Scientific consulting in reinsurance brokerage: models, experiences, developments[1]

By Dietmar Pfeifer, University of Oldenburg and
AON Re Jauch und Hübener, Hamburg

Introduction

One of the central problems in reinsurance during the last two decades is probably the increasing number and severity of natural catastrophes which has independently been reported in numerous scientific investigations. The larger part of these is mostly due to meteorological events like windstorms, hailstorms, flooding or landslide; areas where in particular growing damage rates are observable. Although there seems to be no stringent proof yet for the thesis that these rates are basically due to a worldwide climatic change, there is however much evidence that the observed trends cannot be explained by economic factors like inflation, concentration of values or a growing insurance density alone. For a more detailed discussion of these points, see e.g. Berz (1999). The reinsurance industry therefore has a vital interest in good physical or mathematical models which besides "static" information about model parameters also allow for a simulation of future loss scenarios by the use of suitable computer programs.
In this paper we want to show how scientific consulting in the reinsurance business can profit much from such methods, with a particular emphasis on the use of modern statistical tools. As an illustration, we consider insured losses caused by U.S. hurricanes which are sufficiently documented in public sources.

Geophysical and meteorological models

An important approach to the mathematical analysis of losses caused by climatic events is the modelling of the corresponding physical forces and their impact on the insurance industry. One of the first companies to develop such models was *Applied Insurance Research* (AIR) who have in particular concentrated on claims caused by hurricanes in the south-east of the U.S. For a survey, see e.g. *K.M. Clark:* Current and Potential Impact of Hurricane Variablility on the Insurance Industry, in: Diaz und Pulwarty (1997), 273 – 283; other aspects of such models, in particular w.r.t. earthquakes, are e.g. treated in Woo (1999). Besides a detailed study of relevant physival parameters such as air pressure, wind speed and direction, geographical locations of storm centers etc. the model also relies on a larga data base with informations on the location, type and content of insured buildings. With the aid of high-speed computers the model simulates storm events on the basis of weather records dating back until the early 1900's; a typical study comprises about 1000 simulations which are considered to be representative for future occurrances of such events. By means of suitable mathematical functions the simulated meteorological and physical parameters are then linked to the possible damages at or in the buildings under consideration. This results in the generation of loss potentials which are considered to be representative for today's and future claim scenarios, and allow for an empirical estimate for some PML (Probable Maximum Loss), which statistically corresponds to a (in general high) quantile q of the overall loss distribution. If F(x) denotes the corresponding cumulative distribution function, i.e. the probability for the event that the total losses do not exceed the value x, then formally $\mathrm{PML}(q) = F^{-1}(q)$ where F^{-1} denotes the mathematical inverse function. For practical purposes this quantile is usually also expressed in terms of the so-called return period T, which denotes the time interval within which on average *one* exceedance of the

[1] English translation of the original article *Wissenschaftliches Consulting im Rückversicherungsgeschäft: Modelle, Erfahrungen, Entwicklungen*. Zeitschrift für Versicherungswesen 21 (2000), 771 – 777.

PML is expected; i.e. we have $T = \frac{1}{1-q}$. Clark (1997) provides the following table for the overall loss potential due to hurricanes (insured claims, basis 1993) per *one* yearly hurricane event:

Return period T (years)	quantile q	PML (in Mio. U.S. $)
10	0,90	7800
20	0,95	13200
50	0,98	23600
100	0,99	30700
200	0,995	34500
500	0,998	50900
1000	0,999	51500

Tab. 1: loss potential by U.S. hurricanes

This table which is based on empirical values can also be read like this – from bottom to top: within 1000 simulated storm events (corresponding to 1000 years time horizon) there occurred precisely one claim of U.S.$ 51,5 Mio., precisely two claims of U.S.$ 50,9 Mio. and more, precisely 5 claims of U.S.$ 34,5 Mio. and more, etc.
From the statistical point of view, however, the empirical PML's particularly for large return periods (above 200 years) are critical, since they rely only on 5 simulated (observed) values. Also, the knowledge of only a few such PML estimations does not provide sufficient information about the underlying loss distribution as a whole, which however would be possible if all of the simulated values were taken into account.

Mathematical and statistical models

In contrast to the meteorological and geophysical models the statistical approach to the problem of forecasting potential future losses and PML's is to analyze past or *historic data*. There is some criticism by the physical modellers and in part also by the insurance industry in particular w.r.t. PML estimates for return periods of 200 years and above since no or only sparse loss observations are available here. In principle, however, this objection also applies to the physical models since they base on comparable historic storm events which are likewise extrapolated into the future. This problem is also discussed in Hipp (1999) and Pohlhausen (1999); Pohlhausen writes [conveying to the general sense]: "It is always a difficult task to derive forecasts for future developments out of the past. However, this is a meaningful enterprise. There is no other way to approach the uncertainty of the future."
Interestingly, due to the historical hurricane loss data set (1949 – 1992) published in *Catastrophe Reinsurance Newsletter* of 1993 ending with the 15 billion U.S.$ record loss caused by hurricane Andrew, it is in some sense possible to compare both approaches. Since the data are strongly affected by an exponential trend with a rate of about 10 % yearly average increase (as seen by some simple log-linear regression) the data have to be detrended and adjusted to the year 1993 before they can be compared to the AIR study. The following graph shows the result of such a procedure.

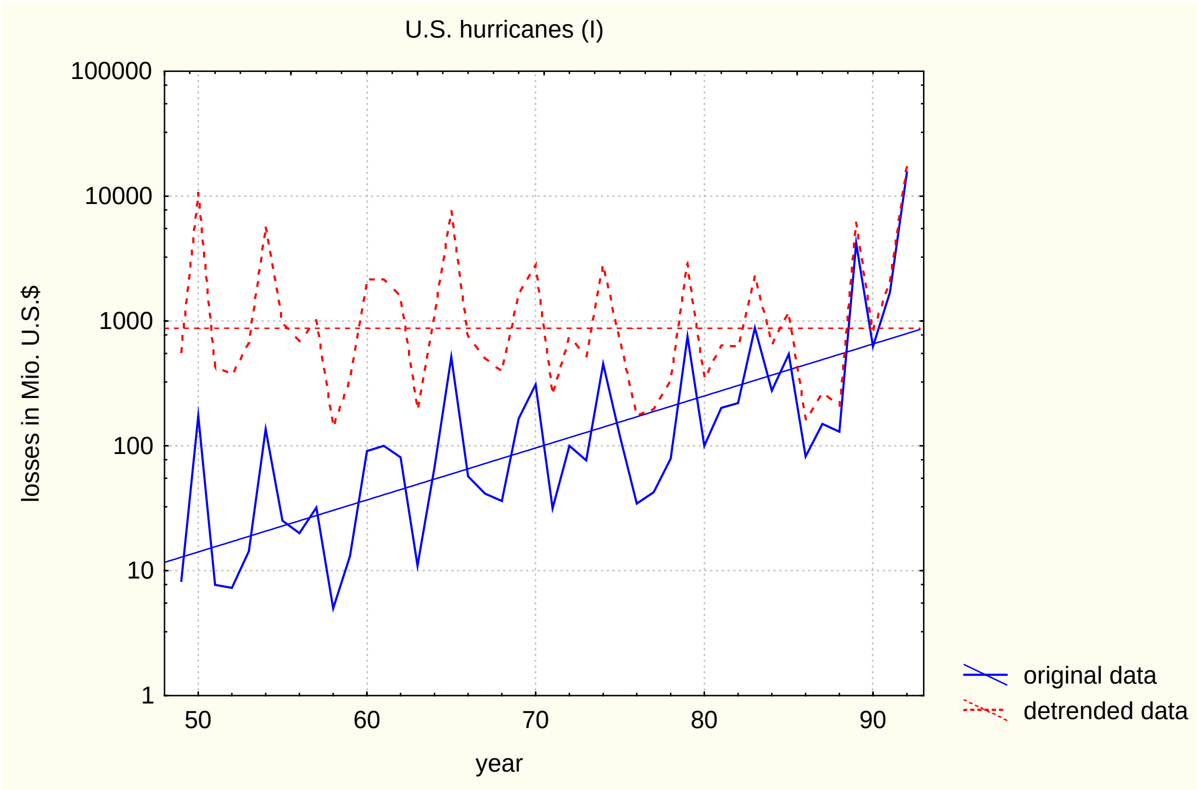

Fig. 1: observed and detrended hurricane losses during 1949 - 1992

Before a statistical analysis is to be performed with such data it is worth while to think a moment about possible candidates for fitting distributional models. Since most commercial statistical software packages offer a great variety of alternatives here, one should take some theoretical results into account which have been derived for large claims, e.g. in the framework of *statistics of extremes*, a field that has fruitfully grown during the last 100 years. See e.g. Hipp (1999), Beirlant et al. (1996), Embrechts et al. (1997) or Reiss and Thomas (1997), who discuss actuarial applications of this theory in greater detail. One central theorem of this theory says that under quite general conditions the normalized extremes (maxima) of independent observations possess a limit distribution function F which must belong to one of the following three classes:

distribution function	distribution class
$F(x)=e^{-e^{-x}}, x \in \mathbb{R}$	Gumbel distribution
$F(x)=e^{-x^{-\alpha}}, x > 0 \ (\alpha > 0)$	Fréchet distribution
$F(x)=e^{-(-x)^{\alpha}}, x < 0 \ (\alpha > 0)$	Weibull distribution

Tab. 2: extreme value distributions

Note that in the literature the Fréchet distribution is sometimes also denoted as *inverse Weibull distribution*, because the negatively inverse values of Weibull distributed random variables are Fréchet distributed. More recently a uniform parametrization of these classes has been developed where the Gumbel distribution is a limiting case between the Fréchet and the Weibull distribution; see e.g. Reiss and Thomas (1997), Chapter 1.3 (so-called γ-parametrization). For practical applications w.r.t. the fitting of loss distributions the Gumbel and the Weibull distributions are inappropriate, however, since they have mass in the negative real numbers. On the contrary, the remaining Fréchet distribution has turned out to be extremely efficient, in particular when fitting losses from windstorm events; cf. Pfeifer (1997) and Rootzén and Tajvidi (1997). The following graph shows the result of a simulation study for 44 claims from a Fréchet distribution fitted to the detrended hurricane data.

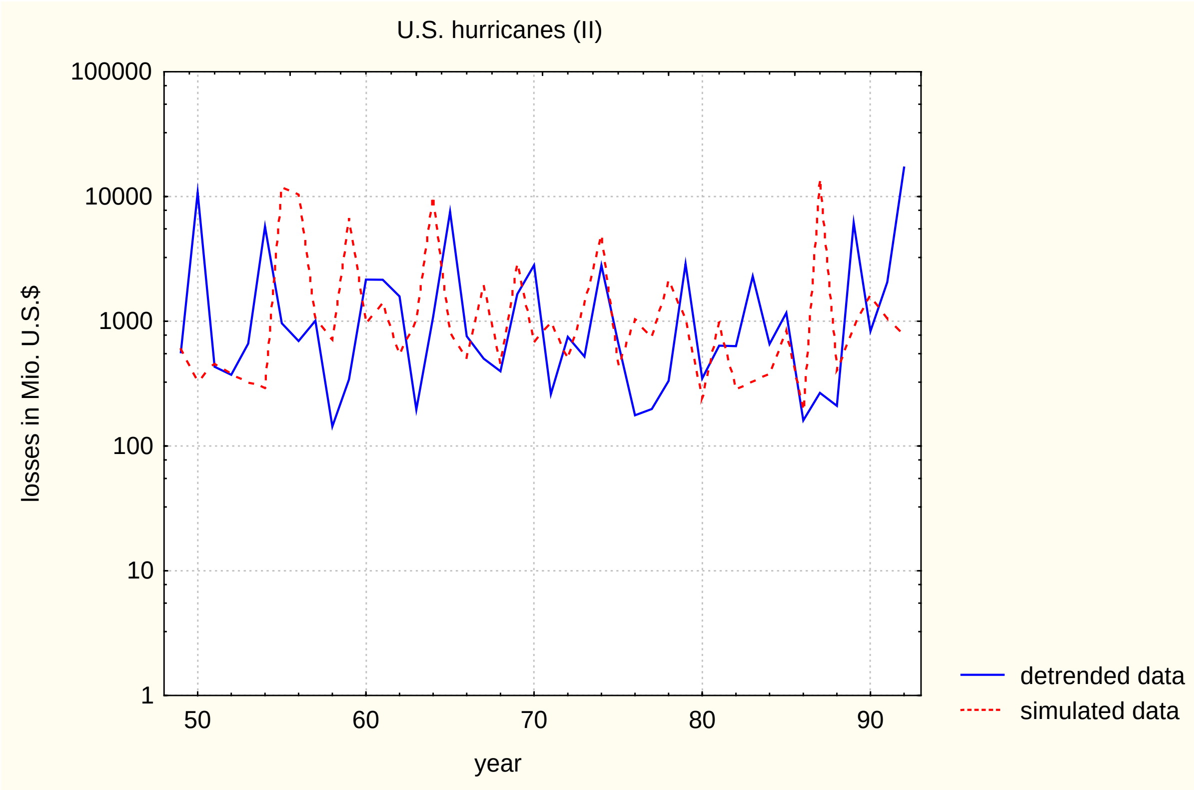

Fig. 2: simulated and observed hurricane losses (detrended)

Seemingly there are no systematic differences between both time series visible, which is also confirmed by appropriate statistical tests. However, it makes sense to include also other distributional classes into the analysis, in particular those which exhibit a tail behaviour similar to that of the Fréchet or other extreme value distributions. Such classes include for instance the Pearson type V (inverse Gamma) or the loglogistic distribution, which are available in some professional fitting packages (see Law and Kelton (1991)). The following table contains the estimated parameters (scale and shape parameter) for some of these distribution classes, for the detrended hurricane data set. The ordering of the models is according to the goodness-of –fit, i.e. the Fréchet distribution provides the best result here.

model	Fréchet distribution	Pearson Type V	Loglogistic	Lognormal
scale parameter	506,8325	566,37823	802,31944	6,77273
shape parameter	1,05681	1,09325	1,50267	1,17497

Tab. 3: estimated parameters for hurricane data (detrended)

The following graph shows the goodness-of-fit between empirical and theoretical distribution function.

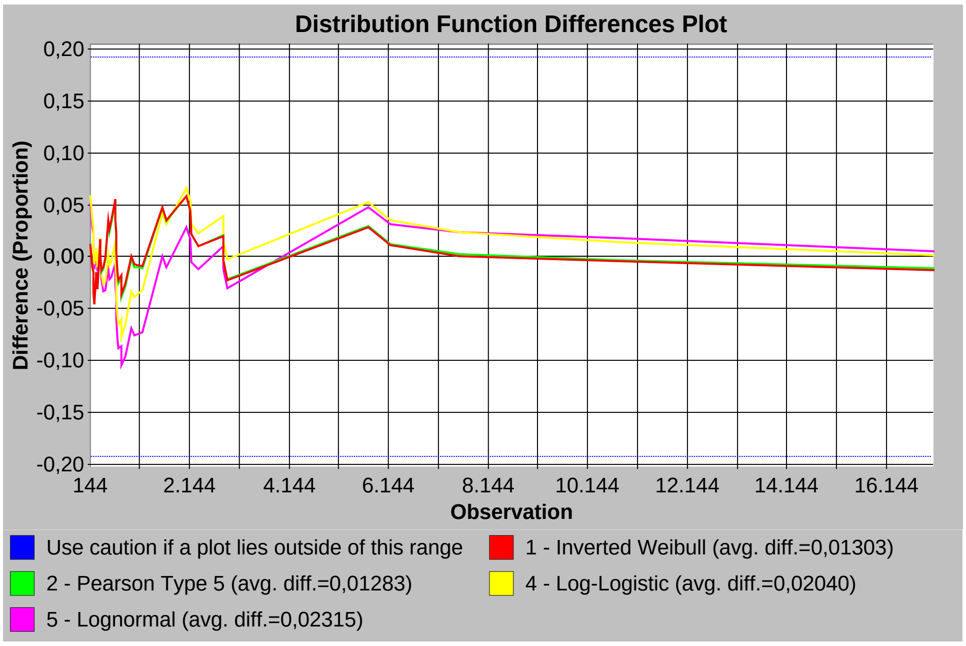

Fig. 3: plot for visualization of goodness-of-fit (detrended hurricane data)

The above graph shows quite well that potentially all four distribution classes could be used for modelling the hurricane losses since all deviations remain between the critical lines (dashed). However, it is also seen that the Fréchet distribution shows the least oscillations here, hence provides the best fit according to this criterion. Another possibility to test the goodness-of-fit graphically is to use the so-called P-P-plot in which theoretical cumulative probabilities are plotted against the empirical ones.

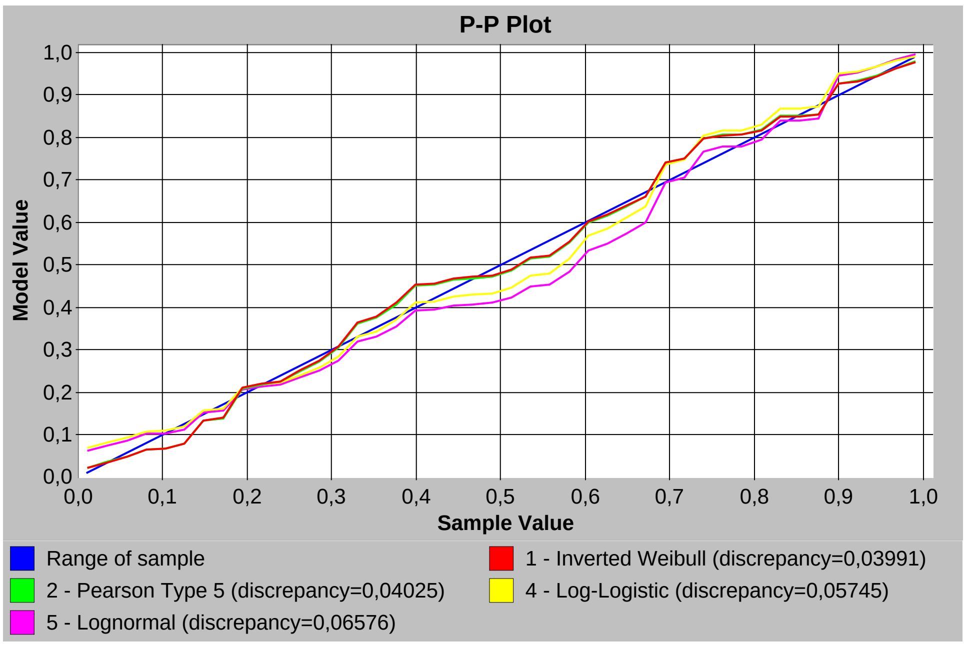

Fig. 4: P-P-plot for detrended hurricane data

According to this criterion, the Fréchet distribution is again best-ranked here. For a very exhaustive discussion of such methods we refer to Beirlant et al. (1997).

Once an appropriate model fitting has been done it is possible to obtain corresponding PML estimates from that. In the case of a Fréchet distribution model, it is even possible to express the PML in terms of the return period T explicitly:

$$\mathrm{PML}(\mathrm{T}) = \sigma \left\{ -\ln\left(1 - \frac{1}{\mathrm{T}}\right) \right\}^{-1/\alpha} \simeq \sigma \mathrm{T}^{1/\alpha},$$

where σ denotes the scale parameter, and α denotes the shape parameter .The above approximation is sufficiently precise for return periods above 20 years already. For the four distribution classes, we obtain the following results.

		distribution class		PML (Mio. U.S. $)		
return period T	quantile q	Fréchet	Pearson Type V	Loglogistic	Lognormal	AIR
10	0,90	4262	4201	3462	3938	7800
20	0,95	8422	8167	5692	6035	13200
50	0,98	20340	19244	10694	9757	23600
100	0,99	39381	36520	17076	13441	30700
200	0,995	76063	69088	27176	18020	34500
500	0,998	181276	160091	50105	25706	50900
1000	0,999	349459	302041	79525	32980	51500

Tab. 4: PML estimates after distributional fit, U.S. hurricanes

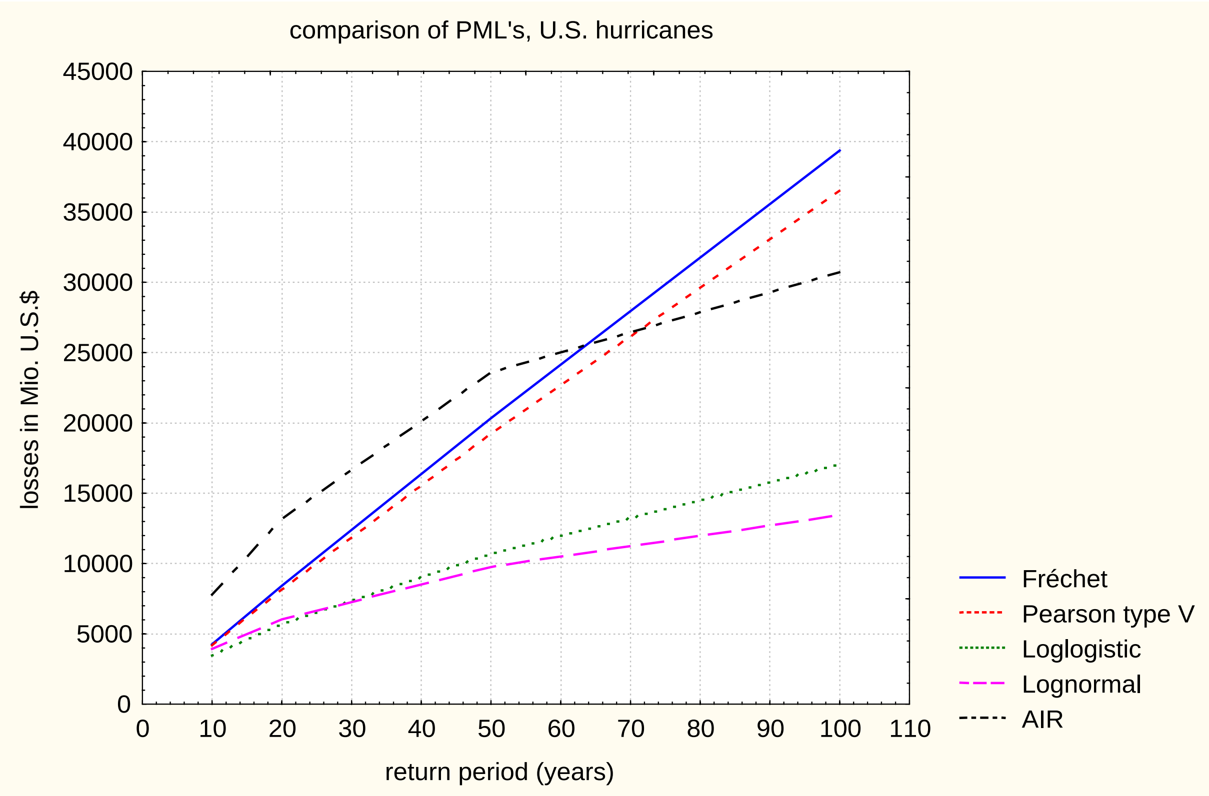

Fig. 5: PML estimates after distribution fitting, U.S. hurricanes (detrended)

Seemingly there is a qualitatively good coincidence between the PML estimates of AIR and those from the Fréchet or Pearson type V model in the range of up to 100 years for the return period T. For larger values of T, however, there are substantial differences in the estimates; one possible aspect is here that the PML estimates of AIR are for a single storm event per year only while the statistical analysis considers the aggregate claims over the whole year. This might explain for a PML estimate which is roughly twice as high in the Fréchet and Pearson type V model compared with the AIR value for a return period of 200 years since the average frequency of hurricanes is definitely more than one per year.

Experiences and recommendations

There is certainly no simple or even unique solution to the problem of estimating the "true" PML in practical situations. Depending on which reinsurance company or broker is concerned with this problem or how large the return period T under consideration is, there will most probably be a variety of different results, not only for the U.S. hurricane market; see also Pohlhausen (1999). However, experiences of more than five years of scientific consulting for a leading reinsurance broker have shown that in most cases PML estimates for catastrophic claims with return periods up to 200 years are still comparable, also under the different approaches outlined above. Perhaps the value of the PML alone is not even the most interesting information; the "dangerousness" of the loss distribution might count here as well. In the case of extreme value distributions and their relatives this aspect is usually described by the shape parameter α. For Fréchet distributions this relationship is immediately seen by the fact that the PML increases with the return period T like $T^{1/\alpha}$; i.e. the closer α is to the value 1, the more „linear" is the increase of the PML with T. It should be kept in mind that for values of α below 1 there does not even exist the mathematical expectation of the loss distribution, which means that theoretically the corresponding risk is not even "insurable". By experience, most storm portfolios worldwide show up with α-values beween 1 and 1,5 indicating that the loss distribution is quite "dangerous" with a tendency to cyclic large loss potentials. Statistical extreme value theory here also provides tools for the estimation of such parameters; see e.g. the monograph by Reiss and Thomas (1997) which comes with a CD ROM containing a statistical software package called **XTREMES** being developed at the University of Siegen, Germany.

Future developments

So far classical analyses of loss potentials with physical as well as statistical methods were mainly restricted to the one-dimensional case , with emphasis on PML estimates for individual risks with univariate loss distributions. In the future, however, multidimensional analyses will become of more importance, e.g. in connection with the rating of insurance bonds or, more generally, the so-called Alternative Risk Transfer (ART) which is concerned with the transfer of insurance risks to the capital market; see e.g. Hipp (1999). Although financial derivatives like CatXL or other catastrophe bonds do not yet play a major role in the reinsurance business one should be prepared to deal with the many possible dependence structures between the different risks involved. The approach via correlations which is quite popular in finance is, however, not appropriate here since due to the particular situation (extreme value distributions) statistical dependencies between risks cannot be modelled suffiently with such tools. A possible approch is here given by the so-called *copula models* which offer the possibility to describe the statistical dependence via backtransformation of the univariate observations by the inverse cumulative distribution function to (theoretically) uniformly distributed marginals; see e.g. Reiss and Thomas (1997), Chapter 8. Their software package **XTREMES** also contains a module for the analysis of different copula models which are particularly important in the range of extreme value distributions. Yet there is still demand for further scientific research in this area, especially w.r.t. the identification of suitable copula models for catastrophic risks.
Other multivariate statistical tools which have turned out to be very efficient in practical applications are classical ordination principles like *principal components analysis* (PCA) or the more modern *multidimensional scaling* (MDS), which belong to the area of *explorative data analysis*. A central

idea of MDS is to model the similarity of time series pertaining to claims caused by different risks and to map these for visualization purposes into the two or three-dimensional Euclidean space. This corresponds to a certain kind of complexity reduction; see e.g. Fahrmeir and Hamerle (1984) and enables the user to identify easily those risks which are similar in their loss potential either due to adjacent geographic locations or due to their dependence on similar climatic factors, or other effects. The following graph shows the result of a correponding analysis for risks like windstorm, hailstorm, flooding, landslide, avalanches and snow pressure, based on long-range times series of claims from Central Europe.

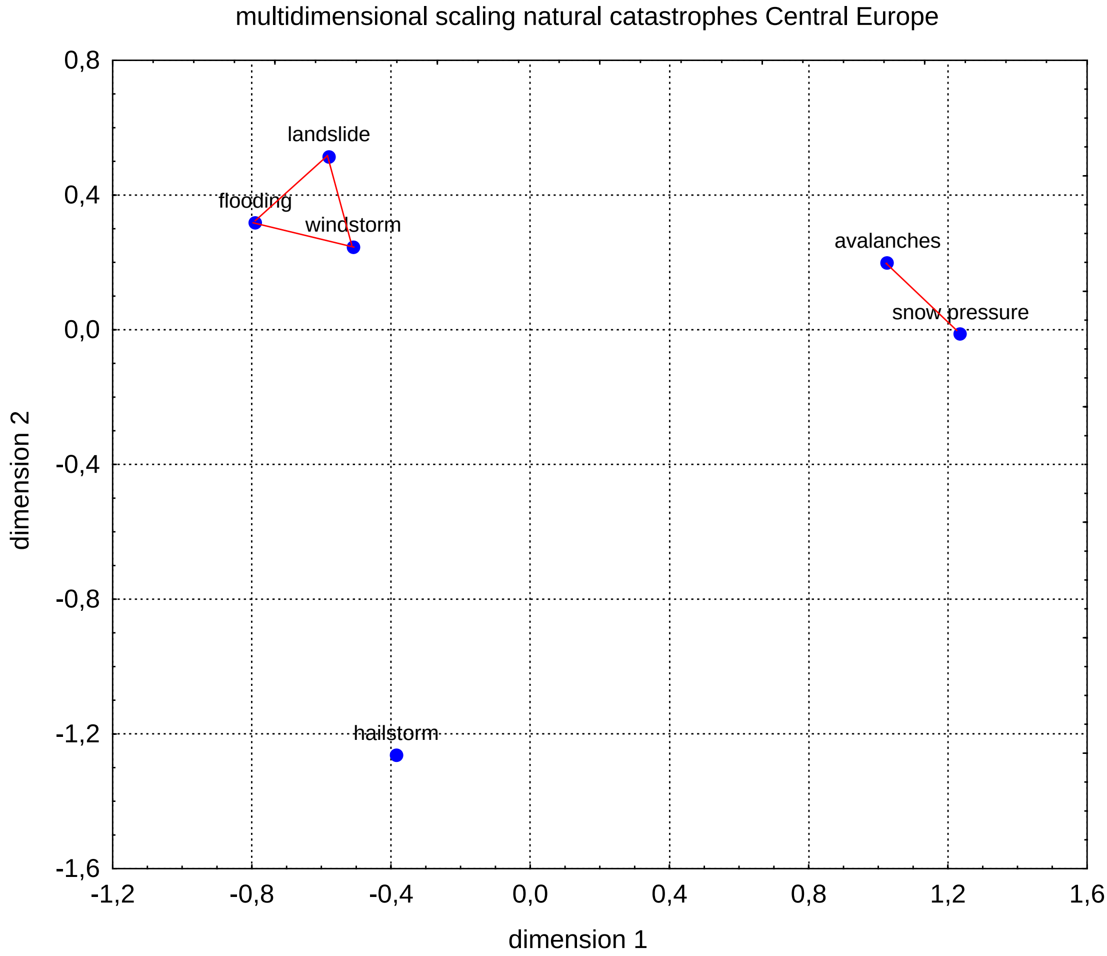

Fig. 6: example for a statistical analysis of claims with multidimensional scaling

Note that the "dimensions" 1 and 2 do not have any direct physical interpretation here, however the physical distances between "points" represent the true similarity of loss potentials for the corresponding catatrophe types. It is interesting to notice that three disjoint groups of risks are obtained here (windstorm / flooding / landslide vs. avalanches / snow pressure vs. hailstorm). This can perhaps be explained by similar climatic triggers for such kind of risks and is in very good coincidence with investigations of various (re-)insurers.

References

Beirlant, J., Teugels, J.L. and Vynckier, P.: Practical Analysis of Extreme Values. Leuven University Press, Leuven 1996.

Berz, G.: Naturkatastrophen an der Wende zum nächsten Jahrhundert – Trends, Schadenpotentiale und Handlungsoptionen der Versicherungswirtschaft. Zeitschrift für die gesamte Versicherungswissenschaft 2/3 (1999), 427 – 442.

Catastrophe Reinsurance Newsletter (1993) No. 2, p. 8.

Diaz, H.F. and Pulwarty, R.S.: Hurricanes. Climate and Socioeconomic Impacts. Springer, N.Y. 1997.

Embrechts, P., Klüppelberg, C. and Mikosch, Th.: Modelling Extremal Events for Insurance and Finance. Springer, Berlin 1997.

Fahrmeir, L. and Hamerle, A.: Multivariate statistische Verfahren. W. de Gruyter, Berlin 1984.

Hipp, Ch.: Risikomanagement von Naturkatastrophen: helfen mathematische Methoden? Zeitschrift für die gesamte Versicherungswissenschaft 2/3 (1999), 443 – 456.

Law, A.M. and Kelton, W.D.: Simulation Modelling & Analysis. McGraw-Hill, N.Y. 1991.

Pfeifer, D.: A statistical model to analyse natural catastrophe claims by means of record values. Proceedings of the XXVIIIth International ASTIN Colloquium, Cairns, August 10 - 12, 1997, The Institute of Actuaries of Australia, Sydney, 45 - 57.

Pohlhausen R.: Gedanken zur Überschwemmungsversicherung in Deutschland. Zeitschrift für die gesamte Versicherungswissenschaft 2/3 (1999), 457 – 467.

Reiss, R.-D. and Thomas, M.: Statistical Analysis of Extreme Values, with Applications to Insurance, Finance, Hydrology and Other Fields. Birkhäuser, Basel 1997.

Rootzén, H. and Tajvidi, N.: Extreme value statistics and windstorm losses: a case study. Scandinavian Actuarial Journal (1997), 70 – 94.

Woo, G.: The Mathematics of Natural Catastrophes. Imperial College Press, London 1999.