



How to deal properly with a natural catastrophe database – analysis of flood losses

W. Kron, M. Steuer, P. Löw, and A. Wirtz

Geo Risks Research, Munich Re, Munich, Germany

Correspondence to: W. Kron (wkron@munichre.com)

Received: 29 March 2011 – Revised: 10 November 2011 – Accepted: 15 November 2011 – Published: 2 March 2012

Abstract. Global reinsurer Munich Re has been collecting data on losses from natural disasters for almost four decades. Together with EM-Dat and sigma, Munich Re's NatCatSERVICE database is currently one of three global databases of its kind, with its more than 30 000 datasets. Although the database was originally designed for reinsurance business purposes, it contains a host of additional information on catastrophic events. Data collection poses difficulties such as not knowing the exact extent of human and material losses, biased reporting by interest groups, including governments, changes over time due to new findings, etc. Loss quantities are often not separable into different causes, e.g., windstorm and flood losses during a hurricane, or windstorm, hail and flooding during a severe storm event. These difficulties should be kept in mind when database figures are analysed statistically, and the results have to be treated with due regard for the characteristics of the underlying data. Comparing events at different locations and on different dates can only be done using normalised data. For most analyses, and in particular trend analyses, socio-economic changes such as inflation or growth in population and values must be considered. Problems encountered when analysing trends are discussed using the example of floods and flood losses.

1 Global databases of natural disaster losses

Natural disasters are unique events. Many factors play a role in how and to what extent the various consequences of a natural event are generated. They interact and produce a complex process. In order to include a disaster in a database, and render different occurrences comparable, its details have to be condensed into a set of descriptive terms and figures, which could be called meta-data. These comprise fatalities, numbers of people injured, homeless, and affected, material damage to buildings and infrastructure, monetary losses, and key concepts describing certain features. The sum total of

these figures constitutes what we call natural catastrophe loss data, or in short, nat cat loss data.

Loss databases relating to current and historical natural catastrophes have become a valuable instrument, serving purposes ranging from risk assessment in the insurance business and socio-economic analyses to providing background for decision-making or simply, bringing natural disasters worldwide into the public eye. They are, therefore, utilised by numerous scientific institutes, researchers, national and international governmental and non-governmental organisations, the media, and of course the financial and insurance sectors.

In the case of the latter, observed losses were formerly an indispensable part of proper risk assessment. This still applies to those regions where mathematical risk models (loss estimation models) are not available. Indeed, where they are available, loss data are needed to calibrate and validate the models. In a general context, such as the United Nations' goal to substantially reduce natural catastrophes, as specified by the Hyogo Framework for Action (UNISDR, 2007), they are one means of measuring progress towards achieving the goal. International organisations such as the United Nations and the European Union, many national and provincial governments and even private companies in all sectors, and from local to global players, include these data in their strategic planning processes and disaster mitigation measures.

This vast usage of nat cat loss data imposes a heavy responsibility on database operators. It is vital to ensure that the underlying data are of the best possible quality and that this same quality standard is maintained throughout all the available datasets. Consequently, catastrophic events must be studied carefully and professionally by the institutions collecting the loss data, and then registered on the basis of defined criteria, irrespective of whether the events in question are past or current.

At present, there are three global, multi-peril database operators whose data are used and quoted regularly:

1. sigma (<http://www.swissre.com/sigma/>), a database of man-made and natural catastrophe losses that was set up by reinsurer Swiss Re in 1970 and has published statistical analyses in annual publications since then (Swiss Re, 2010);
2. Munich Re's NatCatSERVICE (<http://www.munichre.com/touch/naturalhazards>) was established in 1974 and based on an existing physical archive of past loss data. It covers catastrophes caused by natural hazards and has published relevant data in annual reviews since 1989 (e.g., Munich Re, 2011);
3. The Centre for Research on the Epidemiology of Disasters (CRED) at Leuven University (l'Université catholique de Louvain) in Belgium set up its Emergency Events Database (EM-Dat; <http://www.emdat.be/>) in 1988, initially with the support of the World Health Organization (WHO) and the Belgian Government (Vos et al., 2010). It is the most cited database, having been fully accessible to the public until recently. Analyses are published in an annual report and in occasional newsletters focusing on specific topics.

While all three databases contain the same overall information – such as economic losses, fatalities, numbers injured and affected, damage to infrastructure and buildings – the focus of EM-Dat is primarily on the humanitarian aspects, whereas the two reinsurers concentrate more on accurately reflecting the material losses (Wirtz et al., 2012). The three databases also apply different documentation thresholds (see Sect. 2.2). The following descriptions and discussions relate to the NatCatSERVICE database unless otherwise specified. As a rule, the statements made also apply to the other two databases.

2 Nat cat data management

2.1 Definitions and terminology

Clear and explicit standards, methodologies and definitions are essential factors in managing a natural catastrophe database, i.e., combining, merging and supplementing data, and comparing it with other databases. As a first priority, the peril terminology and definitions must be used in an unequivocal way. The three operators, together with UNDP (United Nations Development Programme) and ADRC (Asian Disaster Reduction Center), accordingly developed an internationally recognised standard applying to disaster-category classification and peril terminology (Below et al., 2009). This standard permits the comparison of data records and analyses produced by different organisations.

One definition not included in this agreement is the distinction between natural disaster and natural catastrophe, two

terms that are normally used synonymously. We suggest using them in a more precisely defined way as shown in the Appendix A.

In the following, the generic term natural “disaster” (rather than “catastrophe”) is used to designate events regardless of the extent of loss (c.f. Appendix A). We acknowledge, however, that the short form “nat cat” is a well-established component of a number of terms. Therefore, even though it would, for instance, be more appropriate to refer to “nat dis” rather than “nat cat” loss data, we use the latter in established concepts to comply with existing usage.

2.2 Identification and size of a natural disaster

Natural disasters that qualify for inclusion in the database are defined as losses that occur due to natural phenomena. These losses include loss of life, injury, resulting poor living conditions and material damage, but not indirect economic loss and adverse ecological impacts (as long as they do not entail costs).

The different focuses of the database operators result in different definitions of natural disaster, requiring specific entry criteria. The sigma database (overall loss of US\$ 86.6 m, insured loss of US\$ 43.3 m, both in 2010 values; 20 fatalities/people missing) and EM-Dat (more than ten fatalities and over 100 people affected) use quantitative thresholds as the minimum entry criteria. In EM-Dat, an entry is also made in the event, a state of emergency is declared or an appeal for international assistance made. The general criteria applied by NatCatSERVICE are lower, a loss dataset being created as soon as harm to humans (fatality, injury, homelessness) or property damage is involved; as a consequence, NatCatSERVICE documents more events than the other two databases. The events are classified into seven catastrophe categories (Fig. 1), depending on the extent and severity of impact: from a purely natural occurrence with no impact (Cat 0) to a great natural catastrophe (Cat 6) (Munich Re, 2006).

In addition, the insurance industry needs to clearly define the temporal and spatial extent of a loss event, because this can have major consequences for indemnification. This aspect is discussed in detail in Sect. 3.3.

2.3 Types of natural disasters

The operators of the three electronic global loss databases cover the entire range of natural hazards. They differentiate between six hazard families, consisting of various event types:

- Geophysical and geological events (earthquake; volcanic eruption; tsunamis; subsidence due to geological causes; “dry” landslide caused by earthquake, volcanic eruption, or geological processes; rockfall)

Catastrophe category		Overall losses					Fatalities
		Loss profile	1980s*	1990s*	2000s*	2010*	
0	Natural event	No property damage	-	-	-	-	none
1	Small-scale loss event	Small-scale property damage	-	-	-	-	1-9
2	Moderate loss event	Moderate property and structural damage	-	-	-	-	> 10
3	Severe catastrophe	Severe property, infrastructure and structural damage	>US\$ 25m	>US\$ 40m	>US\$ 50m	>US\$ 60m	> 20
4	Major catastrophe	Major property, infrastructure and structural damage	>US\$ 85m	>US\$ 160m	>US\$ 200m	>US\$ 250m	> 100
5	Devastating catastrophe	Devastating losses within the affected region	>US\$ 275m	>US\$ 400m	>US\$ 500m	>US\$ 650m	> 500
6	Great natural catastrophe „GREAT disaster“	Region's ability to help itself clearly overtaxed, in terregional/international assistance necessary, thousands of fatalities and/or hundreds of thousands homeless, substantial economic losses (UN definition). Insured losses reach exceptional orders of magnitude.					

Fig. 1. Catastrophe categories in NatCatSERVICE; at least one of the two criteria, overall losses and fatalities, must be met to qualify for a higher category. *Losses adjusted to the decade average, in 2010 values.

- Meteorological events (tropical cyclone; extra-tropical cyclone or winter storm; snowstorm; convective storm, including severe storm, hailstorm, tornado and lightning; local windstorm; sand or dust storm)
- Hydrological events (general flood; flash flood; storm surge; glacial-lake-outburst flood; avalanche; landslide caused by rainfall)
- Climatological events (heatwave; drought; wildfire; cold wave; frost; extreme winter conditions)
- Biological events (e.g., locust infestation; plague of vermin; epidemic plant diseases)
- Extra-terrestrial events (asteroid impact; solar storm)

Despite this essentially distinct classification, in practice, to which group or family a given event should be assigned is not always completely clear. A detailed discussion of the problems and the consequences thereof will be found in Sects. 3.1 and 3.2.

2.4 Event data

While the idea of making an entry in the database is to describe a disaster in as much detail as possible, the information and data actually available often fall somewhat short of the quality requirements. Nevertheless, disaster databases are now expected to provide data in greater quantities and more

granular form. A full NatCatSERVICE entry record has up to 200 attributes. The most important are the following:

- Event identification number and categorization: Hazard family, main-event type, sub-event-type (e.g., “Hurricane Katrina”); associated perils and consequences (e.g., famine following a drought, tsunami following an earthquake, etc.);
- Geographical information: Continent, subcontinent, region, country, state or province, town, coordinates (longitude/latitude);
- Date and duration;
- Victims: Death toll, missing, injured, homeless, evacuated, people affected;
- Type of damage and economic sectors affected: Number and type of buildings, vehicles, vessels, oil platforms, etc., damaged or destroyed; infrastructure, farmland, livestock, etc., affected; industrial losses; ecological damage (for information only);
- Monetary losses: Overall losses and insured losses;
- Scientific data: Moment magnitude (earthquake), category (tropical cyclone, tornado), wind speed (winter storm), intensity (precipitation), return period (flood peak), etc.;
- Other information: Socio-economic information; indices such as impact on GDP;
- Description of the event, including main details.

These informational items describe the main characteristics of a disaster event. Of course, each individual attribute is not encountered in every case. The crucial parameters for statistical analyses are: (1) fatalities, (2) overall losses, and (3) insured losses. Figure 2 shows a sample entry (Hurricane Ike).

2.5 Sources of information

Sources of information can be grouped into six categories: (1) insurance industry, (2) science and research, (3) UN, EU, administrations, governmental and non-governmental organisations, (4) meteorological, geological, etc., services, (5) news agencies, (6) other sources.

Basically, no source which may contribute to the database is excluded from the outset. For quality assurance purposes, every report is validated, evaluated, cross-checked with other sources and marked to indicate its credibility. Over the years, this procedure has led to the categorisation of frequently used sources: those with a high rating from the beginning, those that are basically trustworthy, but only to be used if verified, and finally those which may give valuable indications, but cannot be taken at face value without due verification.

In recent years, it has become much easier to identify and investigate natural disaster data – largely thanks to the internet. At the same time, it has become even more important to ensure that the sources are robust and sound. NatCatSERVICE uses around 200 sources identified as reliable for particular regions and/or types of event. Despite first-class sources, the analysis process can be fraught with problems. Typical challenges include erroneous reports, using incorrect currency-conversion factors, double counting of casualties and inconsistent use of terms. These problems will be discussed in Sect. 4.

Disaster loss reports are often duplicated and further disseminated. It is always possible that the information content will be changed due to abridgement, editing or simply human error. Database operators, therefore, have to test the quality of the figures they obtain. NatCatSERVICE's evaluation system assigns a quality level on a scale of 1 (very good) to 6 (inadequate) to every data record. Although data records of quality level 4, 5 or 6 are not up to the database standard and not used for analysis purposes, they are still incorporated in the stored data in order to retain the available information.

3 Complex events

3.1 Multi-peril events

In the aftermath of Hurricane Katrina, a large number of lawsuits were issued hinging on whether houses on the beach were destroyed by the hurricane's wind forces or by its storm surge. The background to these cases was the fact that, in the USA and a number of other countries, (flood) water damage and wind damage are covered under different insurance

contracts. An owner may be compensated if his house is destroyed by a storm, but not if his insurance company can prove that the damage was caused by water. Where both causes are involved, the situation is further complicated by such issues as: How does the damage break down? What happened first? Were there interactions between the different causes?

Most natural events manifest themselves in more than just one way, entailing primary, secondary and even tertiary perils. Tropical cyclones often bring not only high wind speeds, but also storm surges and torrential rain, which in turn may lead to landslides; a convective storm can be accompanied by gusts, hail, torrential rain (causing flash flooding), lightning and sometimes even tornadoes; an earthquake can trigger a tsunami or landslides, or may cause fires; heatwave and drought can result in subsidence due to soil shrinkage and sets the stage for wildfires. These examples show that the task of categorising natural hazard events is not a simple one. On the other hand, unambiguous categorisation is essential so that sectoral statistical analyses of specific hazards can be conducted and data entries compared. It is also crucial in order to resolve certain insurance-related issues, as illustrated above with Hurricane Katrina.

Disaster events are entered into the database according to triggering natural hazard (primary hazard) event or main cause of loss. Hence a tsunami triggered by an earthquake, for example, is listed in the database as “geophysical event/earthquake/tsunami”. This permits analyses on multiple thematic levels, be it looking at the number of geophysical events and the losses involved, or taking a more detailed look at earthquakes, or even specifically at tsunamis. In such cases, classification is clear and easy, whereas a flash flood can be created by a severe (convective) storm also featuring hail, wind gusts, etc., in which case it would be classified as “meteorological event/storm/severe storm/flash flood”, or it may be the sole damaging impact of a thunderstorm, which makes it a “hydrological event/flood/flash flood”. Usually, when analysing one specific peril class, the number of occurrences poses no problem. But even simple comparative, descriptive statistics may cause major problems, such as how to respond to the following request: “How many flash floods, how many hailstorms, how many destructive convective windstorms and how many tornado disasters happened in Italy in a given time period?” Assuming we had 30 floods, 20 hailstorms, 20 windstorms, 5 tornadoes and 25 multi-peril events made up of more than one of the single hazards, altogether this would make 100 disasters. If we added the damaging causes from the multi-peril events to the four specific classes, we would obtain the correct number of damaging flash floods, hailstorms, etc., but the sum would exceed 100. If we assigned the cause of the main losses to the multi-peril event, the total number of events would be maintained, but (some of) the specific events underrated.

The problems become even more acute when loss numbers are attributed to the various components of multiple-peril

Description | Main Selection | Selection Results

Event Header
 Type/MR No. Country: MR200809B044 Period: 12.-14.9.2008
 Event: Hurricane Ike, storm surge Subevent Cat.: tc: Tropical Cyclone
 Area: United States Area Details: AR; IL; IN; KY; LA; MO; NY; OH, Cincinnati; PA; TX, Galveston, Houston, Harris, Orange, San Leon,
 Losses USD m Eco/Ins: 30,000 18,500 Death Toll: 86

Affected people: Injured Damaged Infrastructure: Construction Sites national
 Missing Dam/Destr. 1,000,000 Roads/railways/bridges Agriculture International
 Evacuated Destroyed Water Supply Fishery/Aquaculture GDP (%):
 Displaced Health Centers Electricity/Communication Marine/Offshore **Number of Claims:**
 Affected Public Buildings Infrastructure damage **Overall state/Impact:**
 Diseases Boats Business Interruption State of Emergency
 Cars 100,000 Industry Environmental damage regional

Death Toll
 Deaths, number unknown

Death Toll	Date	Comment	Source	Assessment	Official Data
86	08/01/2009		Press reports		<input checked="" type="checkbox"/>
30	15/09/2008		PCS 60		<input type="checkbox"/>
56	28/09/2008		LL		<input type="checkbox"/>
47	16/09/2008		AP		<input type="checkbox"/>
67	30/09/2008		AP		<input type="checkbox"/>

Description
 Landfall Galveston 13.09.2009, cat 2 storm. Wind speeds of 180 km/h, heavy rain, major flooding. Remnants of Ike moved into the Lower Great Lakes region. Levees burst. >1 million buildings, mobile homes, vehicles damaged/destroyed. Roads blocked. Airport facilities damaged. 52 offshore oil and gas platforms destroyed, >60 damaged, oil spill. Losses to LNG lines (Liquified Natural Gas). Oil and natural gas production disrupted. Trees, communication lines, power lines downed, 2,200,000 residents without electricity. Evacuations. Also affected: Bahamas, Cuba, Haiti, Dominican Republic, Turks and Caicos Islands.

Exchange Rates (GKD rate type 'M'): USD-EUR 0.711921119 USD-CUR 1, as of 30/09/2008 00:00:00

Economic Damage (in millions)

USD	EUR	Avail. CUR	Select. CUR	Value	Source Date	Source	MR Assessment	Official Data
30,000.00	21,357.63		USD:United States Dollar	30,000.00	22/12/2008	MR		<input checked="" type="checkbox"/>
18,000.00	12,814.58		USD:United States Dollar	18,000.00	01/10/2008	WIR	including Bahamas	<input type="checkbox"/>
*								<input type="checkbox"/>

Insured Damage (in millions)

USD	EUR	Avail. CUR	Select. CUR	Value	Source Date	Source	MR Assessment	Official Data
18,500.00	13,170.54		USD:United States Dollar	18,500.00	12/10/2009	PCS+NFIP+MR		<input checked="" type="checkbox"/>
17,000.00	12,102.66		USD:United States Dollar	17,000.00	27/07/2009	PCS+NFIP+MR		<input type="checkbox"/>
15,000.00	10,678.82		USD:United States Dollar	15,000.00	04/03/2009	PCS+NFIP+MR	Next PCS est 3.6.09	<input type="checkbox"/>
12,500.00	8,899.01		USD:United States Dollar	12,500.00	12/10/2009	PCS 8+9	(+4.8.09)-no further surveys	<input type="checkbox"/>

List of Coordinates [Latitude -90 (South) to 90 (North); Longitude -180 (West) to 180 (East)]

Location Name	Latitude	Longitude	Quality	Avail. Damage Classification	Select. Damage Classification	Main Coordinate
USA: TX, Fulshear; District: Simonton	29.6801928311	-95.9692431974	city-centre or cip-code level coordinate			<input type="checkbox"/>
USA: TX, Galveston / Storm surge	29.18435	-94.970452	city-centre or cip-code level coordinate		Red	<input checked="" type="checkbox"/>
USA: LA, Cameron / Storm surge	29.795394	-93.316147	city-centre or cip-code level coordinate		Red	<input type="checkbox"/>
USA: TX, Houston	29.592429	-95.469238	city-centre or cip-code level coordinate			<input type="checkbox"/>
USA: TX, Houston	29.886478	-95.112682	city-centre or cip-code level coordinate		Red	<input type="checkbox"/>
USA: TX, Kemah / Storm surge	29.532845	-95.024527	city-centre or cip-code level coordinate			<input type="checkbox"/>

Fig. 2. Sample NatCatSERVICE entry (Hurricane Ike 2008).

events. It is rarely possible to split the overall loss into, for example, flood losses and windstorm losses. An example is provided in Sect. 7.2.

Furthermore, we have to consider loss occurrences caused by, for instance, flooding after the failure of a tailings dam. This is clearly not a natural disaster if the dam broke because it was poorly maintained. However, had it failed because it overflowed during an extreme rainfall event, it might be difficult to decide whether the failure was due mainly to the natural cause or to a design fault or poor maintenance (a liability rather than a natural-disaster issue). Such events tend to be classified erroneously as natural disasters (e.g., Stava, 1985,

China 1938)¹, showing that strict database management is needed to ensure the right items are included or excluded. If the database operators reach different conclusions on how such cases should be handled, this could, of course, result in discrepancies between the different databases.

¹Tesero, Stava Valley, Northern Italy, 19 July 1985: Tailings dam of a mine failed due to poor maintenance; 268 people were killed. No rainfall was involved. Yellow River, near Zhengzhou, 9 June 1938: Chinese general Chiang Kai-shek blew up dykes to halt the advancing Japanese army. Several hundred thousand people were killed by the water, illness and the subsequent famine.

Many wildfires are started by arson or technical causes (e.g., sparks from an engine), hence clearly by a non-natural hazard. Nevertheless, they are treated as natural disasters as the necessary preconditions for a wildfire to spread are weather-related and the actual trigger is not the crucial factor.

3.2 Storm beats flood

In many countries, at least in the western world, storm (including wind and hail), on the one hand, and flooding, landslide, earthquake, etc., on the other, are covered by different insurance policies. Almost everywhere, the insurance penetration for storm is much higher (typically 70–90 %) than for other perils (often less than 20 %).

Subsequent to multi-peril events, we are normally given only one overall figure for the combined losses. Since the NatCatSERVICE database, historically at least, was intended primarily for insurance purposes, and combined events tend to produce much higher insured losses than flooding, all losses from convective storms, tropical storms and winter storms are classified as storm unless the windstorm losses are virtually negligible compared with other loss causes, e.g., flood. There is, however, one exception. If the vast majority of losses from a convective event are attributable to water, that event is classified as flood. This exception does not apply to tropical storms given a name by the relevant institution (e.g., National Hurricane Center, Japanese Meteorological Agency); in these cases the losses are considered storm losses, even if produced exclusively by flooding or wave action. A prominent example is Tropical Storm Allison, which caused flood losses of US\$ 6.5 bn in the US in 2001, but practically no windstorm damage. These storm losses are included not in the (regular) flood statistics (as hydrological events), but in the storm statistics (as meteorological events). Another example is the low-pressure system Hilal, which struck Central Europe in 2008, causing flood, hail and windstorm losses of US\$ 1.7 bn. What share the individual loss components accounted for was the subject of guesswork.

This procedure may be regarded as arbitrary or even as a database design error, but there is no alternative. The proportions accounted for by storm, flood and other losses are seldom known in the case of complex events. Indeed, it is difficult enough to obtain (reasonably accurate) overall loss figures, and so we have to lay down a standard procedure.

Anyone unaware of these peculiarities inherent in the NatCatSERVICE who runs analyses using the data published solely for information purposes is likely to obtain biased or even erroneous results. It is crucial to know what the data represent before subjecting them to mathematical and statistical procedures.

3.3 Consecutive and multi-country events

In contracts between primary insurers and reinsurers the definition of “loss event” plays an important role. The reinsurer typically has to pay only if the primary insurer’s loss exceeds a specific amount (the priority or retention). Its obligation to pay is also subject to an upper limit (the limit of liability). Therefore, it is important to clearly define what constitutes a loss event in temporal and spatial terms. Distinguishing between different events is usually straightforward in the case of earthquakes and windstorms. The earthquake epicentre and time of occurrence are precisely known and losses occur instantaneously and can be directly related to the natural event. Even aftershocks are rarely the subject of dispute. Windstorms are produced by certain meteorological conditions and usually by a distinct pressure distribution in the atmosphere (or a “low”), which also allows the damage to be attributed in a particular way.

In the case of floods, attribution is often much more difficult, as their occurrence and intensity also depend on weather conditions prior to the event. Extreme rainfall may not necessarily trigger a basin-wide flood if it encounters a dried-up catchment, but lead only to sporadic losses. However, it may set the stage for a subsequent disaster, given even moderately intense precipitation, if the region is saturated and has no remaining retention capacity. Flood losses will then occur everywhere, including locations where the first event has already caused losses. From a hydrological point-of-view, the two events have to be seen in context. The issue the insurance industry has to address is: Do we have one loss event or two? In the case of two events, the primary insurer has to bear the retention twice. In the case of one event, the aggregate loss may exceed the limit. Depending on the losses actually incurred, either of the two may be disadvantageous for the primary insurer.

Two relatively recent examples of large floods illustrate this conflict. The August 2002 floods in Central Europe were produced by two consecutive lows, Hanne and Ilse, following each other on almost the same track. Reinsurers and insurers agreed that this constituted two events. The first produced overall losses of US\$ 5bn, the second US\$ 16.5 bn. The effect of this agreement is that the European flood event of August 2002 is represented by two entries in NatCatSERVICE and sigma, and by only one event in EM-Dat, which is not concerned with the insurance implications.

In 2007, Great Britain experienced a period of rainfall lasting several weeks during June and July. There was no pronounced pattern in the areas hit by rainfall during this period, flood losses occurring in some areas in the first few weeks, in neighbouring areas much later, and in some places more than once. Nevertheless, we were able to distinguish two major meteorological systems and accordingly declared there to have been two events. Since it was practically impossible to relate the insured losses (still less the uninsured, e.g., infrastructure, losses) to one or the other of the two, the

total insured amount of £3 bn (US\$ 6bn) was divided into two equal parts. From a hydrological point-of-view, we could also have chosen one event with a loss of £3 bn. The overall losses were derived from insurance penetration statistics for the UK.

The examples show that, even if all the loss figures are correct, we cannot produce a completely objective statistical analysis. There are always (semi-)subjective aspects involved. In terms of European flood losses, the two floods that occurred in the United Kingdom in 2007 would be ranked sixth and seventh, whereas combined into one event they would constitute the fourth most expensive European flood event. Consequently, we need to openly state how the data are processed, explain the background to the decisions we have taken, point out the problems involved and specify what assumptions we have made. This ensures statistical analyses are transparent and can be seen in their proper context.

Despite this policy, we are sometimes faced with publications in which our data (or rather our figures) or subsets thereof are used for mathematical and statistical operations without the authors' understanding what they represent. The results of these analyses need to be interpreted very carefully and, in some cases, questioned, because the datasets analysed do not reflect what the authors assume.

Financial or human losses sustained in several countries during one event are registered in the database as "region events". A region event is an all-embracing data record, containing information relating to all the countries affected. In addition, detailed information is available in the country records. This hierarchy permits analyses at the event as well as on a national level. Multi-country events are only counted once in the database.

The problem we face with multi-country disasters is that one natural event can produce different types of loss in different countries. A hurricane making landfall on the southern side of the Yucatán peninsula and producing storm surge and windstorm losses in Belize, may cause windstorm losses only in Mexico and flood losses only in Honduras and Guatemala. Based on the above standard, all the losses will be related to the hurricane, i.e., the meteorological event. This event would not be considered for the purpose of a statistical analysis of floods in Honduras by anyone not having access to – and taking into consideration – further knowledge. Only direct access to the database and a sophisticated analysis procedure will yield the correct result.

4 Accuracy and reliability

4.1 Reporting bias

We live in an era of communication. The internet, mobile phones and Twitter services enable us to receive news about anything and everything from even the remotest places on Earth. Satellites are able to show occurrences and measure

physical parameters day and night and at any spot on the globe. Nowadays, we can assume that we will find out about any natural disaster that happens, even local and small-scale ones, wherever they occur. This was not so until about two decades ago, and so we have to deal with a biased news flow to the database when we look back more than 20 yr. The further we go back in history, the greater the bias. However, this applies only in general terms and on a global level. We can assume in the case of Western Europe and North America, for instance, that the bias will not have been too great over the past 30–40 yr.

Apart from developments in communication technologies, political restrictions and boundary conditions have changed. Today access to internal information is denied or hindered by very few countries. Not long ago, many states in the Eastern Block, East Asia, South America and Africa felt they should not share information on the type of events that occurred and the losses they caused with the rest of the world.

One kind of reporting bias we still face is deliberately introduced by certain interest groups within the region affected, and in particular governmental (national, provincial, local) entities. The idea may be to amplify the losses (in order to obtain more international aid) or to understate them (in order to conceal deficiencies in disaster-preparedness, mismanagement or corruption). One example: In 1975, Banqiao dam on the Ru River in the Henan province of China broke and triggered the failure of several dozen other dams downstream. Tens of thousands of people died, but it was not until September 2005, when the files were opened, that the event was made available to the general public. Fortunately, we have developed techniques that enable us to reveal many such inconsistencies. One is simply long-term experience and another (the most important) is to cross-check against reports from other independent sources. There are normally reports from government sources and aid organisations and often even our own researchers in the field give independent loss estimates which can be compared with each other, ultimately giving a reasonably clear picture, at least, in the case of significant disasters. In recent years, mathematical risk models used in the insurance industry to assess potential large losses have acquired increasing importance for the assessment of actual losses. If suitably adjusted to reflect the specific features of the event, they can reproduce the actual total damage sometimes surprisingly well.

The accuracy of reported loss data depends on the country where the disaster occurs. Natural perils for which there is a high level of insurance penetration in the country concerned yield extremely reliable loss figures, e.g., storm losses in fully developed insurance markets (such as North America, Western Europe, Australia or Japan). If most of the losses are insured, uncertainty can only arise in respect of a relatively small residual amount (e.g., infrastructure damage). Countries with a highly developed governmental cost-surveillance administration thoroughly investigate what happened in order to draw conclusions concerning future

protection measures, legislation and preparedness. Societies differ as to how much their actions are based on a formal review of the situation and a decision taken in the light of the respective costs and benefits, culminating in its implementation at the end of the relevant decision-making process. In some cases, the solution may be ordained by a political leader, e.g., Chairman Mao's pronouncement following the devastating flood in China's Hai River basin in the vicinity of Tianjin in 1963. On 17 November 1963, at the "Hebei fighting against flood exhibition" he inscribed the slogan: "We must cure the Hai River from the root" (PictureChina, 2010). The solution was a huge control system that prevented subsequent catastrophes, but at very high monetary cost.

The smaller the consequences of a natural extreme event (i.e., only a small area is affected), the lesser its impact on politics and business, the smaller the effort to produce accurate loss figures, and the greater the need to rely on a handful of sources or even a single source. The resulting reports can only be checked by comparing them with similar events in the same region. Although reporting bias over time may be encountered in the case of small disasters, it can largely be eliminated where great catastrophes are concerned. A great natural catastrophe (GNC) leaves traces in a society's records which allow us to establish what happened, even decades after the event. Moreover, states, communities and the insurance industry have a special interest in knowing the extent of losses from great catastrophes. We at Munich Re have identified GNCs worldwide since 1950 and investigated their respective impacts very carefully. Hence, we can assume that the reporting bias in NatCatSERVICES's catastrophe category 6 (GNC, c.f. Fig. 1) is small and its time series consistent.

4.2 Reporting errors

Reporting errors (hopefully) occur unintentionally. Errors may arise from a number of sources, the most prominent being faulty translations and the conversion of currencies and units. For instance, 1000 square miles may become 1000 km², or RMB 1bn US\$1 bn (almost ten times as much). Simple calculation errors such as multiplying a number by the currency conversion rate instead of dividing it may go unnoticed, especially if the conversion rate is close to 1.0 (as in the case of € vs. US\$). Zeros may be deleted or added. We have to check whether the American "billion" unit (1000 000 000 or 10⁹) has been correctly translated into German, Italian, French "milliard(e)", "miliardo". In these languages "billion(e)" refers to 1000 000 000 000 or 10¹². Even American and British English differ in this respect.

Usage of descriptive words may also lead to wrong conclusions. The term "victims" is used by some to mean "deaths" and by others "those affected". Similarly, "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". Again, "200 000 ha of farmland affected"

can mean "200 000 ha under water" or that some unspecified tract of "farmland in an area covering 200 000 hectares was flooded". An example: the affected area in reports on the 1998 Yangtze flood ranged between 25 million hectares in press reports, 21 million according to Chinese Vice-Premier Wen Jiabao and the World Food Programme, 7.4 million as reported by the UN Disaster Management Team, 2.8 million in the final report of the UN Disaster Assessment and Co-ordination Team (UNDAC) and 497 760, the figure officially released by the Chinese Ministry of Agriculture (Sauer, 1999). While the latter may refer only to agricultural land (although even this is not certain), the other numbers still range within a factor of nine. During the Queensland floods in Australia at the turn of 2010, almost every news report quoted, "floodwaters (...) cover an area the size of France and Germany combined" (e.g., USA TODAY, 2011), i.e., about 904 000 km², which is almost exactly half the size of Queensland (1 852 000 km²). Again "flooded" was confused with "affected by floods". The original quote by Queensland's state Premier Anna Bligh had been: "We now have 22 towns or cities that are either substantially flooded or isolated because the roads have been cut off to them. That represents some 200 000 people spanning an area that's bigger than the size of France and Germany combined" (ABC, 2010). The first step in preventing such errors is to check the plausibility of the figures by relating them to the spatial extension of the region/country for which the event is reported, its overall wealth and former events. This procedure usually identifies gross errors. A cascade of other checks is subsequently applied, but it must be admitted that the final result is almost never devoid of uncertainty.

5 Estimating losses

Financial loss is the most important parameter in the Nat-CatSERVICE database. It is subdivided into two categories: insured losses and overall losses. The figures for the insured losses are relatively reliable because they reflect claims actually paid by insurance companies. Assessing overall losses is more complex.

5.1 What are economic losses?

The term "economic loss" does not have a uniform definition. It is important to differentiate between "direct losses", "indirect losses" and "secondary/consequential losses". While there is some ambiguity as to what exactly is to be understood by these three classes of loss, Munich Re defines them in the following way (Munich Re, 2001): Direct losses are immediately visible and countable (loss of homes, household property, schools, vehicles, machinery, livestock, etc.). They are always calculated on the basis of replacement and repair costs. Problems arise when it comes to estimating the value of historical quarters and cultural heritage that have

been destroyed. Another aspect is that damaged structures (e.g., dykes) are upgraded – while being repaired or replaced – to a higher safety and, hence, value level. While actual loss quantities should refer to damaged items, the figures in fact usually refer to costs.

Indirect losses include, among others, higher transport costs due to infrastructure damage, loss of jobs and loss of rental income. Two types of insured indirect loss are business interruption (BI), e.g., where production is halted because the insured's plant is flooded, and contingent business interruption (CBI), where it 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) apply to the economic impact of a natural disaster, for instance in the form of reduced tax revenues, lower economic output, reduced GDP or a weaker currency. On the other hand, reconstruction efforts normally stimulate the region's economy following a catastrophe, and lead to gains that compensate part of the losses.

5.2 How overall losses are estimated

Amounts entered as overall losses in the NatCatSERVICE database include only direct losses; indirect and consequential losses are not taken into account. As of 1 January 2012, more than 20 000 disaster events had been registered for the period 1980–2011, and stored in about 25 000 datasets (multi-country events requiring one regional dataset plus one for each country affected). About a third of the entries for specified overall losses are based on official sources such as governments and statistical and financial authorities. A similar share is reported by EM-Dat. If no official information is available, overall losses are estimated on the basis of insurance claims and/or other available loss indicators.

Estimating total losses on the basis of insured losses is usually reasonably efficient. Firstly, the latter are much better known because national insurance associations collate the respective inputs of their members. Since the national, regional or local insurance penetration (i.e., the percentage of insureds) for the different perils and countries is normally known, the overall losses can be extrapolated from the insured losses. The greater the insurance penetration, the more accurate the extrapolation results – at least in the case of windstorm. Where floods and, to an extent, earthquakes are concerned, and a far higher percentage of – usually uninsured – infrastructure damage is involved, the extrapolation results are less accurate. Unlike overall losses, indirect BI and CBI losses are included in the insured loss figures and may, as for example following Hurricane Katrina, assume substantial proportions. At that time, cable TV stations and credit card companies received tens of millions of dollars in CBI compensation for lost business (according to Munich Re's internal records). CBI cover is not particularly widespread, however.

If no known insured loss figures are available, the overall loss is estimated on the basis of other parameters. These include the type of natural disaster and its duration, the region affected (urban or rural), population density, level of prosperity, properties damaged, infrastructure, utility and other supplies, number of injured, homeless and fatalities. All available data are plotted in a matrix and weighed. The events are then assigned to a catastrophe category. 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 obtained for individual units (e.g., average value of residential buildings) (for details see Wirtz et al., 2012).

5.3 Loss history

Media reports very often publish loss estimates immediately after a catastrophe. However, such early estimates are not particularly reliable, losses sometimes being overestimated at the outset, in the hope of generating additional emergency aid. More often, the loss figures increase over time as the real extent of the catastrophe evolves.

The loss history reflects changes in the estimated overall and insured losses of a disaster event over time. It can be segmented into three phases:

- The first phase is the evolution of an ongoing crisis, as a natural event progressively affects more and more areas or increases in severity to the point where the hazard returns to normal again. This phase may last a few days in the case of tropical and extra-tropical cyclones, as the wind field moves on, or from several days to several weeks in the case of flood waves in large river basins, and from several weeks to several months in the case of volcanic crises, heatwaves, cold spells, wildfires and droughts. The first phase of short, instantaneous events such as earthquakes, tornadoes, storm surge or flash floods is actually a specific point in time. In such cases, the subsequent (two to seven) days, during which most of the search and rescue activities are completed, should be considered the first phase.
- The second phase is the aftermath of the catastrophe, when loss surveys are conducted. Ideally, such reports are initiated by the governments in the affected regions. However, it may also be necessary to perform the arduous task of putting together pieces of information from different regions like a puzzle. In the case of insured losses, the national insurance associations and supervisory bodies concerned usually provide accumulated loss figures, as do specialised modelling companies and reinsurers, these being based on simulations and calculations performed, among others, with

the aid of their nat cat loss models, if they are available for the region or country affected.

- The third phase is of indeterminate duration. Loss figure updates sometimes become available many years after an event as a result of long-term scientific studies, final reconstruction-cost figures, or the outcome of lawsuits.

The loss history is recorded as shown in the lower lines of the screenshot reproduced in Fig. 2. Figure 3 shows the loss history for insured losses (from various sources) based on the example of Hurricane Andrew (1992). Andrew made landfall on 23 August and swept across Florida. In the first five days following landfall (phase 1), losses were initially estimated to be in the order of US\$ 3 bn–5 bn. As loss adjusters began to report on the situation in Dade County, north of Miami, the focal point of the losses, the figures were initially revised downwards in the light of their observations and then gradually revised upwards (phase 2). They continued to rise in the months that followed – a feature typical of large catastrophes. Usually the final insured loss figure is known about one year after the event. Sometimes, e.g., if lawsuits concerning the payouts are pending, as with the 1994 Northridge earthquake and Hurricane Katrina in 2005, the final figures may be available only several years after the event, a situation encountered more frequently in the case of overall losses.

Following large natural catastrophes, governments, UN agencies and organisations of the UN system, the EU, universities, the World Bank, other development banks and large NGOs (e.g., the International Federation of Red Cross Societies) conduct or initiate detailed assessment reports. In most other cases, the damage is repaired and the event regarded as accidental, no far-reaching and long-term conclusions being drawn, and little effort being devoted to obtaining accurate loss estimates. Since, in the case of smaller events, the figures published in the earlier phases are often the only ones produced, they must be accepted as correct.

6 Comparability and normalisation

A global database must be structured in a way that allows events from different time zones and regions, with different currencies and social conditions, to be combined. US dollars are the main currency used in the database, losses being converted from the local currency into US dollars at the exchange rate which applied when the event occurred.

For many years, database operators have frequently published charts showing the temporal distribution of natural disaster losses. Figure 4 shows the overall losses from weather disasters in Europe in the period 1980–2010 in original values (blue bars). As with GNCs (cf. Sect. 4.1) it is reasonable to assume little reporting bias for this period in Eu-

rope, but clearly comparing today's absolute loss figures with those of the 1980s is futile due to the many changes in social, economic, environmental and, very likely, climatic conditions. Inflation has occurred, population figures have changed, prosperity levels and the exposed values – and their susceptibility to wind and water – have increased, land use has intensified, protective and adaptation measures have been set up (with the adverse effect of reducing risk awareness).

These effects have to be eliminated by normalisation, i.e., adjusting past losses to today's values. The minimum and doubtless easiest adjustment parameter is inflation. The result produced by deflating the losses on the basis of annual US inflation rates (consumer price indices, CPI) is also shown in Fig. 4 (red portion). To establish whether climate change is manifested in this graph, we need to determine if and to what extent the upward trend in the deflated bars is due to the net effect of socio-economic changes, the residue being attributable to a change in the hazard, and potentially to climate change. An intensive research effort is currently under way into developing methods to account for the additional normalisation aspects mentioned and identify any climate change signals in the losses (Neumayer and Barthel, 2011).

7 Flood disasters

When analysing flood losses, the difficulties encountered are greater than those relating to other natural disaster data. Floods – as opposed to wind, earthquake, or volcanic eruption events – are inherently a secondary type of natural event. Their primary causes are rainfall, temperature change (snowmelt), wind (storm surge) and earthquake (tsunami).

7.1 Types of flood

The initial challenge is the fact that there are different types of flood disaster. We distinguish between four main types: river/general flood, flash flood, storm surge and tsunami. There are a number of other types, but they are less significant and can be incorporated into one of the above main categories (e.g., debris flows and glacial-lake-outburst floods bear a resemblance to flash floods, whilst lake floods are similar to river floods).

River floods usually result from intense and/or persistent rain lasting for several days or even weeks and affecting large areas. Inundation emanates from the river channel, and flood control systems, in particular dykes and reservoirs, have a significant impact on the scale of the consequences.

Flash floods (including off-plain floods) usually occur in the form of independent, localised and random events. Unlike river flooding, it is not the total amount, but the intensity of rainfall that counts. On sloping terrain, this can produce a rapidly growing flood wave, on flat terrain, the water accumulates in lower lying areas such as depressions in the terrain, cellars and underground car parks. Flash floods may last anything from a few hours to one day. They produce

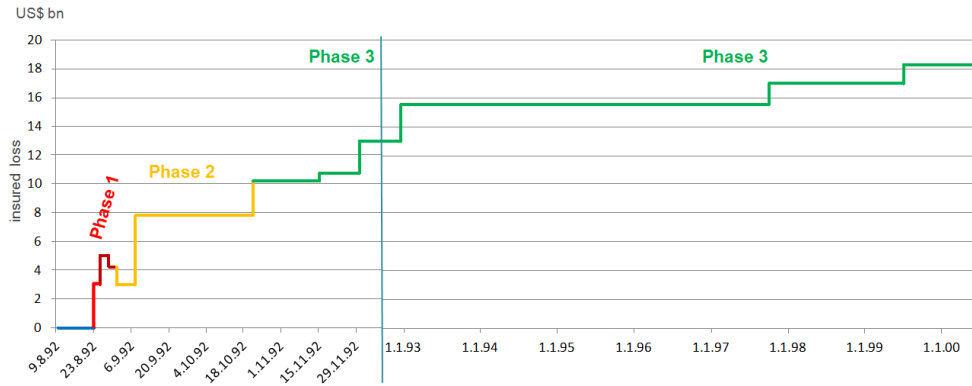


Fig. 3. History of insured losses (US\$ bn) from Hurricane Andrew (23–27 August 1992). (Note: The time scales on the left- and right-hand sides of the blue line differ.)

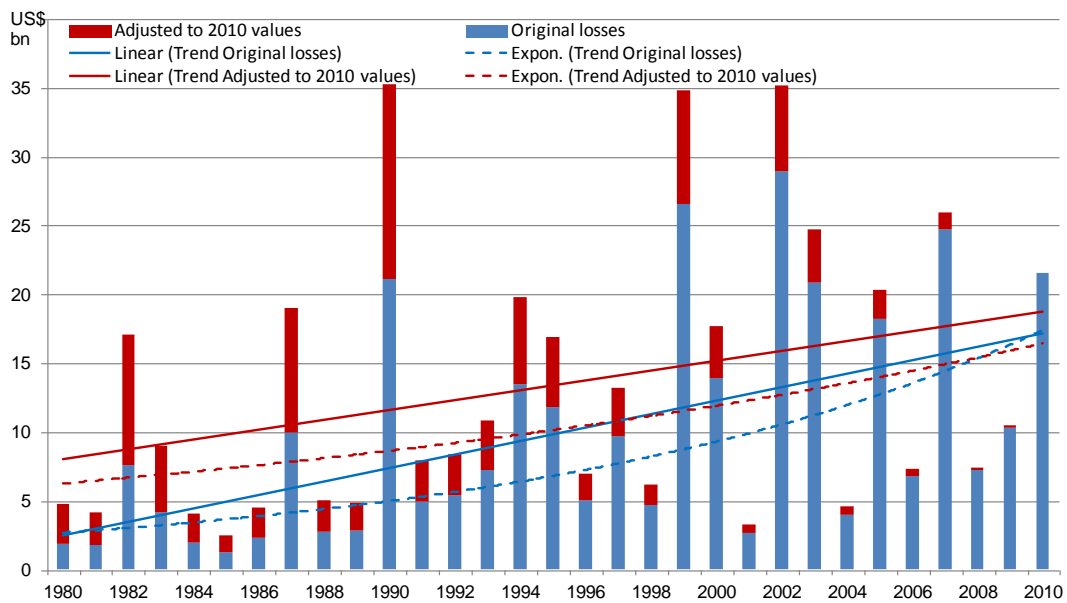


Fig. 4. Overall losses from weather disasters in Europe (1980–2010) in original and inflation-adjusted (2010) values.

high-intensity losses (i.e., high levels of damage per unit of area), but normally the affected area is comparatively limited.

Most floods occur in combination with other hazards. In the case of composite events resulting from tropical and convective storms, loss quantities can seldom be allocated to the various causes. We accordingly have to accept aggregate loss figures that include everything. It is extremely difficult or even impossible – and possibly very time-consuming – to accurately segregate the flood losses in the set of data. Local and regional flash floods, as a rule, fall in Europe within the category “severe storms”. Tsunami losses are never termed flood losses. They are always related to the triggering event, e.g., to an earthquake or a volcanic eruption.

7.2 Example analysis: flood and severe storm disasters in Germany

Germany experiences river floods and flash floods quite frequently. The number of natural disasters in Germany during which a flood loss occurred is shown in Fig. 5 for the period 1980–2010. It comprises the database categories “river floods”, “flash floods” and “severe storms”; of the latter (mostly convective storms with a flash flood component), only those which featured flooding are considered. River and flash flood occurrences vary between none and seven per year, with no discernible trend, while for all flood generating events the increase in the annual number is apparent.

A similar result is produced if we look at a database developed under the German research project URBAS on urban

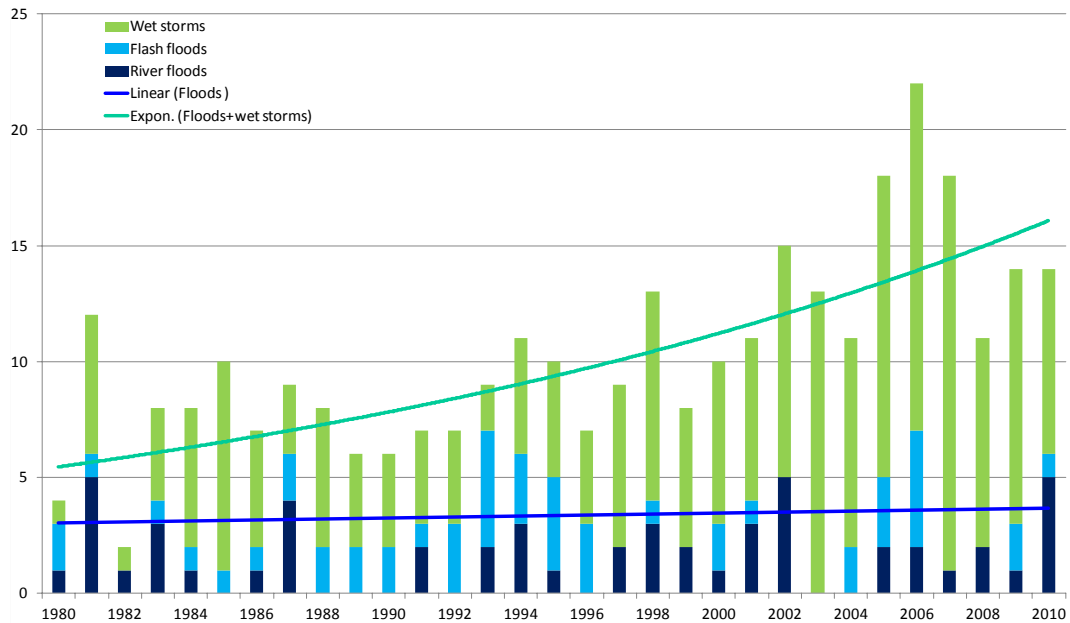


Fig. 5. Number of river floods, flash floods and wet convective events in Germany from 1980–2010 with trends.

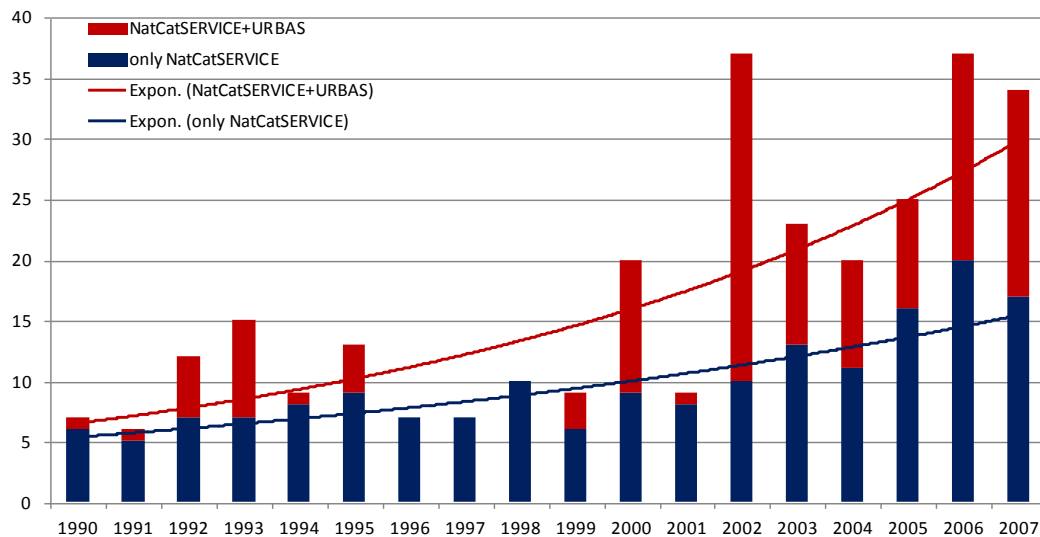


Fig. 6. Number of urban floods in Germany from 1990–2007 estimated from URBAS and NatCatSERVICE data with trends.

floods (Einfalt et al., 2006). Urban floods can be regarded as synonymous to flash floods in the way they are defined. The URBAS project team searched for any available reports on urban floods, such as in local archives and newspapers. This involved a lot of effort, but most of these reports did not include loss figures. The resulting database included original NatCatSERVICE data, but also additional events. Figure 6 shows the annual numbers of urban floods and the portion found in NatCatSERVICE (Note: River floods are not considered here).

The above statistics on the number of flood events can be considered reasonably reliable. However, it has to be ad-

mitted that it is almost impossible to produce valid amounts for the overall flood losses. Figure 7 shows the losses from floods and wet storms simply aggregated for each year. The resulting quantities are no doubt too high, as hail and wind damage are included. The truth lies somewhere between the top of the bars and the top of the flood portion. In a way, the whole wet storm portion of the bars can be regarded as the uncertain areas of the annual loss estimates. As a rough guess, it would certainly not be too unrealistic to assume that half of the wet storm losses are caused by water. To indicate this, the wet storm losses are split into two equal parts in Fig. 7.

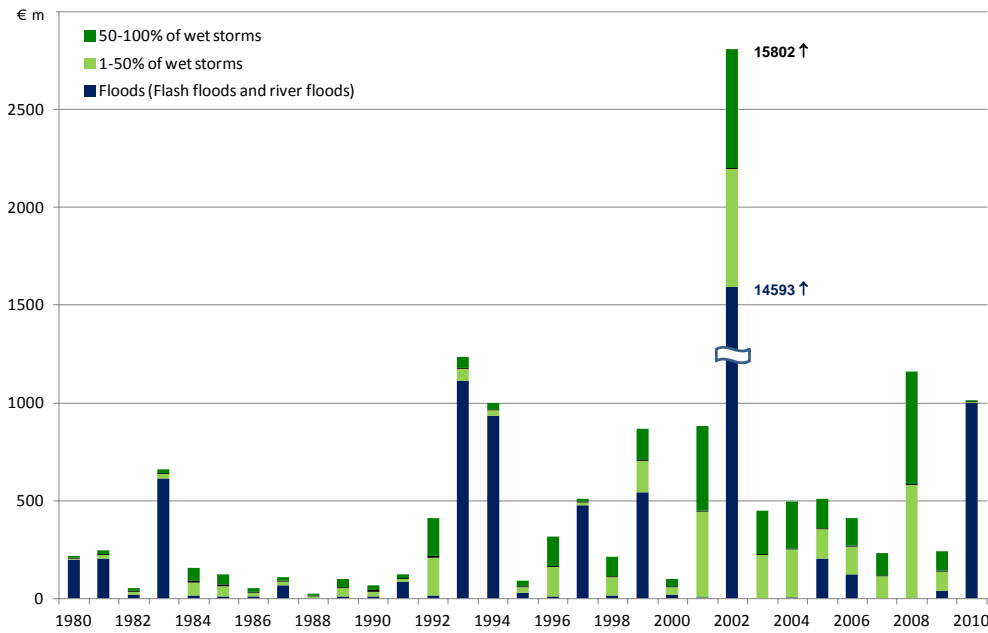


Fig. 7. Losses from floods and wet convective events (50 + 50 %) 1980–2010 in Germany (in 2010 values). *Note: The height of the bar for 2002 is not consistent with the axis.*

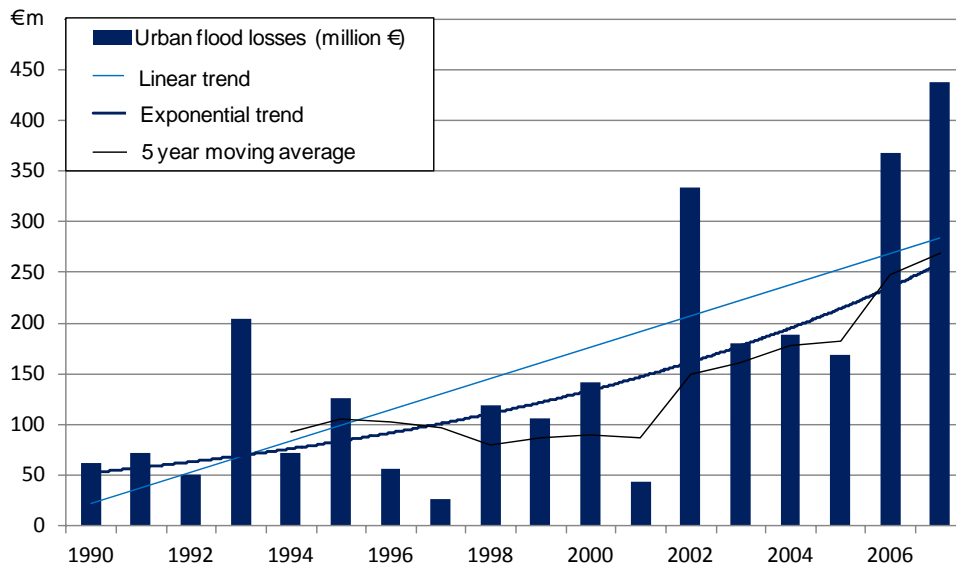


Fig. 8. Flash flood losses in Germany derived from URBAS and NatCatSERVICE databases 1990–2007 (in 2007 values) with different trend curves.

To investigate the typical portion of flood losses in convective events, we used the qualitative descriptions of the urban flood events from the URBAS database together with the description in NatCatSERVICE to compute loss values on the basis of parameters such as number of houses, basements, underground garages, public buildings, roads, railroads, etc., flooded, by assigning average loss values and taking into account the size of the affected regions. This analysis for

Germany (possible due to the limited availability of URBAS data only until 2007) yielded the result shown in Fig. 8. In a preliminary analysis, we now can relate these quantities to the 100 % losses from convective events (upper parts of Fig. 7 plus flash flood losses) and obtain an average percentage of flash flood losses of 34 %. This percentage, however, is obtained for one specific dataset and cannot necessarily be applied generally.

This analysis requires very detailed consideration of all available data, and differs strongly from simply taking the stored individual event losses from the database. It still needs more detailed investigation, but we believe it can give us a reasonable idea of the average distribution of flood and other losses in convective storms. It is certainly not too unreasonable an assumption to split it into two equal portions.

7.3 Trends

Visually, all the figures in Sect. 7.2 display distinct upward trends. But are they significant? The term “statistical significance” plays an important role, but it sometimes seems that its real meaning is not always fully understood. Figure 4 reveals quite large differences between the linear and the exponential trend lines/curves for the respective datasets. Which one is the most representative model? Or is neither of these models valid?

Significance is not an absolute quantity. It depends on the model chosen (e.g., linear trend, exponential trend, etc.), the way the model is applied (e.g., least squares, Mann-Kendall), the level of significance (90 %, 95 %, 98 %, etc.), the length of the time series and the variance of the data points. Significance includes two aspects. First comes the question: Is there a trend (i.e., is the slope different from zero)? Secondly, the significance (certainty) of the calculated value of the trend has to be determined. The most critical, but still sometimes disregarded aspect, is the quality of the database used for the trend analysis. How great is the uncertainty in respect of each single data point? Are the data consistent? Consistent in sampling method, but also consistent in having comparable individual accuracy? Dealing with gauge readings, for example, we usually know where a possible source of inconsistency can be expected (for example, in the year when the gauge was renewed). Some of the shortcomings may not be a problem if we have a large number of data points. But when we look at loss data, we are dealing with (a) (relatively) small samples that have (b) a large size range, in particular when it comes to monetary losses. Often the bulk of the samples consist of small to moderate quantities, with just a few data points that stand out from this bulk and might be regarded as outliers. At the very least, whether all the data should be treated as one combined sample is questionable.

Mathematical trend calculations and their significance are not concerned with what lies behind the data. They assume equal validity for each single value. Sometimes adding or removing a single data point and shortening or extending the length of the series by a single year changes a trend’s significance. And we need to be aware of the fact that one sample includes, for instance, data points from Bangladesh in 1970, China in 1980, Brazil in 1990, the US in 2000 and the UK in 2010. We would sound a general note of caution with regard to the purely mathematical treatment of disaster loss data. They are neither accurate nor of equal uncertainty and still less consistent in their nature. Analyses can and even must

be done, but their interpretation should not only consist of the result of a mathematical procedure. We would argue in favour of a stronger emphasis on plausibility and on the visual picture a time series displays. We do not always need to prove something, but rather to find out something. And the absence of mathematical evidence is not necessarily evidence of an absence of change.

Figure 8 shows three trend curves: a linear trend, an exponential trend and a five-year-moving-average curve. It is obvious that the losses rise with time; with the Mann-Kendall test, an upward trend with more than 99 % significance is obtained. But what is the pattern of the trend? How it is structured is not the subject of the discussion in this paper, but it would seem that an exponential model is more suited than a linear one. The data in the graph can be regarded as sufficiently consistent, but some reporting bias may be involved as there seems to be a step around 2002 (clearly seen in the moving average line) suggesting higher awareness and consequently greater reporting frequency following that year’s millennium floods.

8 Conclusions

Even with a relatively large quantity of data, they must be handled with due care and caution. Despite the fact that Munich Re gives such high priority to quality control, checking every single entry as thoroughly as possible, and correcting entries whenever new information is available, numbers should not be blindly crunched using statistical methods, and instead due attention is required, based on expert knowledge. Statistical analyses of natural disasters require a large set of data, but such sets must also be consistent and their components clearly understood.

A cautioning statement such as the following in EEA (2010) should itself be treated with caution: “*Available information from global disaster databases is limited and suffers from a number of weaknesses. One important consideration concerns increases in the reporting of events during the past few decades as a result of improvements in data collection and flows of information. ... Hence caution is needed in assessing any time series of flood disasters from global databases.*” While this statement is absolutely true, it by no means signifies that databases are useless. On the contrary, global databases contain information which is badly needed and found nowhere else. The key is to deal with the data in the appropriate way, which means always treating the data for analyses and publications with the necessary caution and prudence.

This implies:

1. Being aware that most individual data points represent estimates, not accurate facts.
2. Using every opportunity to confirm the validity of a figure through plausibility checks and cross-checks, and not disregarding other, deviating information, which should be taken as a sign of the need to check again.
3. Considering any possible cause of the distortion (changed environment, changed reporting, changed awareness, etc.).
4. Not using the data as pure numbers and avoiding reliance on mathematical analyses and statistical methods without knowing their “physical” characteristics and how they were produced.
5. Checking the sensitivity of the analysis result to slight changes in the dataset and to large changes in a single value. A trend may change considerably if the length of a data series is altered by adding or omitting just one year, and also if a single – large – value changes (e.g., because it is updated).

Measurements of physical parameters are practically always presented as point information. Such parameters are single components of the set of boundary conditions and input quantities. The situation in a given area is indicated by many points being aggregated on a grid in regionalised form. The loss of a natural disaster, however, is the outcome of the interaction of all input variables and boundary conditions, thus, the consequence of known and unknown interdependencies and interactions, non-stationary, non-uniform, nonlinear processes, random effects, and – sometimes highly erratic and unforeseeable – human behaviour, both of individuals and en masse. Even if we knew each and every physical parameter involved, we still could not calculate the exact outcome of a catastrophe. Therefore, the statistics of outcomes (losses) are indispensable to the assessment of future catastrophes.

Appendix A

Natural disaster vs. natural catastrophe

The two terms, disaster and catastrophe, are often used synonymously in the context of natural hazards. In fact, there is no real consensus among the various groups and people involved. In its “Terminology on Disaster Risk Reduction”

(UNISDR, 2009) the UN International Strategy on Disaster Reduction avoids the term “catastrophe” and defines only “disaster” as “a serious disruption of the functioning of a community or a society involving widespread human, material, economic or environmental losses and impacts, which exceeds the ability of the affected community or society to cope using its own resources.” Thywissen (2006) has compared terms used in the community and it seems that the majority understand “disaster” to be an occurrence that covers the whole range from large to small, while a “catastrophe” refers to large, severe events (Quarantelli, 2006).

A **natural disaster** involves economic or human loss due to a natural event. The literal translation of the Greek root “dis-aster” is “bad star”, taken from an astronomical theme, since the ancients used to refer to the destruction or deconstruction of a star as a disaster. Disasters normally imply sudden onset.

A **natural catastrophe** is an extremely large-scale disaster, a far-reaching event. The Greek “katá-stréphein” means “downturn”, in the sense of turning things upside down. Hence, a catastrophe refers to serious disruption of the functioning of a community or a society caused by widespread human, material, economic or environmental losses and impacts.

It is agreed that, unlike a natural event or phenomenon, both natural disasters and natural catastrophes, albeit the consequence of natural phenomena, have effects on humans and their belongings. This view takes into consideration the respective vulnerabilities, i.e., the capacity to resist injury or loss. Thus, disasters and catastrophes can be regarded as the consequence of inappropriately managed risk.

Acknowledgements. The authors thank the editor, Luca Ferraris, and Davide Miozzo and one anonymous reviewer for their very valuable suggestions and comments which helped to improve this paper.

Edited by: L. Ferraris

Reviewed by: D. Miozzo and another anonymous referee

References

- ABC: Bligh discusses flooding tragedy, Australian Broadcasting Corporation, Broadcast 30/12/2010 (<http://www.abc.net.au/7.30/content/2010/s3104006.htm>, last access: 13 January 2012, 2010.
- Below, R., Wirtz, A., and Guha-Sapir, D.: Disaster category classification and peril terminology for operational purposes, Working paper: Centre for Research on the Epidemiology of Disasters (CRED) and Munich Reinsurance Company (Munich Re), http://cred.be/sites/default/files/DisCatClass_264.pdf, 2009.
- EEA: Mapping the impacts of natural hazards and technological accidents in Europe, European Environment Agency, EEA Technical Report No. 13, Copenhagen, Denmark, 73, 2010.
- Einfalt, T., Friedeheim, K., and Wagner, A.: URBAS: Prediction and management of flash floods in urban areas, Proc. 4th European Conference on Radar in Meteorology and Hydrology

- (ERAD4), Sept. 18–22, 2006, Barcelona, Spain. ERAD Publ. Series Vol. 3, pp. 473–476. O8.6 – ERAD2006-A-00081, online: http://www.urbanesturzfluten.de/materialien/presentation/erad_abstract/view, downloaded on 19 February 2011, 2006.
- Frisch, M.: *Der Mensch erscheint im Holozän*, Suhrkamp, Frankfurt, 1981.
- Kron, W.: Flood Risk = Hazard · Values · Vulnerability, *Water International* Vol. 30, No. 1, March, 58–68, 2005.
- Kron, W.: Flood insurance – From clients to global financial markets, *Journal of Flood Risk Management*, 2, 69–75, 2009.
- Munich Re: Topics Geo – Annual Review: Natural catastrophes 2000, Munich Reinsurance Company, Munich, Germany, 2001.
- Munich Re: Topics Geo – Annual Review: Natural catastrophes 2005, Munich Reinsurance Company, Munich, Germany, 2006.
- Munich Re: Topics Geo – Natural catastrophes 2010: Analyses, assessments, positions, Munich Reinsurance Company, Munich, Germany, 2011.
- Neumayer, E. and Barthel, F.: Normalising economic loss from natural disasters: A global analysis, *Glob. Environ. Change* 21, 13–24, 2011.
- PictureChina: <http://picturechina.com.cn/bbs/watermark.php>, 2010.
- Quarantelli, E.: Catastrophes are different from disasters: Some implications for crisis planning and managing drawn from Katrina, Social Science Research Council, Katrina Research Hub, online article (<http://katrinaresearchhub.ssrc.org>; downloaded 19 February 2011), 2006.
- Sauer, H.-D.: *Das Jangtse-Hochwasser 1998: Ausmaße, Ursachen, Folgen*, *Geographische Rundschau* 51, Heft 6, 1999.
- Swiss Re: sigma – Natural catastrophes and man-made disasters in 2009, Swiss Reinsurance Company Ltd, Zurich, Switzerland, 2010.
- Thywissen, K.: Components of risk, a comparative glossary, Studies of the University: Research, Counsel, Education (SOURCE), publications series of UNU-EHS No. 2, Bonn, 2006.
- UNISDR: Hyogo Framework for Action 2005–2015: Building the resilience of nations and communities to disasters, <http://www.unisdr.org/we/inform/publications/1037>, 2007.
- UNISDR: UNISDR Terminology on Disaster Risk Reduction, <http://www.unisdr.org/we/inform/terminology>, 2009.
- USA TODAY: Australian city of Brisbane braces for flooding, http://www.usatoday.com/news/world/2010-12-30-australia-floods_N.htm, last access: 13 January 2012, 2010.
- Vos, F., Rodriguez, J., Below, R., and Guha-Sapir, D.: Annual disaster statistical review 2009, Centre for Research on the Epidemiology of Disasters (CRED), Université catholique de Louvain, Brussels, Belgium, 2010.
- Wirtz, A., Steuer, M., Löw P., Kron, W., and Höpfe, P.: The need for data – Natural disasters and the challenges of database management, submitted, *Nat. Hazards*, 2012.