone	1076	245,546	100%	79%	57%	0%

2.2.1. *Cyclone data*

There are several types of wind storms. This part of the study focuses on tropical cyclones only, because this phenomenon follows processes that can be approximated by models such as developed by Holland (1980). The minimum required data are the central pressure, windspeed, latitude/longitude and time. Different sources (including the World Meteorological Organisation centres and other national agencies) were found on the Internet, as listed in Table IV (Appendix A). Some procedures were needed to harmonise the different units (e.g. knots *versus* meters per second or mbar versus hpa) used in different countries. An application for automatic conversion has been developed at UNEP/GRID-Geneva in order to provide a standardised version.

The information used were the date, name of cyclones, time, latitude, longitude, estimated central pressure (hPa), estimated maximum sustained surface wind (m/s) – the measures made every 6 hours allow the computation of the average velocity of the cyclone 'migration'.

2.2.2. *Flood data*

No global database on floods, covering the period of the study, was available. Although the Dartmouth Flood Observatory (DFO)[5] is constantly improving the temporal coverage of its flood database (1985 to 2003) and provides highly relevant information on floods, including duration and severity, the files provided on the internet are only gross approximations of areas affected at global scale. Furthermore, global mapping of flood hazards from earth observation satellites will remain difficult for events that occurred when radar images were not available (before 1991, date of the launch of the European Radar Satellite – ERS-1). Therefore the data used to estimate flooded areas were the watersheds derived from the USGS HYDRO1k Elevation Derivative Database (see Table IV in Appendix A for URL and sources) at a resolution of 30 arc-second (or 0.00833 degree; 0.9 km at the equator). The original HYDRO1k data consists of polygons (drainage basins) organised in a hierarchy of watersheds.

2.2.3. *Earthquake data*

Numerous organisations record earthquake epicentres and magnitudes. The most complete and convenient one for the purpose of this study was the Earthquake Catalog provided by the Council of the National Seismic System (see sources in Table IV in Appendix A). This database offers a global

[5] Darmouth Flood Observatory http://www.dartmouth.edu/~floods/index.html.

coverage from 1964 and is updated daily. Three hundred and thirty events are recorded for the period 1980–2000.

All the information needed for the model was available: date, time, latitude/longitude (in decimal degrees), depth and magnitude. The epicentre location has a precision of 0.01 degrees ($\cong$ 1.1 km at the equator).

2.2.4. *Volcanoes*

The magnitude of volcanic eruptions is provided according to the Volcanic Explosivity Index (VEI). This is a magnitude established by Newhall and Self (1982), integrating quantitative data as well as descriptions of observers. The scale (0 to 8) describes an increasing explosivity. Each level corresponds, among others, to a particular volume of explosive products, eruptive cloud height and descriptive terms (Simkin and Siebert, 1994).

The information contained in the database covers the year of eruption, the month and day, latitude/longitude, the magnitude (VEI), the name of volcano and the country.

2.3. PRECISION AND LIMITATIONS

Several general remarks and caution must be discussed before any analysis and interpretations of results are performed.

2.3.1. *Victims*

Table I depicts the number and percentage of EM-DAT events with different levels of geo-reference. Although all of the events include information on the country, only a portion of them includes a more precise location (i.e. place names, geographical coordinates).

2.3.2. *Natural events*

Core geo-spatial data for most of the events were available, except for floods. In several cases, geophysical datasets show missing years and/or missing countries. For example: cyclone tracks are available for 20 years world-wide, except for India, Pakistan and Bangladesh where only 9 years were found.

The main contribution of this study is to extrapolate an area for each potentially hazardous event from point coordinates or place names.

3. Modelling Methods

3.1. TROPICAL CYCLONES

Three types of physical impacts have to be taken into account to estimate the area and population affected by a cyclone: the wind velocity, the storm surge[6] as well as heavy rain producing floods. The possibility of geo-referencing the CRED database was tested using a wind model, but other models including storm surge or precipitation could be used in the same way, as long as detailed data on temperature, elevation, bathymetry and coast lines are available.

Moving from tracks of cyclones to an area can be modelled using a complex formula taking into account central pressure, windspeed and other variables. For transforming tracks of cyclones into a windspeed surface, several models are available. One of them is derived from a model developed by Holland (1980) from an original approach by Schloemer (1954) and further adapted by UNEP/GRID-Geneva.

Automated procedures were applied in order to process 21 years of data for cyclones world-wide to produce buffers according to the maximum sustained surface winds and central pressure, the outputs consisting on Saffir-Simpson windspeed classes (Mouton and Nordbeck, 2003). This dataset called Project of Risk Evaluation, Vulnerability, Information & Early Warning (PREVIEW) – Global Cyclone Asymmetric Windspeed Profile, is available for download.[7]

Once the buffers are computed, each file name includes the cyclone name and year allowing the creation of a double 'key' based on year and cyclone name. Alternatively, if the name of the cyclone was not introduced into the CRED database, links could be made using the precise date (year, month, day) and the intersection with country reported in EM-DAT. This last method is not as straightforward and requires a little intervention from the GIS operator. Figure 1 describes the procedures as well as an example for the cyclone Angela.

For cyclones that have a name in EM-Dat, a link was made using the cyclone's name and the year from both tables (PREVIEW Global Cyclones Asymmetric Windspeed Profile and EM-Dat). A verification was applied by looking at the match with month and day, and by intersecting the cyclone's buffer with the country indicated in the EM-Dat record. If the name was not mentioned in EM-Dat, the vector borders of the concerned country were intersected with the cyclone buffers for the selected year, month and region (as described in Figure 1). A link was created for any cyclones that intersected with the country during the specific month.

[6] The sea level may rise up to 5 m under the action of winds during a hurricane. This higher wave can rapidly flood a shallow coastal area. The local topography can also contribute to a concentration of these waters when returning to the sea.

[7] http://www.grid.unep.ch/data/gnv200.php.

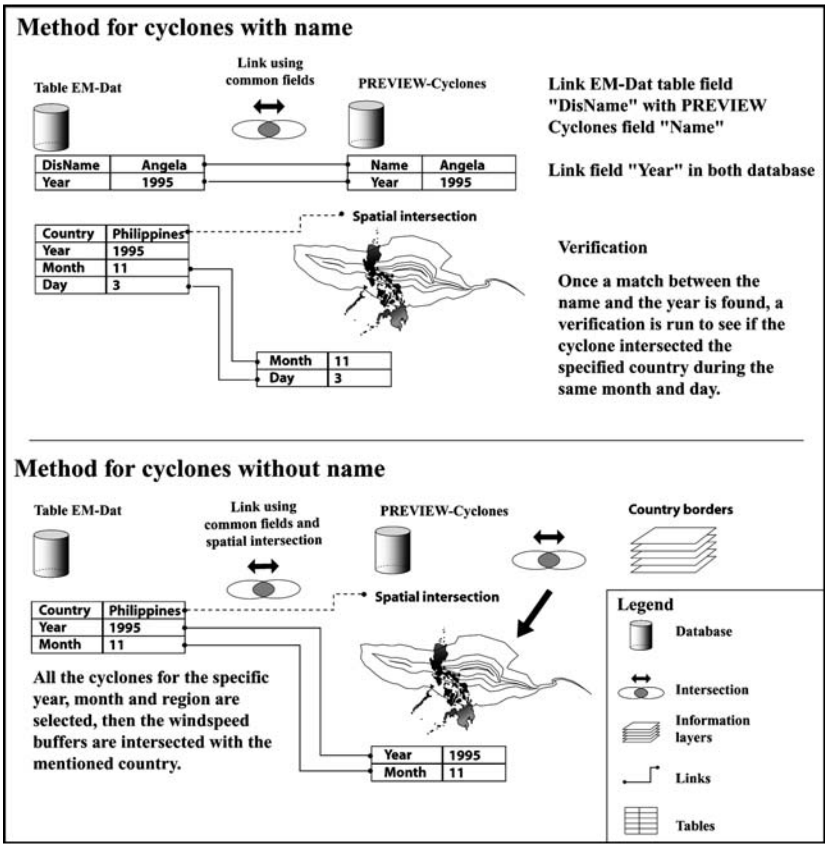

Figure 1. Links between EM-DAT and geo-spatial cyclone datasets.

3.2. FLOODS

Flood is the only event type where no automated procedures could be applied, because event names are non-existent for floods and no geographical coordinates are provided in EM-DAT. Some improvements were also necessary to remove data artefacts in the HYDRO1k data: watersheds in Asia had to be closed in high latitudes; some watersheds of small size (a few pixels) did not have any ID code (mainly in coastal areas), and in such cases they were absorbed in the surrounding watersheds. Furthermore, the watershed data about Australia was not available in electronic format and had to be digitised from the Australia's River Basins 1997 map (see Table IV in Appendix A).[8] In order to have all the watersheds in the same projection,

[8] At the time the present study was carried out, the digital version of the Australia's River Basins 1997 was not yet available on-line.

they were re-projected in decimal degrees from the original azimuthal Lambert projection.

The HYDRO1k watershed map represents only an approximation of flooded areas but a more complex modelling based on elevation, slopes, presence of vegetation, type of soils and climatic data was beyond the scope of this study.

A straightforward procedure was then applied: names of cities, streams, regions, provinces or states were identified in the EM-DAT columns ‘location’ and ‘comments’, and their coordinates were obtained from gazetteers (see Table V in Appendix A). These points were then used to select the affected watersheds as described in Figure 2. Finally, for each flood event, the disaster number is assigned to all the affected watersheds.

This procedure requires the most intensive human intervention, firstly in searching the geographical coordinates of the locations and entering them into a GIS, but also for verifying possible errors, paying attention to the following aspects:

– every point should be in the country defined by the disaster number;
– only one point by watershed per event,

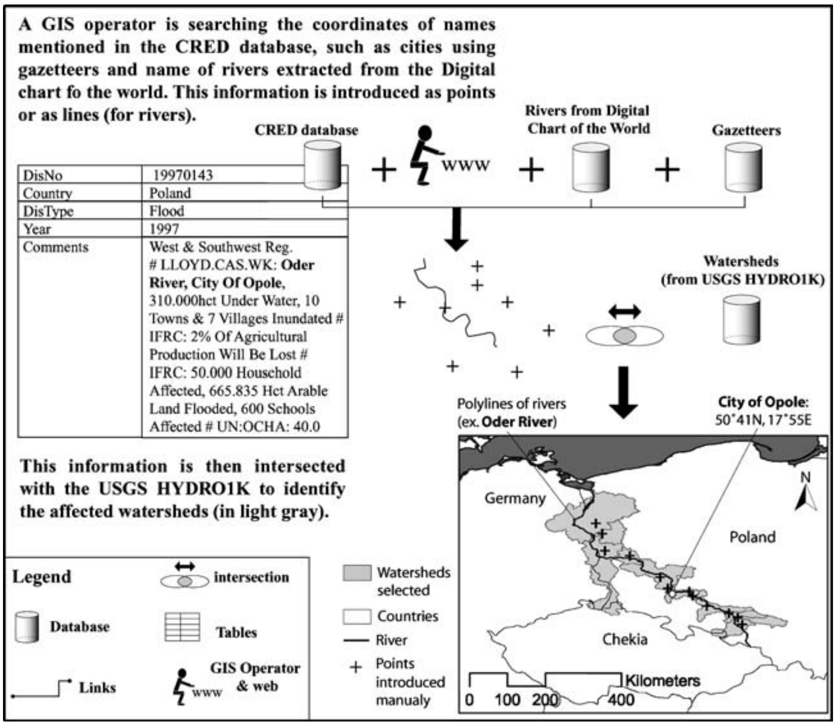

Figure 2. Information on floods from EM-DAT linked with geographical information.

– no duplicates (points having the same coordinates and disaster number);
– no points having coordinates not corresponding to a watershed (points out).

The method could be improved by adding detailed flood models. The release in 2004–2005 of a global Digital Elevation Model (DEM) at 90 m from the shuttle would drastically increase the resolution. Radar imagery could be used directly for determining flooded area, as already performed by ESA or the DFO and other agencies as seen in Figure 5. Obviously, this can only be applied to floods that have occurred after the launch of such satellites. The price of radar imagery is becoming more affordable allowing more possibilities for mapping. For previous events, a model of the surface affected is needed. Even for floods as detected by radar images, the area exposed is greater than the actual recorded flooded area: slopes, fields, houses are also affected by water going down-stream, whereas the flooded area is only recorded by satellite as large surfaces of water lying at the bottom of valleys.

3.3. EARTHQUAKES

The columns 'latitude' and 'longitude' of EM-DAT (epicentre location) were used to link the table of earthquakes with impacts as depicted in Figure 3. Buffers were drawn around each point to reflect the area affected by the earthquakes. Earthquake effects on the earth's surface are direct consequences of magnitude, depth of the hypocentre, distance to epicentre and subsoil effects, but once again, the model had to be simplified as the soils and fault orientations were not available globally.

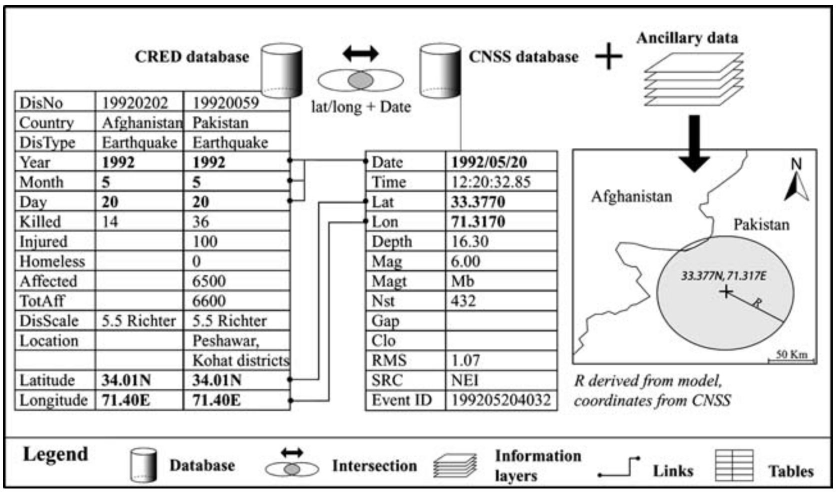

Figure 3. Earthquakes connection between EM-DAT information and geodatasets.

Table II. Buffer distances in relation to magnitude.

Magnitude	Buffer radius
<6.3	75 km
6.3–6.7	125 km
6.8–7.2	150 km
7.3–7.7	175 km
>7.7	200 km

A choice was made to produce seismic hazard zones using the seismic catalogue of the CNSS. Records of the 21 year period (1980–2000) were grouped into five magnitude classes. The thresholds for radius were fixed based on magnitude for which an estimation of ground motions duration, for specific acceleration and frequency ranges, is higher or equal to one second, as described in the table "*Bracketed duration in second*" from Bolt *et al.* (1975). Table II shows the resulting buffer distances.

Figure 3 describes the procedure followed to locate earthquakes.

The method presented in this study is automatic; this would not have been the case if geo-referencing was performed using the administrative names found in the column 'comments', a time consuming approach when hundreds of events are concerned.

3.4. VOLCANOES

The names of volcanoes are provided in the CRED database under the column 'DisName'. This information associated with the name of the country – to reduce possible confusion – allows a link of CRED data with the Worldwide Volcano Database from the National Geophysical Data Center (NGDC; see Table IV in Appendix A for references) and to extract crater coordinates as well as the VEI used to derived the area around each volcano (from the NGDC: World-wide Volcano Database). Figure 4 describes the procedure for geo-referencing the EM-DAT for volcanoes.

The manifestations of volcanic activity vary significantly, depending on volcanological and regional characteristics. For instance, lahars are linked to many parameters such as pluviometry, seismicity, topography and soils characteristics, among others. Tephra falls are directly influenced by dominant wind direction, and may affect areas hundreds of kilometres away from the eruption. Ground water access to the magma may produce phreatomagmatic eruptions and thus significantly increase the level of explosivity. Each volcano should be (and even *could be* – as they are well studied and not too numerous) modelled individually. To give a rough idea on what could be performed as a first evaluation, a procedure taking into account VEI was applied to draw a surface around volcanoes according to their magnitude of eruption.

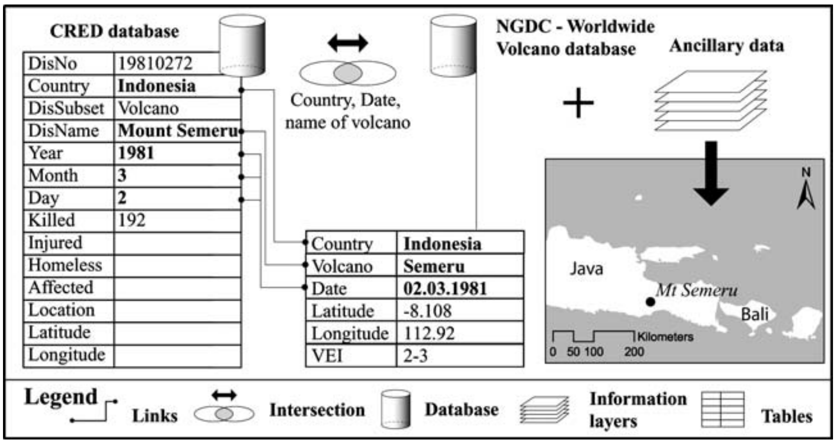

Figure 4. Volcanic eruptions, EM-DAT information linked with geo-spatial datasets.

As the principal causes of direct deaths are linked to explosive events, two groups of magnitudes were defined. The first one corresponds to VEI levels 2 and 3, described as explosive eruptions. The second corresponds to levels 4 to 8, described as cataclysmic, paroxysmal or colossal eruptions. The distances of 10 km were assigned to the first group and 30 km to the second. The choice of the distances was based on numerous maps of area affected as drawn from ground measures using various sources. The first two VEI levels (0 and 1) were disregarded, as they are usually not dangerous. Although casualties may be caused by small eruptions, they mostly result from secondary effects (e.g. lahars following precipitation, snow and ice melting,...). These are impossible to model at a global scale and request *in situ* precise data.

Ultimately the use of remote sensing and/or ground observation, would be the only way to achieve a proper definition of the area affected. A model of lava and lahar flows cannot be produced with efficiency due to the possible deformation – or even transformation – of the volcano during the eruptions. The simplified version provided here is a first cut-off for the area affected. In Figure 4, only the country borders are provided as background but information on slopes, elevation, population and other spatial layers could be added once the geo-referencing is achieved, in order to better estimate the risk faced by the population and the potential needs for prevention or relief actions.

4. Results and Discussion

The objective of the work was to geo-reference disastrous events as recorded in EM-DAT. Automated methods such as described in this research do not

Table III. Proportion of EM-DAT events which could be geo-referenced (1980–2000).

Hazards	EM-DAT records			After geo-referencing			
	Number of events	Victims	Events with coord. %	Geo-ref. events	Geo-ref. victims	% Geo-ref. events	% Geo-ref. victims
Earthquake	617	158,551	41	238	98,035	39	62
Flood	1628	170,010	0	1339	145,191	82	85
Volcano	88	25,977	50	40	22,736	46	88
Cyclones	1076	245,546	0	666	216,785	62	88

aim to compete with local impact mapping. As explained in the introduction, the quick geo-reference method aims to answer needs at global level using global data sets. A first quantitative evaluation consists in assessing the percentage of events that could be geo-referenced using the present methods. A second qualitative verification was carried out in order to assess the validity of the geo-reference method. A certain number of verifications were made using available maps of areas affected using either remote sensing techniques, or on ground measurements. These methods are described in parts 4.1 to 4.5. The four following examples illustrate either successes or difficulties of the geo-reference. They were chosen because they are well studied cases allowing multiple cross-references. Comparisons with the other alternative ways (e.g. using administrative units) are included; advantages and inconveniences of both methods are discussed.

The information in Table III depicts the number of events that could be geo-referenced.

Using the information in EM-DAT, the location and the spatial extent could be determined for up to 82% of the events and 88% of the victims within the period 1980–2000. The usually higher percentage of victims as compared with the percentage of events geo-referenced shows, quite logically, that severe events are generally more completely described, typically for earthquakes and volcanoes. It is to be noted that the percentage of geo-referenced earthquakes and volcanoes are less than the percentage of EM-DAT records with geographical coordinates. This is partly because it was not possible in some cases to find a match between the database on volcanic eruptions and EM-DAT due to differences in coordinates, names and/or dates, but also due to the fact that only major earthquakes and volcanic eruptions were modelled (respectively with a magnitude equal to or greater than 5.5 and with a VEI equal to or greater than 2). Therefore some EM-DAT events, probably caused by low intensity earthquakes or eruptions, were not geo-referenced.

Cyclones give an other example of the importance to have an event name in EM-DAT. If an average of 61.9% of the cyclones could be geo-referenced, this proportion is in fact 85.6% for cyclones with names and 30.5% for the ones without names. In the latter case (463 cyclones), 322 events could not be linked to the geo-spatial datasets because either no correspondences were found (259 events) or more than one correspondence was identified if several cyclones hit the country during the same month (this case occurred 63 times). In such cases, a manual verification might be performed to check which cyclone had the best match, using the day and information included in the column 'comments'.

Illustrated descriptions of the modelling results are provided further below for each hazard type.

4.1. FLOODS

Modelling the extent of disastrous flood events can be illustrated by the case of Mozambique floods in 2000, which drew particular attention from the media and the relief community.

Figure 5 identifies watersheds affected by the flood (dark grey): the areas are better defined than using the geo-reference at first administrative level

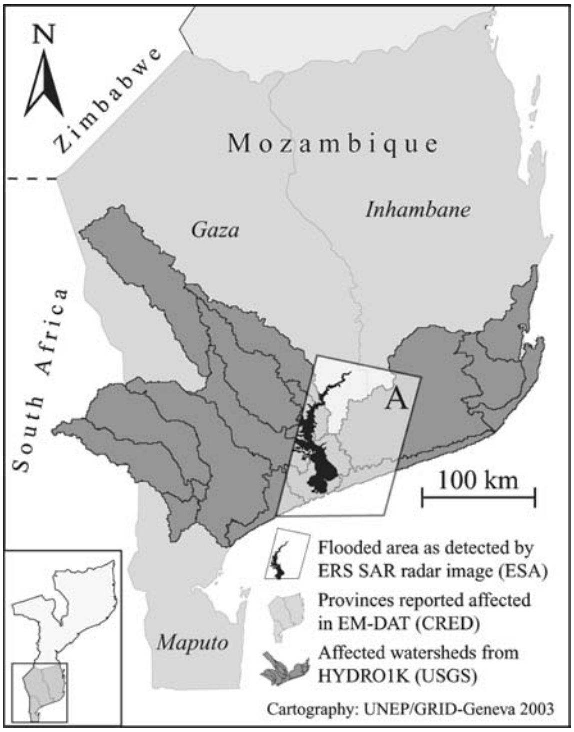

Figure 5. Map of floods in Mozambique (February–March 2000) comparing satellite observations and geo-reference models. Sources of radar image: http://earth.esa.int/ew/floods/mozambique_00/.

(light grey). An example of possible direct observation is provided by the radar detection (in black), but only covers a portion of the area affected (highlighted by the polygon A).

As already stated, this paper does not focus on precise modelling procedures, but on how to provide rapid geo-referencing of EM-DAT database using existing global datasets. However, two weaknesses of the method must be stressed. Firstly, entire watersheds are so far selected, whereas only a portion of them are actually flooded. The second aspect to be considered is the size of watersheds in the Hydro1K dataset, some being very large and some much more detailed. This is particularly true for central Asian countries, where the DEM lacks appropriate precision. All these parameters lead to an exaggeration of the area flooded and not necessarily in a consistent way.

Improvements on flood magnitude could also be achieved by incorporating more precise models including elevation, soils, land cover and computation of daily potential evapotranspiration fields computed with the Penman–Monteith equation (Allen *et al.*, 1998) as performed by Artan *et al.* (2002) for the Famine Early Warning System Network (FEWS) for Mozambique and by Funk *et al.* (2003) for developing an historical African rainfall.

4.2. EARTHQUAKES

To test the relevance of the simplified model developed for the purpose of this study the example of Izmit earthquakes of 1999 (Figure 6) was chosen because of the large availability of on-ground information and maps. A map of impacts based on ground measures (black areas) delineates Modified Mercalli Intensity Scale levels VII and VIII (considerable damage to poorly constructed buildings).

Both geo-references from administrative borders and from the model are close to the actual impacts. In this example, the administrative units are relatively small and the event is large. However, with a lower magnitude event and larger administrative divisions, the result using first administrative level would be much poorer as compared with the model.

The general lack of data on habitat quality, the complexity of earthquake modelling (chaotic impacts depending on population activities at the moment of the event), implies a drastic generalisation. However, the model could still be improved if types of soil as well as orientation of faults become available. The use of a formula for transforming magnitude into intensity such as developed by Kawasumi (1951), would lead to the improvement in the size and shape of the ellipsoid for identification of the area affected, as well as introducing a measure of intensity, necessary for determining potential harm to population.

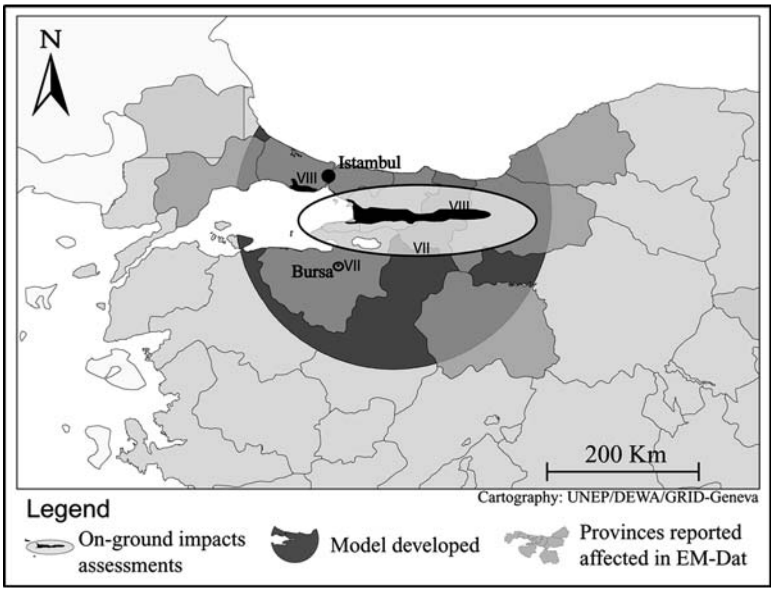

Figure 6. Map of Izmit earthquake comparing observations and geo-reference models. Sources of extent derived from on-ground measures according to the map from the Department of Earthquake Engineering, Kandilli Observatory and Earthquake Research Institute, http://www.koeri.boun.edu.tr/depremmuh/rms2.jpg.

4.3. CYCLONES

The method for geo-referencing tropical cyclones (see an example of the windspeed model in Figure 7) appears to be efficient: windspeed buffers are a realistic representation of the area affected by different windspeeds through time. In many cases, the information reported through press agencies and other sources fail to specify if the location is a city or a province. In this example – cyclone Angela in 1995 – the entry in EM-DAT specifies three affected locations: Luzon, Visayas and Calauag. Four administrative units of the first level can be selected using this information: Central Luzon and the three units of Visayas: East, Centre and West. The name of Calauag corresponds to three different cities located in different administrative units. This ambiguous indication therefore limits geo-reference using place names in EM-DAT. The spatial approach is quicker and allows an identification of affected areas independently from reporting sources. It also provides a spatial distinction between categories of wind.

In Figure 7, the model based on independent climatic data provides the area affected through time; to verify the accuracy of the model, one can superimpose the satellite image (GMS-5 of 2 November 1995 at 5:03 UTC). The

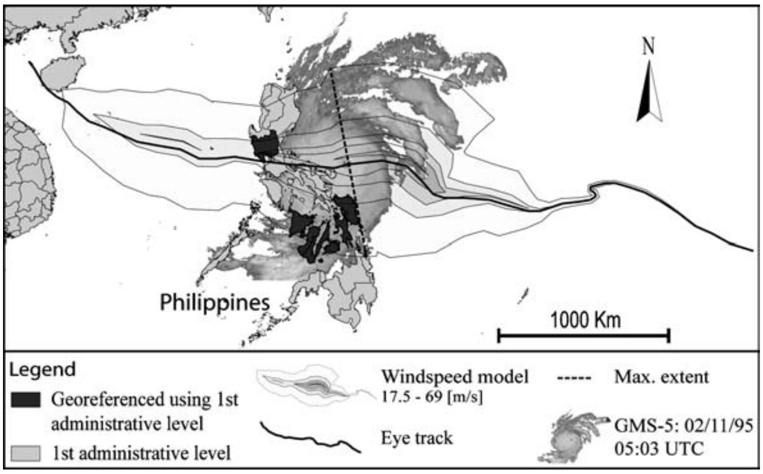

Figure 7. Map of Cyclone Angela (1995) as geo-referenced using different models. Original data sources: Cooperative institute for meteorological satellite studies, University of Wisconsin (USA) http://cimss.ssec.wisc.edu GMS-5 satellite image (visible, 1 km resolution).

diameter of the cyclones matches well with the buffer size. Once compared with reported impacts as mapped using first administrative units (in dark grey), the modelled buffer includes these areas and identifies additional affected units for which no victims were reported. The other advantage of using the PREVIEW Global Cyclones Asymmetric Windspeed Profile Data set relies on the possibility to overlay Saffir-Simpson classes of windspeed with the population, thus making it possible to derive how many persons have been affected by a certain windspeed. Such possibilities can easily lead to the estimation of population vulnerability while comparing differences in casualties.

4.4. VOLCANOES

Due to the small areas affected by lava flows/lahars, the automated method for geo-referencing volcanic eruptions provides a significant improvement as compared with the first administrative level approach. In the example of Etna eruptions shown below (Figure 8), the 10 km buffer is a good approximation of the area affected, and definitely much more precise than considering the extent of the whole province of Sicily.

A DEM adding slope information might further improve the extent of the mapping. However the random incidental source (and thus direction) of the lava flow is still mostly unpredictable as explained previously, which calls for individual modelling of volcanoes, if more precise impacts are needed.

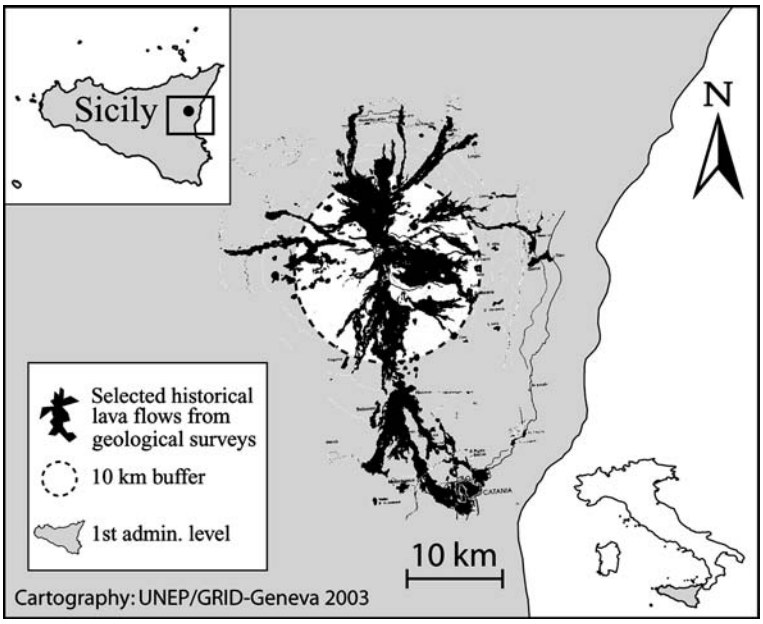

Figure 8. Map of Etna VEI 2-3 events as geo-referenced using a 10 km buffer. Sources of lava flows historic map: Behncke B., Dipartimento di Scienze Geologiche, Catania (IT), http://boris.vulcanoetna.com/ETNA.html.

4.5. COMPARING ADMINISTRATIVE AND SUSCEPTIBILITY MAPS BASED APPROACHES

To better understand the advantages and inconveniences of geo-referencing disasters on the basis of susceptibility maps, three examples of cyclone events have been taken (Figure 9), each of them representing a different situation and revealing several issues.

Firstly, the size of administrative units varies extensively (see West Australia *versus* units in Philippines). However the main difference between the geo-reference based on natural limits of hazardous events or on administrative units from EM-DAT, lies in the identification and mapping of the affected areas independently from the reporting of losses.

Furthermore, comments in EM-DAT concerning the location of impacts are provided by various sources. The column 'location' may include cities or administrative units. Geo-referencing EM-DAT based on administrative entities requires searching gazetteers for the latitude and longitude of each city mentioned. This strenuous task is often impeded by the difference in spelling or by the fact that many cities have the same name.

This can lead to three kinds of situation. In the first case (A: tropical cyclone Vance) the only entry of location in EM-DAT is "Exmouth, West

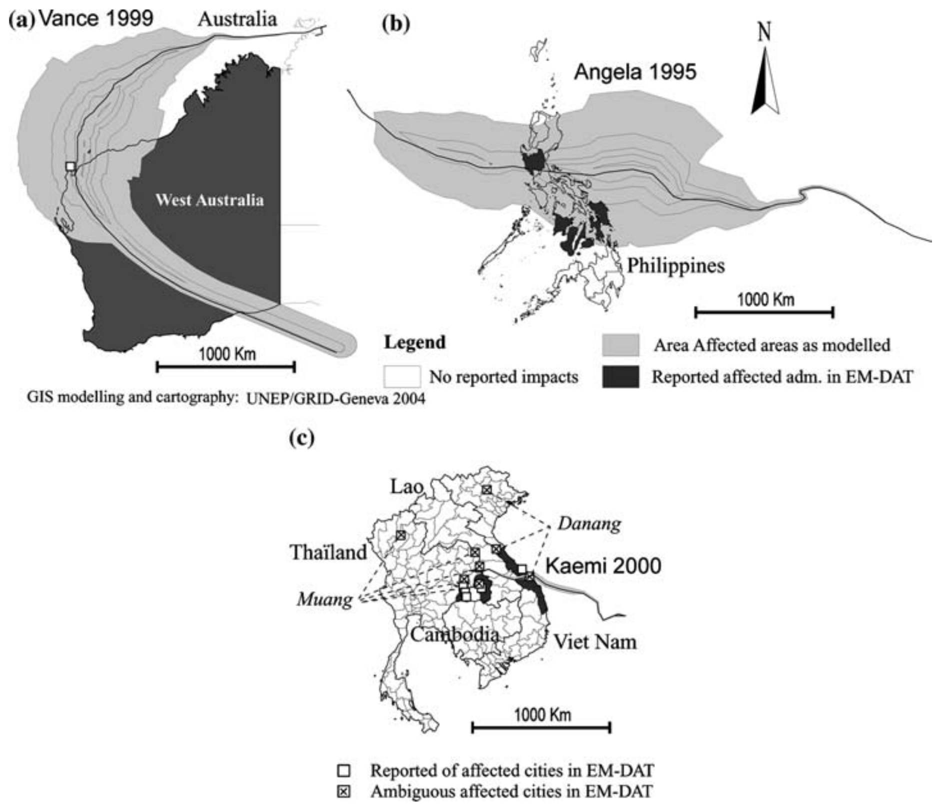

Figure 9. Comparison of geo-reference using administrative boundaries or geo-physical limits.

Australia". If the first administrative level is mapped, the huge area of the whole province of West Australia is irrelevant, but if only the city of Exmouth is mapped, a large area (in light grey) is dropped. In the second case (B: tropical cyclone Angela) there is a good match between the two geo-referencing methods. However, some provinces have no reported losses although they are in between two provinces with reported losses. In the third example (C: tropical cyclone Kaemi), the cities reported in EM-DAT are ambiguous, e.g. there are five different *Muang* in Thailand and three different *Danang* in Vietnam. Using the windspeed buffer, the affected area is easily identified. Linking EM-DAT with the PREVIEW Cyclone Asymmetric Windspeed Profile is limited by this model that only takes into account winds. In this case (C), the impacts would rather follow heavy rains (floods and landslides) since tropical cyclone Kaemi did not have significant winds; hence the affected area drawn from the buffers does not include all the reported administrations mentioned in CRED. However, the geographical information is useful complementary information, e.g. for identifying the correct Danang. In the same case, impacts are reported for Vietnam and Thailand, but not for Laos located in between the two. This type of case is interesting as it may provide useful information on the coping capacity and the reporting facilities of the countries.

5. Conclusion

Altogether, the method provides additional spatial information independent from the usual administrative levels. The automation allows quick mapping without involving a large amount of work (with the exception of floods). It improves information on disasters by providing a visual appraisal of the situation, with the possibility to compute area extents, population affected as well as ratios of casualties versus affected population, which can be used as a proxy of vulnerability.

Based on existing global datasets, the results presented here highlight the limitations of achieving comparable outputs when precise datasets with global coverage are still rare (although their availability is drastically improving). If a more precise location of events is needed, then the introduction of improved inputs will be needed.

The essential role of the EM-DAT information fields used (as described in Figures 1–4) must be stressed, in particular the names of cyclones and volcanoes, and the coordinates of earthquake epicentres. In the case of floods, one can eventually consider indirectly geo-referencing the events by introducing the standardised watershed identifiers (as the Pfafstetter codes of the Hydro1K database), this being a special case for which an interaction between CRED and Dartmouth would be useful. The lack of completeness of the required information (see Table I) was the main cause preventing the geo-referencing of a proportion of the events (see Table III).

Other outputs of this research consist of a series of global databases on natural hazards, which can be visualised or downloaded freely from the UNEP/GRID-Geneva website (http://www.grid.unep.ch).

The event numbers ('DisNo') from CRED were implemented as unique identifiers of events, allowing a link of impacts with all sorts of other spatial information as well as with socio-economic parameters from existing sources available at UNEP, World Bank, FAO or World Resources Institute (WRI). The foreseen use of a single standard identification number for each disaster, such as the GLIDE number,[9] will ease the process of geo-referencing hazards. In the case of cyclones, for EM-DAT unnamed events, the adoption of date and region as an interception key is needed and this will constitute a future improvement.

[9] http://www.glidenumber.net/.

This research highlights a great potential for basing EM-DAT on a GIS (at least for the four hazards studied), but this would generate higher costs and processing time for the CRED and the University of Louvain. It has to be acknowledged that, with the amount of resources available, the CRED staff has produced so far the best publicly available collection of events which covers the entire world.

Ultimately, these techniques open the doors to several applications, among them the assessment of trends in population vulnerability by comparing the casualties with the exposed population. From the outputs of this study, spatial queries and links with other ancillary spatial information can be made, thus allowing a better understanding of the contextual parameters that turn hazards into disasters. Such studies could play a significant role in identifying the most frequently affected areas and vulnerable populations. This is particularly true for climatic hazards such as floods and cyclones, the recurrence of the phenomena being higher than for tectonic events for which the long return period can prevent the computation of a corresponding vulnerability.

In conclusion, the methods for mapping disasters using administrative units and geo-physical modelling of hazards are complementary, since the former refer to intervention entities, while the latter reflect the physical processes. They both serve the purposes of better understanding the causes and consequences of natural disasters, as well as of more efficiently targeting the actions to be taken.

Appendix A

Table IV. Data sources for casualties and hazards.

Data on	Organisations	URL address
Casualties	Centre of Research on Epidemiology of Disasters (CRED) EM-Dat	http://www.em-dat.net/
Cyclones	UNEP/GRID-Geneva PREVIEW Global Cyclone Asymmetric Windspeed Profile	http://www.grid.unep.ch/data/gnv200.php
Cyclones	Unisys Weather:	http://weather.unisys.com/hurricane/
Cyclones	Typhoon, 2000:	http://www.typhoon2000.ph/
Cyclones	Australian Severe Weather:	http://australiasevereweather.com/cyclone/

Table IV. Continued.

Data on	Organisations	URL address
Cyclones	Atlantic Hurricane Track Maps & Images:	http://fermi.jhuapl.edu/hurr/index.html
Cyclones	Hawai'i Solar Astronomy:	http://www.solar.ifa.hawaii.edu/index.html
Cyclones	Bureau of Meteorology, Australia:	http://www.bom.gov.au/
Cyclones	Fiji Meteorological Service:	http://www.met.gov.fj/
Cyclones	Japan Meteorological Agency:	http://www.kishou.go.jp/english/
Cyclones	India Meteorological Department	printed material (years 1992 to 2001)
Floods	CRED EM-Dat database	http://www.em-dat.net/
Floods	U.S. Geological Survey, HYDRO1k Elevation Derivative Database (1997)	http://edcdaac.usgs.gov/gtopo30/hydro/
Floods	Australia's River Basins, 1997	http://www.ga.gov.au/nmd/products/thematic/basins.htm
Earthquakes	Council of the National Seismic System (CNSS): Earthquake Catalog	http://quake.geo.berkeley.edu/cnss/
Volcanoes	National Geophysical Data Center: Worldwide Volcano Database.	http://www.ngdc.noaa.gov/seg/hazard/vol_srch.shtml

Table V. URL for websites providing geographical coordinates.

Organisations	URL
Alexandria Digital Library Gazetteer Server	http://fat-albert.alexandria.ucsb.edu:8827/gazetteer/
Arizona State University Library	http://www.asu.edu/lib/hayden/govdocs/maps/ geogname.htm#us
Astrodienst	http://www.astro.com/cgi-bin/atlw3/aq.cgi?lang = e

Table V. Continued.

Organisations	URL
GEOnet Names Server by NIMA	http://164.214.2.59/gns/html/index.html
Getty Thesaurus of Geographic Names	http://www.getty.edu/research/tools/vocabulary/ tgn/index.html
Global Gazetteer	http://www.calle.com/world/
Infoplease Atlas	http://www.infoplease.com/atlas/mapindex.html
Introduction to Geographic Names by NIMA	http://gnpswww.nima.mil/geonames/GNS/
Map.com - Online World Atlas	http://www.maps.com/explore/atlas/
MapBlast	http://www.mapblast.com/myblast/index.mb
MapQuest	http://www.mapquest.com/
Multimap Interactive Atlas	http://uk2.multimap.com/
Nations and their administrative divisions	http://stephen.walsh.net/admindiv.html
Principal cities and agglomerations	http://www.citypopulation.de/cities.html
University of Iowa	http://www.cgrer.uiowa.edu/servers/servers_references.html
University of Texas: A&M Map Library	http://www.lib.utexas.edu/Libs/PCL/Map_collection/ Map_collection.html

Acknowledgements

A. Martin Diaz, O. Nordbeck, T. Ton-That, B. Widmer for their contribution to the development of geo-spatial datasets; F. Mouton for his help in the mathematical modelling.

References

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M: 1998, Crop evapotranspiration – Guidelines for computing crop water requirements, *FAO Irrigation and drainage paper* **56**, http://www.fao.org/docrep/X0490E/x0490e00.htm#Contents.

Artan, G. A., Restrepo, M., Asante, K., and Verdin, J.: 2002, A Flood Early Warning System For Southern Africa, *Pecora 15/Land Satellite Information IV Papers*, http://www.isprs.org/commission1/proceedings/paper/00030.pdf.

Blaikie, P., Cannon, T., Davis, I., and Wisner, B.: 1996, *At Risk: Natural Hazards, Peoples Vulnerability and Disasters*, Routledge, London and New York.

Bolt, B. A., Horn, W. L., Macdonald, G. A., and Scott, R. F.: 1975, *Geological Hazards,* Springer-Verlag, Berlin.

Clarke, J. I. and Rhind, D. W.: 1992, *Population Data and Global Environmental Change*, IISC/UNESCO, Paris.

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

FAO: 2001, Asia FIVIMS For Disaster Preparedness, In: *Report of the FAO Asia-Pacific Conference on Early Warning, Prevention, Preparedness and Management of Disasters in Food and Agriculture*, Chiangmai, Thailand, Annex XIII, http://www.fao.org/DOCREP/005/AC120E/AC120E00.HTM.

Funk, C., Michaelsen J. Verdin, J., Artan G., Husak, G., Senay, G., Gadain, H., and Magadazire, T.: 2003, The collaborative historical African Rainfall Model: description and evaluation, *Int. J. Climatol.* **23**: 47–66, http://www.environmental-center.com/magazine/wiley/0899-8418/pdf3.pdf.

Guha-Sapir, D. and Below, R.: 2002, The quality and accuracy of disaster data: a comparative analyses of three global data sets, technical document, The ProVention Consortium and the Disaster Management Facility, The World Bank, WHO Centre for Research on the Epidemiology of Disasters, University of Louvain School of Medicine.

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

IFRC (International Federation of the Red Cross and Red Crescent Societies): 2003, *World Disasters Report 2003*, Kumarian Press, Bloomfield, CT, http://www.ifrc.org/publicat/wdr2003/.

Kawasumi, H.: 1951, Measures of earthquake danger and expectancy of maximum intensity throughout Japan as inferred from the seismic activity in historical times, *J. Soc. Precis. Eng*. **29**, 469–482.

Mouton, F., Nordbeck, O.: 2003, Cyclone database manager : a tool for converting point data from cyclone observations into tracks and wind speed profiles in a GIS, University of Geneva, http://www.grid.unep.ch/activities/earlywarning/download/cylconesmodeling.pdf.

Newhall, C. G. and Self, S.: 1982, The volcanic explosivity index (VEI): An estimate of explosivity magnitude for historic volcanism, *J. Geophys. Res*. **87**, 1231–1238.

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 prepared for the UNDP/BCPR, Geneva, Switzerland.

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

Schloemer, R.W.: 1954, Analysis and synthesis of hurricane wind patterns over Lake Okechobee, *Florida. Hydromet. Rep*. **31**, Dept. of Commerce, Washington, DC.

Simkin, T. and Siebert, L.: 1994, *Volcanoes of the World*, Geoscience Press, Tucson, AZ.

Smith, K.: 1996, *Environmental Hazards, Assessing Risk and Reducing Disaster*, Routledge, London and New York.

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

UNDP: 2004, *Reducing Disaster Risk. A Challenge for Development*, UNDP, Bureau for Crisis Prevention and Recovery, New York, http://www.undp.org/bcpr/disred/rdr.htm.

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

Verelst, Luc: 1999, Recommendations for Geo-referencing EM-DAT Data Base, Consultancy Report prepared for the Asia FIVIMS Project and the CRED, Rome.