Nat Hazards (2014) 70:135–157
DOI 10.1007/s11069-012-0312-4

ORIGINAL PAPER

The need for data: natural disasters and the challenges of database management

Angelika Wirtz · Wolfgang Kron · Petra Löw · Markus Steuer

Received: 19 August 2011 / Accepted: 23 July 2012 / Published online: 28 August 2012

Abstract Hundreds of natural catastrophes occur worldwide every year—there were 780 loss events per year on average over the last 10 years. Since 1980, these disasters have claimed over two million lives and caused losses worth US$ 3,000 billion. The deadliest disasters were caused by earthquakes: the tsunami following the Sumatra quake (2004) and the Haiti earthquake (2010) claimed more than 220,000 lives each. The Great East Japan Earthquake of 11 March 2011 was the costliest natural disaster of all times, with total losses of US$ 210 billion. Hurricane Katrina, in 2005, was the second costliest disaster, with total losses of US$ 140 billion (in 2010 values). To ensure that high-quality natural disaster analyses can be performed, the data have to be collected, checked and managed with a high degree of expertise and professionality. Scientists, governmental and non-governmental organisations and the finance industry make use of global databases that contain losses attributable to natural catastrophes. At present, there are three global and multi-peril loss databases: NatCatSERVICE (Munich Re), Sigma (Swiss Re) and EM-Dat (Centre for Research on the Epidemiology of Disasters). They are supplemented by numerous databases focusing on national or regional issues, certain hazards and specific sectors. This paper outlines the criteria and definitions relating to how global and multi-peril databases are operated, and the efforts being made to ensure consistent and internationally recognised standards of data management. In addition, it presents the concept and methodology underlying the NatCatSERVICE database, and points out the many challenges associated with data acquisition and data management.

A. Wirtz (✉) · W. Kron · P. Löw · M. Steuer
Munich Reinsurance Company, Koeniginstrasse 107, 80802 Munich, Germany
e-mail: awirtz@munichre.com

W. Kron
e-mail: wkron@munichre.com

P. Löw
e-mail: ploew@munichre.com

M. Steuer
e-mail: msteuer@munichre.com

Keywords Natural disasters · Natural catastrophes · Database · Loss data · Economic losses · Insured losses · Natural perils · Natural hazards

1 Introduction

Loss data relating to current and historical natural catastrophes can be used for a variety of analyses and purposes, and are requested by numerous scientific institutes, researchers, government agencies, UN and EU organisations and NGOs, as well as by the insurance and financial sectors. They are necessary for risk assessment and for measuring the progress towards achieving the goal of the Hyogo Framework of Action: the substantial reduction in catastrophe losses (GRIP 2011). The financial and insurance sectors use loss data to perform risk analyses and develop risk transfer products. Furthermore, disaster loss data enable global or regional trends in the consequences of different natural perils to be identified, and the resulting knowledge may lead to the initiation of loss prevention measures. It is, therefore, exceedingly important for the underlying data to be of the best possible quality and for this standard of quality to be consistent throughout the available data series. To achieve this, catastrophic events must be analysed carefully and professionally by the institutions collecting the loss data, and then registered in accordance with a framework of defined criteria, regardless of whether the events are past or current. Only if these criteria are met can users trust the data on which they base their analyses.

At present, there are three global and multi-peril database operators and several hundred national, regional, hazard-based, and specialised sector-based databases worldwide. This paper discusses whether, and subject to what conditions, data from global and national databases can be combined or compared, in order to obtain maximum information.

Clear and explicit standards, methodologies and definitions are needed to combine, merge and supplement data. The Centre for Research on the Epidemiology of Disasters (CRED) and Munich Re have been endeavouring to establish international standards for many years. In 2002, CRED carried out an analysis of natural catastrophe losses using data from three databases (Guha-Sapir and Below 2002). This analysis highlighted the problems resulting from the use of different peril terminologies and definitions. As a consequence, CRED and Munich Re together with Swiss Re, UNDP (United Nations Development Programme) and ADRC (Asian Disaster Reduction Center) developed an internationally recognised standard for disaster category classification and peril terminology (Below et al. 2009). This disaster category classification and peril terminology are described in the following section.

The objectives of this paper are to give an overview of existing disaster loss databases and outline their input criteria. It discusses the need to apply agreed standards and methodologies to disaster loss data collections in order to increase the transparency for the users, and shows the current status of agreed standards already implemented. In order to provide a deep insight into a disaster loss database, the paper presents in detail the structure, methodology and range of data of one example database, the NatCatSERVICE, and shows the challenges that arise from data collection and management. This is followed by an analysis of “great natural catastrophes 1950–2010”. At the end, it suggests next steps that should be taken in order to further increase the quality of disaster loss data in global and regional databases.

2 Global databases

2.1 Overview

The operators of the three electronic global loss databases have continuously and comprehensively collected data covering the entire range of natural hazards for many years. They distinguish between

- geological/geophysical events,
- meteorological and hydrological events,
- (so-called) climatological events and
- other events (e.g. space weather, biological epidemics).

Sigma, a database of man-made and natural catastrophe losses set up by reinsurer Swiss Re, has published statistical analyses in annual publications since 1970. Munich Re's NatCatSERVICE was established in 1974 and built upon an already existing physical archive of historical losses. It only covers catastrophes that are caused by natural hazards (Munich Re 2003). The Centre for Research on the Epidemiology of Disasters (CRED) at Louvain University in Belgium started its Emergency Events Database (EM-Dat) in 1988 with the initial support of the World Health Organisation (WHO) and the Belgian Government (CRED 2010). The focus is on humanitarian aspects, while the two reinsurers' main interest lies in accurate numbers of material losses. However, all three databases include the financial and the humanitarian loss parameters.

In addition, there are a number of databases that have systematically collected and stored data from around the world, but have concentrated on specific hazards and do not simultaneously deal with all natural perils. They usually also do not focus on losses but on a complete record of events and their intensities. To mention but a few: The USGS database on earthquakes (http://earthquake.usgs.gov/earthquakes/eqarchives/epic/), the Dartmouth Flood Observatory at the University of Colorado (http://floodobservatory.colorado.edu/), the tropical cyclone database of the Earth Observation Research Center of the Japan Aerospace Exploration Agency (http://sharaku.eorc.jaxa.jp/TYP_DB/index_e.shtml) and the Smithsonian Institution on volcanic eruptions (http://www.volcano.si.edu/).

2.2 Definition of natural catastrophe and entry criteria

The different foci of the database operators bring about differences in defining the term "natural catastrophe" and specifying entry criteria. Sigma uses strict thresholds as entry criteria: a loss data set is created if a given annual inflation-adjusted economic loss (2010: US$ 86.5 million) and/or 20 fatalities/people reported missing, and/or 50 people injured and/or 2,000 homeless are reached or exceeded (Sigma 2011). EM-Dat defines a disaster as "a situation or an event which overwhelms the local coping capacity, necessitating a request to a national or international level for external assistance; an unforeseen and often sudden event that causes great damage, destruction and human suffering" (CRED 2010). For a disaster to be entered in the database, again thresholds are applied, as at least one of the following four criteria must be met: (1) ten or more people reported killed; (2) 100 or more people reported affected; (3) declaration of a state of emergency; and/or (4) call for international assistance (Below et al. 2009). The general entry criteria of NatCatSERVICE are lower. A loss data set is created as soon as harm to humans (fatality, injury, homelessness) happens or property damage occurs. However, the events are classified in six catastrophe classes (Category 1 to Category 6), depending on the severity of the monetary or humanitarian impact: from a natural occurrence with very low economic impact (Category 1) to a "great natural catastrophe" (Category 6) (Munich Re, 2006).

2.3 Outline of the global multi-peril databases

2.3.1 Sigma

Swiss Re's database includes both man-made and natural disasters. The entries go back to 1970, and roughly 300–350 new events are added each year. In addition to metadata such as peril and date and location of the event, the data include overall and insured losses. Information on the victims is also recorded (fatalities and people missing, injured or homeless). The sources of the database are mainly primary insurance and reinsurance periodicals, internal reports, Lloyd's list, online databases and daily newspapers. Annual lists of all the individual events are published in the periodical "Sigma—Natural catastrophes and man-made disasters". Published data can be downloaded from the Swiss Re website but there is no further public access to the database. For some projects, raw data are provided.

2.3.2 Em-Dat

Em-Dat is operated by CRED at the Catholic University of Louvain, Belgium. It encompasses technical (man-made) and natural disasters. Natural disasters also include biological events, such as epidemics or insect infestation. The loss information dates back to 1900 and about 300 events are added each year. In addition to metadata, data records primarily encompass humanitarian data, such as fatalities and missing persons, injured, homeless or evacuated. Loss data (overall losses and insured losses) are mainly based on information from UN agencies, government offices, IFRC (International Federation of Red Cross and Red Crescent Societies), research organisations, insurance periodicals (Lloyd's List) and reinsurance publications. The database is publicly accessible, but registration is required. All essential data are published once per year (Annual Statistical Review: Number and Trends).

2.3.3 NatCatSERVICE

Munich Re's database encompasses natural catastrophes of all kinds. Historical information dates back to AD 79 (eruption of Mount Vesuvius). Data suitable for analytical evaluation are available for great natural catastrophes (Category 6) as of 1950 and for all events worldwide from 1980. About 800 events are added each year. In addition to metadata, the data records include insured and economic losses. These are mainly obtained from internal company reports and expert appraisals. Further sources include insurance associations, scientific sources, governmental and non-governmental organisations. Each data record also includes a systematic description of the event. The database is not publicly accessible. Selected analyses and tables are available from a download centre, which can be accessed upon registration. Essential data are published once per year in a review of natural catastrophes (Topics Geo Review of Natural Catastrophes). For specific research projects, raw data are provided.

3 National and regional databases

3.1 Overview

Ideally, each country should have its own national database of natural disasters. However, such databases exist in a few countries only and are often operated not under government supervision, but through the initiative of private or non-governmental institutions, usually in response to a huge disaster. Following the tsunami in 2004, for example, the countries affected found that reliable, systematically collected historical loss data were not available. As a consequence, the development of national—and in some cases regional—natural catastrophe databases in these countries was initiated and organised by a United Nations Development Programme (UNDP), in 2005. UNDP organised training courses and provided technical support and guidelines for establishing and institutionalising such databases.

As a result of this initiative, 25 country profiles and data sheets from 21 national databases and four regional databases are now available at the National Loss Data Observatory of the Global Risk Identification Programme (GRIP) (http://www.gripweb.org/grip). The National Loss Data Observatory is defined as a sustainable institutional arrangement for the systematic and homogeneous collection of data on losses from natural disasters. GRIP will expand and improve the basis of evidence for disaster-related losses by promoting and supporting the systematic collection of loss data in databases to make them available for various analyses. The final goal of GRIP is to implement (50–100) national disaster loss data observatories in highly disaster-prone countries around the world (GRIP 2011).

3.2 From national to global databases

With the availability of national and regional databases, the following question arises: To what extent is it possible to aggregate or comparatively analyse data records from these databases? Firstly, the databases in the National Loss Data Observatory are of different lengths: for instance, the records of the regional database in Florida cover the period up to 2001, while those of the national databases in Indonesia and Sri Lanka end in 2008 and 2007, respectively. Secondly, different designations are used for the same perils. The Florida database, for instance, includes the perils "storm, strong wind, hurricane and tornado", while the Sri Lanka database refers to "gale, storm, strong wind, tornado and cyclone". For the preparation of the ProVention Consortium's GRIP Workshop on the Compilation of Reliable Data on Natural Disaster Occurrence and Impact, UNDP and CRED undertook an analytical review of existing historical databases (Tschoegl et al. 2006). A selection of international, regional, national and sub-national databases, as well as peril-based databases were analysed. The results show that the databases are very difficult or even impossible to compare, mainly due to the usage of different definitions and peril terminologies.

A very similar conclusion was reached in a study by Melanie Gall (Stephenson Disaster Management Institute, Louisiana State University, Baton Rouge, Louisiana) who compared data records in the EM-Dat, NATHAN (a sub-set of NatCatSERVICE), SHELDUS and Storm Events databases (Gall et al. 2008). The main factors that prevented comparative analyses and led to misinterpretations were hazard bias (wrong or uncertain representation and distribution of losses between different hazard types), threshold bias (under-representation of minor events) and accounting bias (under-reporting of indirect, uninsured and other losses). To overcome these problems, the study recommends that clear guidelines and standards be created, for example, to establish a common naming convention for similar phenomena.

4 International standards are essential

As a direct consequence of these analyses, CRED and Munich Re established a collaborative initiative, to implement a common "Disaster Category Classification and Peril Terminology for Operational Databases" in 2007. This common classification was developed by several expert meetings and working groups involving CRED, Munich Re, Swiss Re, the Asian Disaster Reduction Center (ADRC) and the United Nations Development Programme (UNDP). The result, an internationally recognised disaster category classification and peril terminology, represents an important step in the development of international standards (Below et al. 2009). The classification differentiates between two generic disaster categories: natural and technological disasters. The following chapters refer to natural disasters only. The natural disaster category is divided into six disaster groups (geophysical, meteorological, hydrological, climatological, extraterrestrial and biological), each of which is further divided into sub-groups, as shown in Table 1. This classification was presented at the 2007 Global Platform in Geneva, and discussed and agreed by representatives from WMO, IFRC and OCHA. The EM-Dat and NatCatSERVICE databases have used this peril classification since 2009, and it has also been implemented by other regional databases like Sheldus, a national disaster database in the United States. Consequently, the comparability and clarity of data records has already been vastly improved so that now, for instance, the effects of a tsunami are clearly classified under "Geophysical events" and not, as sometimes in the past, under "Flood events". It is very important to have this standardised form of classification insofar as global and national database analyses, for example, show which peril groups account for the heaviest monetary losses or the highest death tolls. Policy-makers also use this information when deciding on disaster risk reduction measures.

The following section describes how a global database is organised and takes as an example Munich Re's NatCatSERVICE database. It describes why NatCatSERVICE was established and gives an overview of the most important attributes used in describing a natural loss event. Areas where there is need for better or additional information are also discussed. In general, most of the information and descriptions given also apply to the other global databases, EM-Dat and Sigma.

5 The NatCatSERVICE: history and methodology

5.1 History

Natural catastrophes have always played a significant role for reinsurance companies throughout the world. After being set-up in 1863 and 1880, respectively, the two large reinsurance companies Swiss Re and Munich Re faced their first major challenge with the San Francisco earthquake of 1906: in the early morning of 18 April 1906, an earthquake struck San Francisco, destroying one-fifth of the city—28,000 buildings in total, most of which fell victim to a fire that burned for four full days and nights. Over 3,000 people died. The overall economic losses came to US$ 524 million, of which US$ 200 million were insured. Munich Re paid out claims totalling US$ 11 million (original value), which was

Table 1 Grouping of disasters

Hazard family	Main event	Sub-peril
Geophysical	Earthquake	Earthquake "ground shaking"
		Earthquake "fire following"
		Tsunami
	Volcanic eruption	
	Mass movement (dry)	Subsidence
		Rockfall
		Landslide
Meteorological	Storm	*Tropical storm*
		Hurricane, typhoon, cyclone
		Extratropical storm
		Winter storm, blizzard/snowstorm
		Convective storm
		Severe storm, thunderstorm, tornado, hailstorm
		Local storm (orographic storm)
		I.e. Foehn, Bora Bora, Mistral
Hydrological	Flood	General flood
		Flash flood
		Storm surge
		Glacial lake outburst flood
	Mass movement (wet)	Avalanche
		Landslide
Climatological	Extreme temperature	Heat wave
		Cold wave/frost
		Extreme winter conditions
	Drought	
	Wildfire	Forest fire, bush fire, grassland fire, brush fire
Extraterrestrial	Meteorites, asteroids	
	Space weather	
Biological	Epidemic	
	Insect infestation	
	Animal stampede	

14 % of its net premium income at the time. For Munich Re, this still represents the biggest single loss from a natural disaster relative to its premium income.

In the following six decades, Munich Re collected data on natural disaster losses in a relatively haphazard and non-centralised fashion, usually based on the country departments dealing with the respective losses. The natural catastrophe database, NatCatSERVICE, was born out of these historical archives in 1974, when a geoscience department was founded. From then, all data relating to natural hazards and actual loss events were collected systematically and organised in a single archive covering all countries and all hazards. The first electronic version was launched in 1988. The following sections contain a description of the composition of the individual data items and the methods applied in researching the data, determining the losses and applying quality validation measures.

Table 2 Overview of data

Years	Number of data sets
0079–0999	35
1000–1499	200
1500–1899	1,000
1900–1949	1,200
1950–1979	2,900
1980–2010	23,300

5.2 Data records

The database currently contains 29,000 data records (as at December 2010). The entry that goes farthest back in history concerns the eruption of Mount Vesuvius on 24 August in the year AD 79, which destroyed the cities of Herculaneum and Pompeii. In all, 1,235 disasters are recorded between AD 79 and 1899. These records on historical events cannot be used for statistical analyses, as they are incomplete and only available in the case of certain regions and events. For instance, earthquakes in China during the last two thousand years are well documented, while systematic records on North America have only been kept since the mid-nineteenth century.

Great and devastating natural catastrophes are included from 1900 onwards. For all natural disasters worldwide, reasonably complete and consistent data record exist as of 1980, permitting analyses at global, continental and national level. For some countries (Germany, Japan, the USA), complete records are available going back to 1970 (Table 2).

5.3 Catastrophe classes

So that changes in event severity can be analysed, and events are allocated to different catastrophe classes. Depending on their monetary or humanitarian impact, the documented events are classified into seven classes, ranging from a natural occurrence with negligible economic and human impact to a great natural catastrophe (Munich Re 2006). Events from catastrophe class 0 are disregarded for analytical and statistical considerations, but nevertheless included in the database, because a small event (e.g. a minor wildfire) can potentially develop into a huge natural disaster, which then necessitates a change of catastrophe class. At the upper end of the scale, there are the "great natural catastrophes", which constitute event class 6. In line with definitions used by the United Nations, a "great natural catastrophe" clearly overstretches the affected region's ability to help itself and interregional or international assistance is consequently required. As a rule, this will be the case when thousands are killed and hundreds of thousands are made homeless, or if the overall loss reaches exceptional dimensions, depending on the economic capacities and conditions of the country concerned. These great natural catastrophes can be used for long-term analyses starting in 1950, as major disasters of this type have always been reported in detail and the analyses are not distorted by reporting bias. Two examples of this type of natural catastrophe are the eruption of Mount Lamington in Papua, New Guinea, in 1951, in which 3,000 people lost their lives, and the major storm surge that caused over 2,300 deaths in Europe in 1953. As with all other great natural catastrophes in the period from 1950 to 1980, both events were described and documented by a number of different and excellent sources.

Catastrophes are classified according to number of deaths and economic losses. If only one criteria is met (i.e. either number of deaths or overall loss), this is sufficient for classification. Classification also has to take into account the different economic conditions involved. Thus, an event that produced losses of the order of US$ 50 million in 1990 constitutes a severe (Category 3) catastrophe in the USA and a devastating (Category 5) catastrophe in Bangladesh. The monetary thresholds are inflated using the consumer price index of the relevant country, and then converted into US dollars as the common denominator (Tables 3, 4).

5.4 Event data

The objective of making an entry in the database is to describe a catastrophe in as much detail as possible. On the other hand, actual available information and data often fail to fully meet the requirements. Nowadays, disaster databases are expected to provide more and more granular data. In the NatCatSERVICE database, a full entry record consists of up to 200 attributes. The following are the most important:

- Event identification number
- Event categorisation
  - Hazard family, main event group, sub-peril type (see Table 1)
  - Name of catastrophe (e.g. “Hurricane Katrina”)
  - Associated perils and consequences (e.g. famine following a drought or tsunami following an earthquake)
- Geographical information
  - Continent, subcontinent, region, country, state/province, town or city, coordinates (longitude/latitude)
- Date and duration

Beginning and end of a disaster are strictly connected to the occurrence of losses. For instance, the beginning of a hurricane is defined as the date on which the first loss occurs; the hurricane’s meteorological genesis and history are not considered. The end of a disaster is deemed to be when the last direct loss (in terms of monetary loss or people affected) occurs. In the case of floods, for instance, there can be still high water levels on rivers, but if no further damage is reported, the day on which the final direct loss is recorded is entered as the final day. It is important to document the beginning and end of a disaster in order to separate events that may overlap in time and/or geographical area. This allows the number of events to be counted as a whole or by region.

However, it is difficult to apply this definition to natural hazards that are of slow onset (i.e. droughts) because droughts develop gradually, and their effects accumulate slowly over a substantial period of time, and may linger on for years after the termination of the event. As far as possible, the start dates for drought disasters in NatCatSERVICE represent the month and year in which the drought began—typically, the start is that officially declared by the affected governments or by drought-related organisations. The end date is identified using the drought monitors of the different regions (e.g. the US National Drought Mitigation Center, the Global Drought Monitor of the UCL Department of Space or the Beijing Climate Center drought monitor).

- Victims

Table 3 Criteria of catastrophe classes for the United States (by way of an example)

	Catastrophe classes	Loss profile	Overall losses (in million)				And/or fatalities
			1980s*	1990s*	2000s*	2010	
0	Natural event	No property damage	–	–	–	–	None
1	Small-scale loss event	Small-scale property damage	US$ 0.63	US$ 0.91	US$ 1.18	US$ > 1.33	1–9
2	Moderate loss event	Moderate property and structural damage	US$ 5.08	US$ 7.28	US$ 9.40	US$ > 10.6	>10
3	Severe catastrophe	Severe property infrastructure and structural damage	US$ 29	US$ 42	US$ 54	US$ > 61	>20
4	Major catastrophe	Major property, infrastructure and structural damage	US$ 114	US$ 164	US$ 212	US$ > 239	>100
5	Devastating catastrophe	Devastating losses within the affected region	US$ 305	US$ 437	US$ 564	US$ > 636	>500
6	Great natural catastrophe "GREAT disaster"	Region's ability to help itself clearly overtaxed, interregional/international assistance necessary, thousands of fatalities and/or hundreds of thousands homeless, substantial economic losses (UN definition). Insured losses reach exceptional orders of magnitude					

* Decade average—categorisation uses annual thresholds, adjusted by US CPI

Table 4 Number of entries per catastrophe category

Natural catastrophes 1980–2010	
Catastrophe category	Number of entries
0 + 1	8,400
2	8,650
3	3,200
4	1,700
5	1,000
6	295

- They were as follows: death toll and missing, injured, homeless, evacuated or displaced people. It is important to know that the usage of descriptive words can lead to misunderstandings. For example, the term "victims" is sometimes used to indicate number of deaths, sometimes "deaths and those reported missing" and sometimes "those affected". Other indicators are even more ambiguous and imprecise. For instance, "people affected" can mean "actually suffering from damage or harm caused by a natural event" or simply "living in the area where the event took place". Reliable information on the number of people affected and exactly how they are defined is extremely useful for enabling aid organisations, government agencies and other decision-makers to effectively apply disaster risk reduction measures. In order to avoid uncertainties in terms of terminology and definitions, CRED initiated the "Expert Working Group on Human and Economic Disaster Impact Characterization" in 2010. The Group's main objectives are to discuss the current status of applied disaster loss definitions used by database holders and field agencies and to reach a consensus on a set of basic definitions for quantifying the human losses. Once the human indicator definitions have been agreed (probably by the end 2012), the global CRED, Sigma and Em-Dat databases will implement them.

– Damage and affected economic sectors

- Number and type of buildings damaged or destroyed (e.g. public buildings, hospitals)
- Vehicles, vessels affected
- Type of infrastructure affected (e.g. bridges, roads)
- Arable farming, livestock, aquaculture
- Transport/marine (e.g. shipping, oil platforms)
- Industrial losses
- Ecological damage (e.g. oil spills)

– Losses

- They were as follows: overall losses with the development of the loss estimates over time following the event and precise identification of the source of each update; losses in national currency, US$ and Euros; automatic currency conversion at the rate applicable on the last day of the respective month; and losses in original values. Aggregated annual losses or individual event losses are adjusted for inflation via the country's CPI (consumer price index) in the local currency and then converted into US$.
- Insured losses (as above)

- Scientific data
  - Characteristics of the corresponding natural event (magnitude, intensity, wind speed, precipitation, etc.)
- Other information
  - Impact on GDP
  - Classification of the country concerned by World Bank income group (high-income, upper-middle income, lower-middle income, low-income countries)
- Brief description of the event with all relevant details. This provides a quick overview of the scale and impact of the catastrophe.

These items describe the main characteristics of a disaster event. Of course, not every individual attribute is encountered in every case. The crucial parameters for statistical analyses are death toll and monetary losses.

5.5 Multi-peril and multi-country events

Catastrophic events are entered in the database according to triggering natural hazard event or main cause of loss. A tsunami triggered by an earthquake, for example, would be listed in the database as a "geophysical event/earthquake/tsunami". This permits analyses at multiple thematic levels, be it a general look at the frequency and the losses due to all types of geophysical events or a more detailed look at one type of geophysical events, for instance, tsunamis. In some cases, the loss event may be assigned to different main groups. A flash flood, for instance, may be part of a convective storm that also features hail, wind gusts, etc., and is therefore classified as "meteorological event/convective storm, flash flood". However, it may also occur in the form a single event, which qualifies it for the category "hydrological event/flood/flash flood".

Only on the basis of an unambiguous categorisation of events is a comparison between entries in different databases possible. This is illustrated by the comparison of all events between 2000 and 2008 (Fig. 1) in the EM-Dat and NatCatSERVICE databases, broken down by standardised event group. They are broadly similar, the only exception relating to meteorological and hydrological events, where notable differences are found. The reason for this is that the NatCatSERVICE database contains a larger number of storm events registered on account of the financial losses involved, but not meeting the humanitarian entry criteria of EM-Dat—namely minimum number of fatalities and/or declaration of a state of emergency.

Events resulting in financial or humanitarian losses in several countries are recorded in the database as "region events", an all-embracing data record item containing the data for all affected countries. Detailed information is also available in the country records. This hierarchy permits analyses at event level (e.g. number of events and losses due to all hurricanes) and national level (e.g. number of events and losses due to all hurricanes in the USA). Multi-country events are only counted once in the database.

5.6 Economic and insured losses

Financial loss is an important parameter in loss databases. It is subdivided into two categories: insured losses and economic losses. Figures for insured losses are very reliable

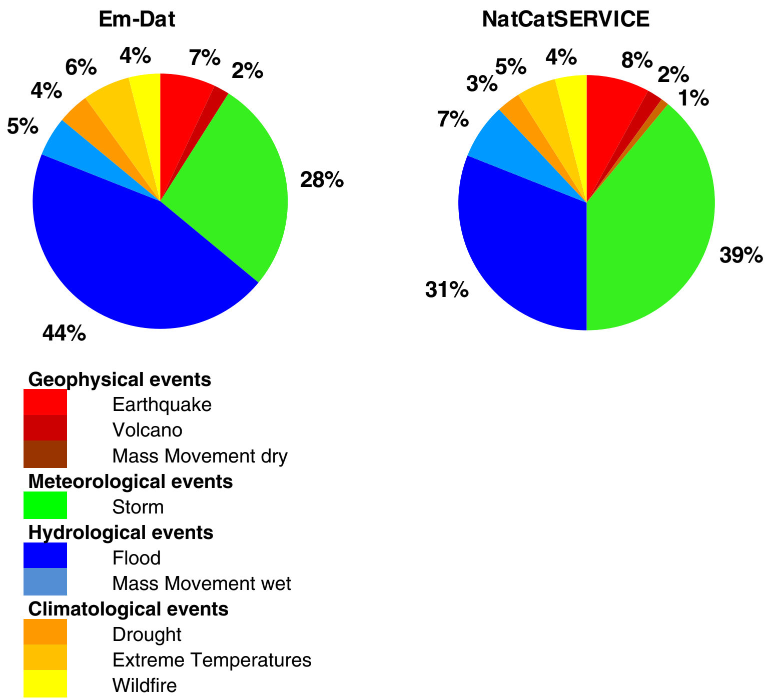

Fig. 1 Comparison of events in EM-Dat and NatCatSERVICE: percentage distribution by main disaster event, 2000–2008

because they reflect claims actually paid by insurance companies. Assessing economic losses is, however, more complex.

Media reports very often publish loss estimates immediately after a catastrophe. However, such estimates are very unreliable. The initial damage estimates are often overstated, sometimes in the hope of mobilising more emergency aid. But, in other instances, the loss numbers develop over time, as the catastrophe evolves. Examples are Hurricane Katrina (2005) and the Pakistan flood of 2010, which progressively involved more and more regions. Governments, the UN, the EU, universities, the World Bank and other development banks generally only elaborate on detailed loss analyses following particularly large or severe natural catastrophes. Therefore, in the case of smaller events, the figures published at the beginning are often the only ones available, and widely accepted as correct.

The term "economic loss", moreover, does not have a uniform definition. It is important to differentiate between "direct losses", "indirect losses" and "secondary/consequential losses". While there is some uncertainty as to what exactly is to be understood by these three classes of loss, NatCatSERVICE defines them in the following way:

Direct losses are immediately visible and quantifiable (loss of homes, household property, schools, vehicles, machinery, livestock, etc.) (Munich Re 2001). They are always calculated on the basis of replacement and repair costs. However, it is difficult to estimate the value of historically significant areas and cultural heritage assets. Also, damaged structures (e.g. dykes) are upgraded—when repaired or replaced—to a higher safety and therefore value level. While actual losses should refer to the items damaged, the figures given in fact often refer to the costs incurred.

Unlike direct losses, which are immediately apparent, *indirect losses* are not generally obtainable close to an event. They arise as a result of the physical destruction of assets. Examples of indirect losses include higher transport costs due to damaged infrastructure, loss of jobs, production losses or loss of rental income. Two types of insured indirect loss are business interruption (where production is halted because the insured's plant is not operational—if, for example, it is flooded) and contingent business interruption (where production is halted because a supplier's plant is flooded or where finished products or parts cannot be delivered because the recipient company is not operational).

Consequential losses (secondary costs) are also not available directly after the event, but often only months or years after. Consequential losses concern the economic impact of natural disasters, for instance, in terms of diminished tax revenues, lower economic output, reduced GDP or lower currency exchange rates.

To measure the full extent of a natural catastrophe, all three loss indicators would be needed, and should be included in loss databases as three separate loss information items. Due to the insufficient availability of information on indirect and consequential losses, only direct economic losses are currently documented in NatCatSERVICE. Information on indirect and consequential losses is only documented if the analyses have been carried out by bodies such as ECLAC (Economic Commission for Latin America and the Caribbean), the World Bank or universities. The Forensic Investigations of Disasters Working Group, set up by ICSU/IRDR (the International Council of Science/Integrated Research on Disaster Risk, www.irdrinternational.org), developed methodological guidelines and suggestions for the design and conduct of a set of internationally organised case studies of disasters in 2010. This information can help ensure standardised assessment of the full extent of great natural catastrophes in the future and should definitely be documented in disaster loss databases by way of additional information.

5.6.1 Estimation of direct economic losses

NatCatSERVICE has registered 23,300 loss events since 1980. The number of events where the economic loss was assessed, verified and published by governments, statistical and financial authorities, etc., amounted to only one-third of these events. The same was observed by the CRED. In their EM-Dat database, economic loss figures in the period 1973–2003 were reported for less than 40 % of all natural disasters (Guha-Sapir et al. 2004).

NatCatSERVICE estimates the loss on the basis of insurance claims, information on insurance penetration for the different perils and different countries, and available loss indicators. The loss information only includes direct (tangible) losses with the exception of "business interruption", which is included under insured market loss. Secondary effects and consequential losses are not taken into account. They are, as mentioned above, not yet consistently available. The assessment procedure used to obtain direct economic losses is illustrated by the following two examples:

Example 1: Assessment of the overall loss with known insured losses based on insurance density, a parameter known for all countries and for the various natural hazards. The natural peril concerned, the regions of that particular country affected (urban or rural area, population density, quality of buildings) and the branches of industry which suffered losses are included in the calculation. With this information, the loss can be estimated realistically.

Example 2: If there are no known insured losses, the overall loss is estimated on the basis of the following parameters: type and duration of the natural disaster, region affected (urban or rural area), population density, standard of living, damaged properties, infrastructure and sources of supply, as well as number of victims and fatalities. These data are plotted in a matrix and weighted. The events are then assigned to a catastrophe class. All comparable disasters in the region for which detailed and well-referenced overall loss data are available are additionally filtered with the aid of an approximation process. The events are clustered, and realistic values are obtained for individual units (e.g. average value of a residential building). In this way, the loss attributable to the event can be allocated to an order of magnitude.

5.7 Geocoding

Towns and administrative units affected by natural disasters are documented in the NatCatSERVICE database. Important supplementary information on the towns affected, such as the main cause and relative severity of the damage, is also recorded, if known. All locations are geocoded, that is, their exact geographical position is expressed in latitudinal and longitudinal coordinates.

The analysis of Winter Storm Xynthia (2010) shows that the administrative units affected correlate with the wind field. Thus, the heterogeneity of the natural catastrophe which caused damage, due to storm surge, wind and river flood, can be clearly seen (Fig. 2).

5.7.1 Analysis options

5.7.1.1 Validation of exposure maps Geographical analysis of loss data is an ideal tool for validating exposure maps. Since the data reflect real-life conditions, modelled results can be checked for plausibility. For instance, populated areas that are highly exposed to a particular natural hazard can be identified by plotting the towns and administrative units affected by a certain type of event on a map.

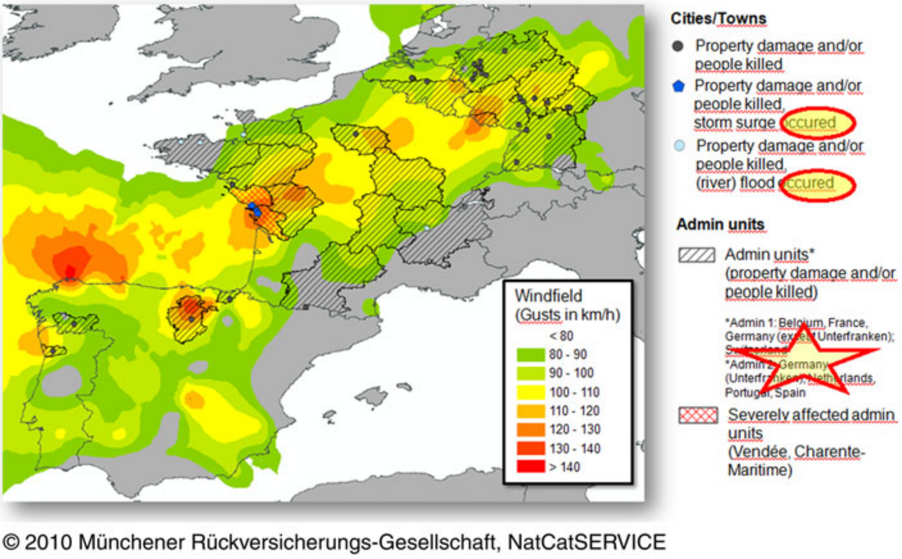

Fig. 2 Winter Storm Xynthia—wind field and loss areas

5.7.2 Higher granularity of loss figures

Loss figures are usually reported in aggregated form for administrative units. If the area affected by a natural catastrophe only accounts for part of this administrative unit, the loss data can be downscaled to a smaller area. The effect of the loss event can be determined in detail from these data and further analyses generated, such as a standardisation of losses or comparison with modelled results.

Hurricane Ike, which occurred in 2008, is a good example of an event with a very high information density (Fig. 3). Once the areas affected by Ike have been defined, the insured market loss—which is known for each state—can be narrowed down precisely into loss areas and broken down into smaller areas.

Geocoding permits numerous analysis options, which would be impossible if location data were recorded without geographical reference. In this way, loss data can be analysed professionally and risk areas identified more effectively.

5.8 Sources and data quality

In recent years, it has become much easier to investigate data on natural events—largely because of the Internet. At the same time, it has become even more important to ensure that sources are robust and sound. NatCatSERVICE employs around 200 sources that have been identified as first rate for a particular region and/or type of event. The main source groups are as follows:

- Insurance industry information,
- Meteorological and seismological services,
- Reports and evaluations by aid organisations or NGOs, governments, the EU, the UN, the World Bank and other development banks,
- Scientific analyses and studies,
- News agencies.

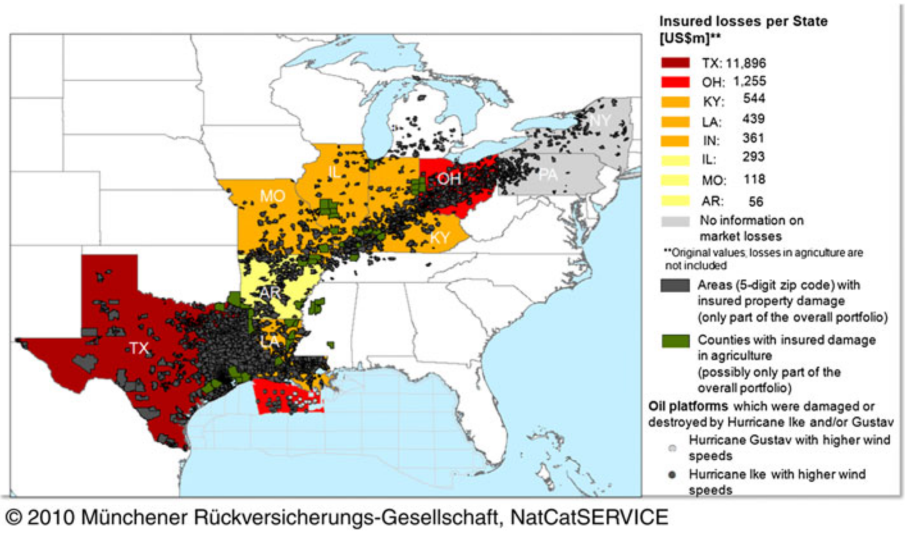

Fig. 3 Loss track of Hurricane Ike in 2008

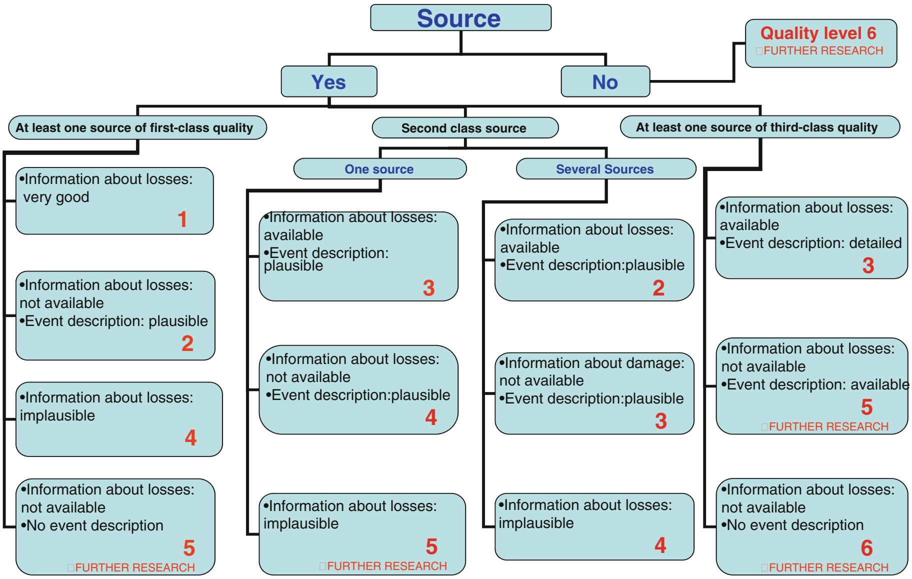

Fig. 4 Quality check—decision-making tree

Despite first-class sources, the analysis process is sometimes fraught with problems. Typical challenges include false reporting, the use of incorrect conversion factors and double counting of casualties. Such data are often copied and further disseminated. Therefore, database operators must test the quality of the figures they obtain. The validation process in the NatCatSERVICE database checks the quality and number of the sources referred to as well as the plausibility of the loss figures and the description of the event. An evaluation system has been developed for this quality check, which assigns every data record to a quality level on a scale from 1 (very good) to 6 (inadequate). Data records on quality levels 4, 5 or 6 do not meet with the quality standards of the database and are not used for analyses (Fig. 4).

5.9 Analyses

The following are intended as examples of possible analyses of the data in a global loss database. The analyses are only based on the great natural catastrophes from catastrophe class 6 (see Table 3: Criteria of catastrophe classes), which means that they claimed more than 2,000 lives and/or left more than 200,000 people homeless and/or caused an overall loss of at least 0.5 % of the gross domestic product and/or rendered the region or country dependent on national or international assistance. Reporting bias can be ruled out for this group of great natural catastrophes as it is very unlikely that events on such a major human and/or financial scale would not have been thoroughly reported (see Chapter 5 3).

In the period from 1950 to 2010, 295 disasters met the criteria for great natural catastrophes (disaster category 6). Roughly 60 % of the events were included solely on account of their economic losses, and around 10 % on account of their humanitarian impact, that is, the number of fatalities or people left homeless. About 30 % of the events

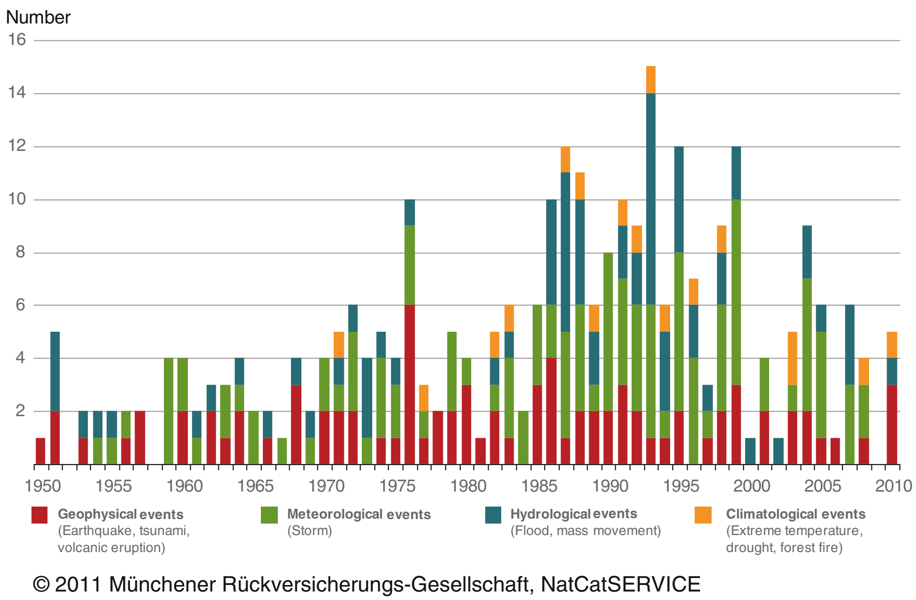

Fig. 5 Great natural catastrophes 1950–2010—number of events

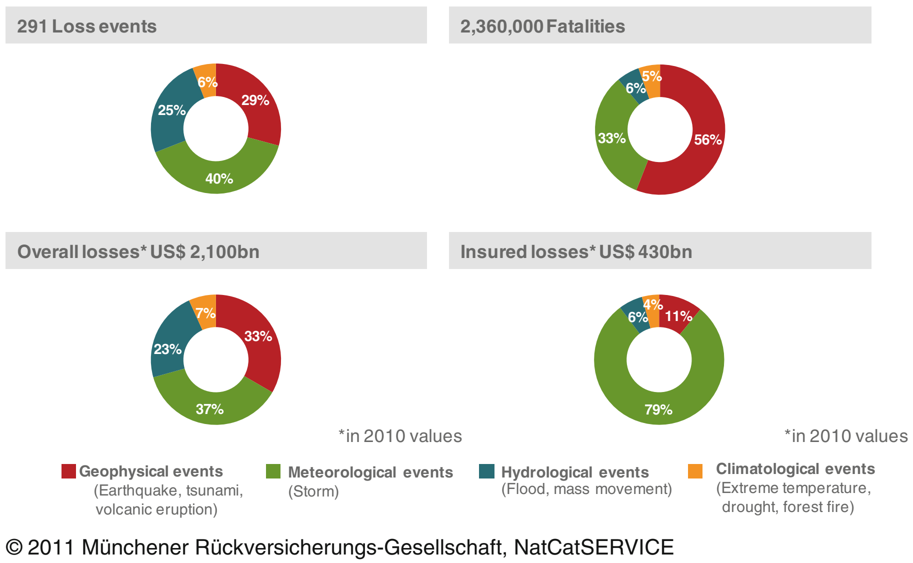

Fig. 6 Great natural catastrophes 1950–2010—percentage distribution of perils

met all the criteria. Since 1950, there have only been 3 years without any "great natural catastrophes": 1952, 1958 and 2009 (Figs. 5, 6).

Broken down by main peril, 29 % of great natural catastrophes were geophysical events, 40 % meteorological, 25 % hydrological, and 6 % climatological events. More

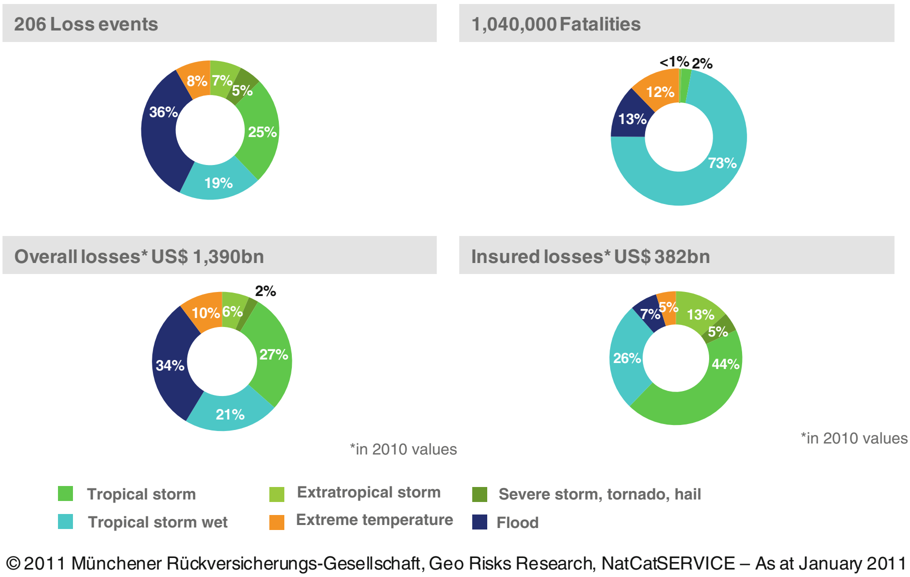

Fig. 7 Great weather catastrophes 1950–2010—percentage distribution of sub-perils

than half of the fatalities were caused by earthquakes. The deadliest earthquakes between 1950 and 2010 were the 1976 Tangshan quake in China (242,000 deaths), the earthquake in Haiti on 12 January 2010 (225,000 deaths) and the 2004 earthquake that triggered a tsunami in Southeast Asia (220,000 deaths). Storm surges can be similarly deadly. Figure 7 shows that storm surges caused by tropical storms account for 73 % of all fatalities. The deadliest events were the storm surges in Bangladesh (1970) with 300,000 fatalities, Myanmar (2008) with 140,000 fatalities and Bangladesh (1991) with 139,000 fatalities.

The costliest natural disaster between 1950 and 2010, in terms of both overall and insured losses, was Hurricane Katrina, which struck the US Gulf Coast in 2005. Very often, however, the highest losses in economic terms are caused by earthquakes: the second, third and fourth costliest disasters between 1950 and 2010 were of geophysical origin. With an estimated overall loss of US$ 210 billion (as at July 2011), the earthquake and tsunami of 11 March 2011 in Japan broke all previous records (Table 5).

A breakdown of the four main perils by continent shows the following picture:

Fatalities — Most fatalities caused by geophysical events occurred in Asia (68 %) and North and Central America (20 %). Asia also accounts for the great majority of fatalities from meteorological events (94 %) and hydrological events (71 %)

Overall losses — More than half of the overall losses caused by earthquakes and volcanic eruptions occurred in Asia (57 %), while North and Central America rank second (18 %) and Europe ranks third (14 %). As far as meteorological events are concerned, North and Central America account for roughly three-quarters (74 %) of all losses. Asia (58 %) was the continent hardest

Table 5 Costliest and deadliest natural catastrophes 1950–2010

Years	Event	Country	Overall losses (US$ million, original values)	Insured losses (US$ million, original values)	Fatalities
2005	Hurricane Katrina	USA	125,000	62,200	1,322
1995	Earthquake	Japan	100,000	3,000	6,430
2008	Earthquake	China	85,000	300	84,000
1994	Earthquake	USA	44,000	15,300	61
2008	Hurricane Ike	USA, Caribbean	38,000	18,500	168
1998	Floods	China	30,700	1,000	4,159
2010	Earthquake, tsunami	Chile	30,000	8,000	521
2004	Earthquake	Japan	28,000	760	46
1992	Hurricane Andrew	USA	26,500	17,000	62
1996	Floods	China	24,000	445	3,048

Years	Event	Country	Fatalities
1970	Cyclone, storm surge	Bangladesh	300,000
1976	Earthquake	China	242,750
2010	Earthquake	Haiti	222,570
2004	Earthquake, tsunamis	Southeast Asia	220,000
2008	Cyclone Nargis, storm surge	Myanmar	140,000
1991	Cyclone, storm surge	Bangladesh	139,000
2005	Earthquake	Pakistan, India	88,000
2008	Earthquake	China	84,000
2003	Heatwave, drought	Europe	70,000
1970	Earthquake, mountain slide	Peru	67,000

hit by hydrological events, followed by Europe (25 %). Half of the overall losses caused by climatological events occurred in North and Central America, while around one-third (31 %) occurred in Asia

Insured losses — Insured losses from geophysical events were mainly located in North and Central America (59 %), while South America and Asia each accounted for 18 %. North and Central America also top the losses list for meteorological events (77 %) and climatological events (90 %), while Europe accounts for the great majority (66 %) of losses caused by hydrological events (Table 6)

In total, the economic loss attributable to the 291 classified great natural catastrophes since 1950 amounts to US$ 2,090 billion. In the time series analysis (Fig. 8), both the economic loss and the insured losses are extrapolated using the applicable nominal consumer price index to adjust the losses in line with general prices. The bars representing losses show the monetary consequences at 2010 prices due to the disasters under exactly the same conditions.

Table 6 Great natural catastrophes 1950–2010—comparison by continent

	Number	Fatalities	Overall losses (US$ million, in 2010 values)	Insured losses (US$ million, in 2010 values)
Africa				
Geophysical events	5	19,097	18,008	12
Meteorological events	1	9	888	0
Hydrological events	3	2,027	2,607	565
Climatological events*	4	0	5,344	0
Total	13	21,133	26,846	577
Asia				
Geophysical events	41	897,683	401,045	8,614
Meteorological events	45	733,507	114,744	26,491
Hydrological events	39	95,057	273,873	3,699
Climatological events	3	342	44,337	1,376
Total	128	1,726,589	833,998	40,180
Australia/Oceania				
Geophysical events	1	13	2,220	1,240
Meteorological events	5	138	6,982	3,263
Hydrological events	1	28	8,872	493
Climatological events	1		3,402	363
Total	8	179	21,476	5,359
Europe				
Geophysical events	15	33,379	99,929	995
Meteorological events	14	645	80,500	46,968
Hydrological events	18	3,559	117,684	16,981
Climatological events	3	126,400	19,162	47
Total	50	163,983	317,275	64,990
South America				
Geophysical events	11	101,440	52,082	8,618
Meteorological events	1	2,412	151	0
Hydrological events	5	31,296	10,704	386
Total	17	135,148	62,937	9,005
North America				
Geophysical events	12	270,182	125,070	27,612
Meteorological events	46	41,921	577,484	262,081
Hydrological events	7	2,185	59,386	3,504
Climatological events	10	627	73,449	16,593
Total	75	314,915	835,390	309,790

* Fatalities caused by famine not included

The analyses of “great natural catastrophes 1950–2010” shown are intended to illustrate ways of examining global loss data and identifying possible trends in terms of peril or geographical area.

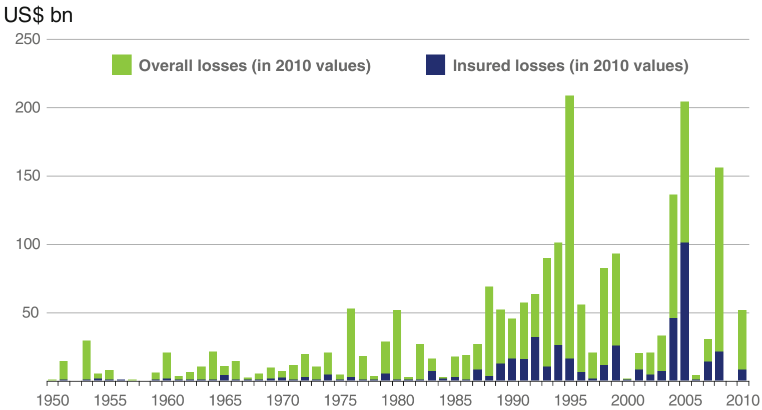

Fig. 8 Great natural catastrophes 1950–2010—overall and insured losses

6 Conclusions

Data on natural disasters are used by a wide variety of organisations, such as governments, non-governmental organisations (NGOs), scientific institutes and the private sector, including the financial sector with the insurance industry. Such data are, in many cases, a valuable instrument in providing background information for better risk assessment and decision-making. It is, therefore, essential that the data provided by the different databases meet the highest quality standards. The methods used to collect data and assess losses must be consistent throughout the entire collection period. Methods must be transparent and data unambiguous for the stakeholders using them. In order to minimise uncertainties, it is important that global and national database providers use common standards and definitions. Fortunately, major steps have already been taken in this direction, one example being the consensus reached on the hierarchy and terminology of natural hazards by CRED, Munich Re, Swiss Re, UNDP and ADRC. Initiatives are currently under way to establish guidelines in respect of geocoding and the definition of human loss indicators. The next steps in enhancing the quality of disaster loss records will have to focus on the exact definition of loss categories such as economic losses, indirect and consequential losses. Although the complexity of economic impact indicators is certainly a challenge, a joint effort is needed involving database operators and data providers, economists and organisations dealing with disaster loss assessment in order that the quality of disaster loss data may be further enhanced.

References

Below R, Wirtz A, Guha-Sapir D (2009) Disaster category classification and peril terminology for operational purposes, Annex II. http:\\emdat.be

EM-Dat (2011) The International Disaster Database at http:\\emdat.be. Accessed 24 July 2011

Gall M, Borden KA, Cutter SL (2008) When do losses count? Six fallacies of natural hazards loss data. American Meteorological Society, June 2009, pp 799–809

Global Risk Identification Programme (GRIP) (2011) Disaster profiles. National data based profiles. http:\\gripweb.org. Accessed 24 July 2011

Guha-Sapir D and Below R (2002) The quality and accuracy of disaster data. A comparative analyses of three global data sets. ProVention Consortium

Guha-Sapir D, Hargitt D, Hoyois P (2004) Thirty years of natural disasters 1974–2003: The numbers, p 38. http:\\emdat.be
Munich Re (2001) Topics geo. Annu Rev Nat Catastrophes 2000:16–21
Munich Re (2003) NatCatSERVICE® a guide to the Munich Re database for natural catastrophes, p 5. www.munichre.com
Munich Re (2006) Topics geo. Annu Rev Nat Catastrophes 2005:14–17
Swiss Re (2011) sigma No. 1/2011, p 35. http:\\swissre.com
Tschoegl L, Below R, Guha-Sapir D (2006) An analytical review of selected data sets on natural disasters and impacts. http:\\emdat.be