NFT 1/1997

On predicting the next very severe wind storm loss

by fil.kand. **Håkan Pramsten**, actuary at Länsförsäkringsbolagens AB[1]

Håkan Pramsten

The practical problem of predicting the loss amount that will hit an insurance company in the next very severe wind storm has recently been analysed in the context of extreme value statistics. The new assessment methods proposed will be shown to be easy to apply and also well justified from a theoretical point of view. It is also argued that recent years' windstorm and hurricane catastrophes could have been predicted by proposed methods.

1. Introduction

Losses caused by natural and man made catastrophes on human lives and welfare is a recurrent threat to society. Without appropriate measures, e.g. a reinsurance program matching exposure, such events also jeopardize the survival of any insurance company exposed. This paper addresses in particular the problem of predicting the accumulation of losses due to a wind storm event. Planning next year's reinsurance cover, one crucial figure, for any non-life insurance or reinsurance company with a wind storm exposure, is the answer to the question: "What accumulated loss on a gross basis will the company suffer, when hit by the most severe wind storm event next year?"

Although, due to genuine uncertainty, time, place and amount for such an event seems to be unpredictable, Rootzén et al [13] recently report that such outcomes can, de facto, be predicted and uncertainty can be measured.[2]

The ultimate aim of this paper is – in a less technical way and requiring less mathematical skills – to make the insurance community aware of this quite new risk theoretic approach, which is elaborating on an idea put forward by the German insurance mathematician, Erhard Kremer [10], [11], [12]. It will not only give a predicted monetary amount, but also a measure of rational belief whether

[1] Now Chief Actuary at WASA SAK Försäkringsaktiebolag.

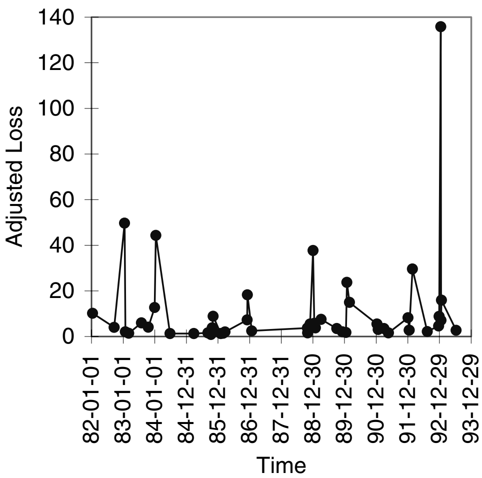
[2] This research has been supported by Länsförsäkringsbolagens forskningsfond.

or not, in a specified amount of time, a loss will occur in excess of the predicted amount or not and, if so, by how much. It will also be argued that the proposed methods, although not yet fully explored, are of practical importance and well justified from a theoretical point of view.

It may be that the audience gets alarmed by the potential risk in wind storm insurance, underwritten without any contractual aggregate limit as illustrated in the worked out examples below.

The paper is divided into two parts. The first part is practically oriented, including worked out examples of the proposed new methods, showing how simple they are to apply. The second part of the paper try to prove that they are well justified from a theoretical point of view. At end there is a list of questions still not resolved and needs of further research.

Figure 1. Länsförsäkringsbolagen's adjusted wind storm loss amounts in MSEK per episode 1982-93



2. The storm loss data used

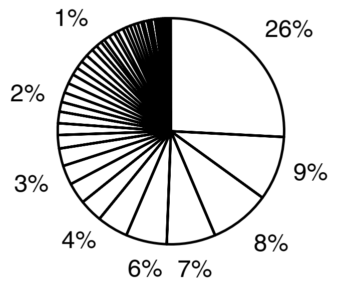
For illustrative purposes the proposed techniques will be used on data as in Figure 1. The data should be viewed as 46 wind storm losses, adjusted for inflation, exposure and climatological changes, hitting Länsförsäkringsbolagen (LF) in 1982 until 1993.

The question now is, what could be said about the most severe wind storm loss a predetermined number of years ahead, considering the information available in the Figure 1 time series.

3. The prediction problem in light of experience

The prediction problem is not an easy one. This is illustrated by Figure 2, where the events are ordered in magnitude according to share of period total loss.

Figure 2. Events in Figure 1 ordered by share of period total loss.



It can be seen that the big one in January 1993 amounts to one quarter of all losses during the period. Also, the big one turns out to be three times the second largest. What about the next big one, then? Will it at most equal the biggest one so far? Will it double it? Will it again be three times the biggest one known before? Or will it be even worse, say, perhaps ten or twenty times as big?

Before 1993 the biggest loss occurred in January 1983, rather early in the observation period. Already one year later there was another one almost as big. Yet we had to wait another ten years for a new record event, but – as we already have seen – when it occurred it was three times the old record event. This relation is not exceptional. In US the present hurricane record, 1992 hurricane Andrew, happens to be three times the former record, 1989 hurricane Hugo. In Japan the 1991 thypoon Mireille is at least six times the previous worst case. In Europe the present record, 1990 winter storm Daria, doubled the so called winter storm 87J, see [4]. Thus, perhaps unexpectedly, when a wind storm or hurricane record is beaten, then, by experience, the excess loss amount can be overwhelming. As will be shown in the worked out examples this potential risk is a built in property of the proposed methods.

4. Traditional accumulation control methods

4.1. The concept of PML

The traditional approach seems to be to think in terms of the worst possible event. The most extreme position is to put an upper limit to exposure in any one event by accumulation control, summing up sums insured, and by not underwriting beyond this limit, either for the portfolio as a whole or by what is believed to be mutually exclusive zones. We believe this to be the original Probable Maximum Loss (PML) approach in the context of wind storm exposure.

The only information used in this approach is sum insured on policies with wind storm exposure, possibly with location of the exposure. Here, oddly enough, loss experience, thus, is thought of as containing no information at all on future losses.

4.2. The engineering simulation PML approach

A PML estimate like the one described above is far too conservative, since most losses in wind storm insurance are only partial losses. Accordingly, companies and risk assessment agencies today seem to be less restrictive and allow for an ”engineering simulation PML” approach, which typically accounts for partial losses by means of modelling loss ratios as some function of the maximum of mean or gust wind speed. When such a function has been accomplished it is applied either to historical events with known maximum wind speeds at different locations but with todays exposure, or else to some kind of ”worst case”, as to path and as to maximum wind speed.

From a methodological point of view there appears to be at least three serious draw backs with this approach. Firstly, a worst case prediction necessarily is ambiguous; it may be that no two opponents can agree on the same worst case. Secondly, the prediction error seems to be unmeasurable, i.e. the quality of the prediction will be unknown. Thirdly, perhaps less obviously, according to experience the modelling of the complex physical processes which are supposed to cause the damages turns out to be rather delicate.

This approach typically requires a lot of information: for each event

- wind velocities
- sum insured
- loss incurred

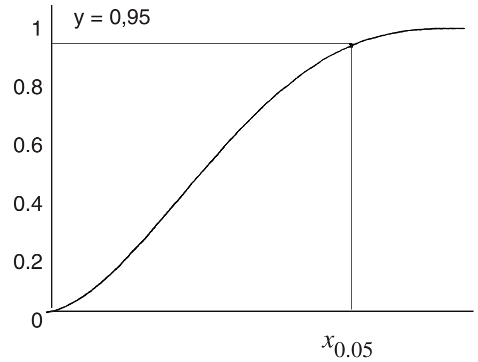
distributed over a grid surface of suitable resolution. Thus, the data in Figure 1 alone is not sufficient in this approach.

5. Quantile PML – a new means for accumulation control

In this paper our main objective is to present, in the context of assessing wind storm exposure, the quantile PML concept. This approach is in the tradition of economic decision theory, which postulates that optimal business decision making involving an uncertain outcome, has to consider all possible outcomes and their associated probabilities.

The quantile PML concept was introduced by Kremer [10],[11],[12] and recently further explored by Rootzén and Tajvidi [13], as an alternative risk assessment method. From a statistical point of view the concept is trivial. To the layman the "p×100:th upper quantile", x_p, is a statistical concept denoting a quantity along the x-axis such that there is a certain probability, p, to obtain a random outcome equal to or in excess of x_p, see Figure 3.

Figure 3. The probabilty distribution function and the upper 5% quantile of a continuous random variable.



Thus, percentiles are points along the x-axis, such that the distribution function grows one percent unit moving right from one percentile to the next. If the distribution is concentrated in a small interval, then the percentiles will be close to each other in that interval; if the distribution is dispersed, as is the case with wind storm losses, then the percentiles will be spread out. Accordingly, the perhaps more familiar "quantiles" known as the quartile and the median cuts the distribution into pieces such that each interval contains 25% and 50% of the distribution, respectively.

On time horizon T years ahead the upper p×100% quantile PML, $x_{T,p}$ in the model discussed in Section 7.3.3 below is given as

$$x_{T.p} = u + \frac{\sigma}{\gamma} \times \left(\left(\frac{\lambda T}{-\ln(1-p)} \right)^{\gamma} - 1 \right) \quad (1)$$

where p is a number in the interval [0,1] , and (u, λ, σ, γ) are unknown parameters to be suitably estimated. Formula (1) is defined in the context of the Generalized Extreme Value distribution (GEV), which will be explained in Section 7. The formula (1) is derived from (14) by solving for x.

Example 1. *For the LF portfolio it seems reasonable to let* $(\hat{u}, \hat{\lambda}, \hat{\sigma}, \hat{\gamma}) = (0.9, 3.83, 3.87, 0.7)$.[3] *Inserting this into* (1) *and using* $x_{T,p}$ = (1, 0.1), *gives*

$$\hat{x}_{1.0.1} = 0.9 + \frac{3.87}{0.71} \times \left(\left(\frac{3.83 \times 1}{-\ln(1-0.1)} \right)^{0.71} - 1 \right) = 66 \quad (2)$$

which should be interpreted as follows: "Looking one year ahead the upper 10% LF-group quantile PML in this example is estimated to be 66 million Swedish crowns." Or, equivalently: "The probability that the most severe wind storm loss one year ahead will exceed 66 million Swedish crowns is estimated to be 10% for a portfolio like LF's."

Table 5.1 summarizes other quantile PML estimates when the same procedure is used for selected values of (T,p).

Table 5.1 Estimated quantile PML for various risk levels and time periods

Risk	*1 year ahead*	*5 years ahead*	*15 years ahead*
10%	66	215	473
1%	366	1149	2497

At this stage, it is essential to point out the difference between formula (1) and values obtained by (2) from a methodological point

[3] According to statistical practice the " ^ "-sign above a parameter indicates that the true value is unknown and has been replaced by an estimate. How the estimated parameter values has been selected will be explained in Section 8.2.

of view. It is important to realize that this difference has far-reaching implications, when results obtained are evaluated by decision makers, i.e. top management. Thus, by varying p in (1), for known values of $(u, \lambda, \sigma, \gamma)$, the obtained values of $x_{T,p}$ reflects genuine uncertainty of the natural phenomenon which governs wind storm loss. Without altering underwriting guidelines or building construction regulations, this uncertainty can not be reduced. Management just have to live with it. Decision makers, however, are even worse off, since they are not fully informed on the loss generating process. Thus, although – at the very best – our model perhaps is (approximately) correct, yet the true values of the parameters are unknown and have to be estimated from experience. But since experiences change over time, so do the estimates, thus, illustrating that values computed from (2) necessarily also include a statistical uncertainty, which is added to the unavoidable genuine uncertainty already commented on. The point is illustrated in the following example.

Example 2*. Suppose that company board management guidelines state that there has to be a cover for the company portfolio such that expected waiting time for a wind storm loss in excess of the cover will be at least 100 years. Suggesting a cover ending at 366 MSEK from Table 5.1 might then be too optimistic, since management don't know if this is the true value of the first percentile from above or not, since it is estimated. It can be in error. If the true model, given by (1), were known it might be that the level 366 MSEK was more like, e.g., the fifth percentile from above, making expected waiting time equal to 20 years instead. Taking this into account, management have to be cautious and decide on a more conservative estimate than the values suggested by Table 5.1.*

The statistical estimation problem will be discussed further below. It will be seen that the problem neither can be neglected nor that it yet is fully resolved. Still, for the time being, postponing this problem, suppose the quantiles given by (2) are true. Then from Table 5.1 the expected waiting time for a loss in excess 66 MSEK is ten years, which is well in accordance with observed data. The very same model predicts expected waiting time for a loss in excess of 366 MSEK to 100 years. This predication can not be evaluated by a comparison with data available, since the observation period is too short. But this is not a reason to reject the prediction. On the contrary, we argue that since (2) predicts ”the ten year recurrent wind storm event” well according to data, then it might be that it also predicts ”the one hundred recurrent wind storm event” properly.

The prediction arrived at is also supported by empirical evidence already reported, why it is hypothesized that last years world wide wind storm catastrophes, i.e. 1990 winter storm Daria, 1989 hurricane Hugo, 1992 hurricane Andrew and 1991 typhoon Mireille, all could have been predicted by the quantile PML approach.

6. The Spill Over distribution

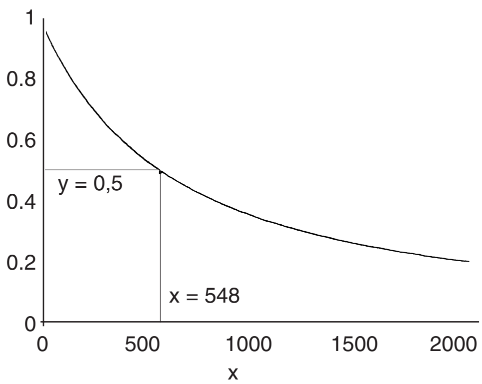
The most common wind storm reinsurance cover is the per event excess of loss (XL) cover. From such a cover the recovery is ”at most l in excess of r, each and every wind storm event”. Here r is the cedent's retention and l is the limit of the cover. Thus, in each and every wind storm event, the reinsured is exposed to ”planned” losses for own account less or at most equal to r. But, because of the limitation of the cover, due to cedent's cost considerations or limited market capacity, the cedent is also exposed to all ”unplanned” losses in excess of $r + l$. In the terminology of reinsurance such losses are called ”spill over” losses. With limited cover such losses can not be neglected from a logical point of view, but, using the traditional PML approach, in practice they often are, recently with fatal consequences for some companies.

Suppose that the next loss will be a spill over loss. The crucial question to the company then is, what size of loss can be expected. Again, with reference to extreme value statistics (cf. Section 7.3 below) and with $v = r + l - u$, it can be argued that the function

$$\bar{F}(x \,|S > r + l) = \left(1 + \gamma \frac{x}{\sigma + v\gamma}\right)^{-1/\gamma} \tag{3}$$

could be used to calculate the probability of a spill over loss, S, of at least size x. Inserting $r + l = u + v = 850$ and once again using LF estimates $(\hat{u}, \hat{\sigma}, \hat{\gamma}) = (0.9, 3.87, 0.71)$ the Figure 4 is obtained.

Figure 4. Estimated conditional probability that a spill over loss will exceed x MSEK excess 850 MSEK.



Putting (4) equal to 0.5 and solving for x, it follows that the median $m\ (r + l)$ of a loss excess of reinsurance limit $r + l = u + v$

$$m\,(r + l\,) = \frac{\sigma}{\lambda}(2^{\lambda} - 1) + v\,(2^{\lambda} - 1) \tag{4}$$

where still $v = r + l - u$.

Example 3*. With $r + l = u + v = 850$ again, the LF-group estimate becomes*

$$\begin{aligned} \hat{m}\,(850) &= \frac{3.87}{0.71}\left(2^{0.71} - 1\right) \\ &\quad +849.1\left(2^{0.71} - 1\right) \\ &= 548 \end{aligned} \tag{5}$$

Thus, given the mathematical model, the LF-group parameter estimates, and an assumed per event catastrophy XL cover ending at MSEK 850, we estimate that a spill over loss is above MSEK 548 is as likely as a spill over loss below MSEK 548.

In practice, again, decision makers have to consider the statistical errors contained in the curve in Figure 4 and the estimate (5), and analyse them according to the principles discussed in Section 8.

7. The modelling of maximum loss amount

7.1. Introduction

Embrechts et al [5] and Beirlant et al [1] discuss modelling of extremal events in insurance from a mathematical point of view. Early Swedish contributions are Benktander [2] and Jung [9].

In practical wind storm loss modelling the most compelling question is, if the existence of such a model can be justified by rational reasons. The question posed will not be answered here. Instead we will tactically show that, if it is reasonable to model wind storm losses, then the proposed approach seems to be the only possible one.

The model from which formulas in Sections 5 and 6 are derived are not new. E.g. Hosking et al [8] report that the same approach is widely used in United Kingdom to model annual maximal flood level arising from separate wind storm events.

7.2. The notion of maximum loss amount

In the context of rate making, the probability distribution of period total claim amount has been analysed extensively in the mathematical theory of risk, or often simply, "risk theory". The probability distribution of the maximum individual loss within a certain period of time, is related to the distribution of period total claim amount, since it is contained in that total, sometimes with a great influence on the distribution, affecting e.g. rate making. Still it has been considered far less in risk theory so far.

Besides, the distribution of the maximum is of interest on its own, since it in a certain sense mirror relevant total exposure to a single natural or man made disaster, which in many countries is supposed to be limited according to public supervision and regulation conditions.

Using powerful standard mathematical notation, it is a very simple exercise to write down a true expression for the maximum individual gross loss amount, M_N, for a book of business underwritten during a certain period of time. The maximum individual loss among the N losses occurring, $X_1, ...,X_N$, can, thus, simply be written as

$$M_N = \max (X_1, ...,X_N) \qquad (6)$$

Seen as a mathematical function, the max-function simply returns the largest outcome obtained.

Using order statistics notion, the expression for M_N could be further simplified as

$$M = X_{[1:N]} \qquad (7)$$

where $X_{[1:N]}, X_{[2:N]}, ...X_{[N:N]}$ denotes the N observations $(X_1, ...,X_N)$ ordered decreasingly by size, i.e.

$$X_{[1:N]} \geq X_{[2:N]} \geq .... \geq X_{[N:N]}$$

However, to be of any value in practical decision making, we have to look for a means to distribute total mass of probability over all possible outcomes of M_N..

7.3. Modelling the distribution of maximum loss amount

7.3.. Asymptotic theory

In asymtotic theory the behaviour of a random variable, such as e.g. M_N, with a distribution which depends on size, N, is analysed in the case where one let N grows towards infinity, i.e. one assumes access to a sample of unlimited size. Maybe unexpectedly, it often turns out that mathematics can be simplified by doing this. The logic of the reasoning is: if there is an asymtotic distribution, then the finite distribution will come successively closer – or converge – towards this limiting form, step by step, when sample size is increased. If the limiting distribution exists, it is further hopefully assumed that the limiting form can be a good substitute or approximation for the exact finite distribution which is not obtainable.

Early useful asymtotic results on the distribution of the maximum, M_N, in a sequence of random variables goes as far back in history as to a famous 1928 paper [6] by the two distinguished English statisticians, sir Ronald A. Fisher and L.H.C. Tippett. In that paper they proved a theorem which afterwards sometimes is considered "the fundamental theorem of extreme value theory". They proved the following remarkable result:

Theorem. *(Fisher-Tippett Limit Law of Maxima Theorem) If the distribution functions of the maximum in a sequence of maxima* $\{M_N, N = 1, 2, ...\} = \{\max (X_1, ...,X_N), N = 1, 2, ...)\}$ *converges to a non-degenerate distribution function,* $G(x)$*, under a linearly normalization, i.e. if, when* $N \to \infty$*, we have*

$$P((M_N - a_N) / b_N = \leq x) \to G(x) \qquad (8)$$

for some constants a_N *and* $b_N > 0$*, then* $G(x)$ *has to be*

$$G(x) = \exp\left\{-\left(1+\gamma\frac{(x-\mu)}{\sigma}\right)_+^{-1/\gamma}\right\} \quad (9)$$

for some shape parameter γ, location parameter μ and scale parameter $\sigma > 0$.

The "+" signifies "positive part", which means that $x_+ = \max(0, x)$.

The quoted theorem uses a strict mathematical language. For practical purposes, however, it simply states as an unavoidable logical truth, that if there is a linear normalized limiting distribution for the maximum outcome in your application, then it has to be on the form (9). However, it doesn't state, that the maximum outcome in your application *has* such a limiting distribution. It just says, if there is one, then it has to be (9). Another possibility is that there is no limiting distribution, or – typically if the outcomes are limited in size upwards – the distribution will be degenerated and equal this limit with probability one.

The distribution function *G* in (9) is known as the Generalized Extreme Value distribution (GEV). It is a result in what has been called "classical extreme value statistics".

7.3.2. Compound loss distributions

The result in the preceding Section may be important in some cases, but yet it seems ill-conditioned in the context of predicting the most severe wind storm loss next year, since the assumptions do not fit nicely into the problem at hand. E.g. we do not expect a lot of wind storm events in any year.

Still the classical model has proved useful through a re-parametrisation suggested by the so called "peaks over thresholds" (POT) models of modern extreme value statistics. These models allow for a compound loss distribution, i.e. a loss generating process which is built up by two different random processes; one describing the occurrences of losses in time and another describing loss severity when a loss actually occur. This is similar to the most common approach in modelling total claim amount, historically known as "collective risk theory", see e.g. Daykin et al [3].

Before discussing POT models in more detail, we have to present one important candidate of each kind; the Poisson distribution and the Generalised Pareto distribution, respectively.

The occurrence distribution candidate is the much used Poisson distribution. It can be defined in the following way:

Definition 1. A *random variable, N, which has a probability function of the following form*

$$\begin{aligned} p_N(x) &= P(N=x) \\ &= e^{-\lambda T}\frac{(\lambda T)^x}{x!} \end{aligned} \quad (10)$$

when x = 0, 1, 2, ... is said to have a Poisson distribution.

A typical case where the Poisson distribution is called upon is given in the following example:

Example 4. *Suppose wind storm events occur purely randomly in time. We state, without a lenghty discussion on what is meant by "purely random" and on mathematical details, that traditional mathematical modelling of this situation ends up with the Poisson point process for the occurrences, and the Poisson distribution, as given by (10), as the counting process counting the wind storms occurrences per every T years. Thus, with N denoting the random variable "number of wind storm events in next T years", N has a Poisson distribution and the probability of the event 'N = x', is given by (10), where λ is a parameter denoting an unknown, theoretical entity properly interpretable as "the expected number of wind storm events per year".*

A less traditional statistical distribution in insurance is the Generalised Pareto distribution:

Definition 2. *A random variable, X, which has a probability distribution function on the following form*

$$\begin{aligned} H(x) &= P(X \leq x) \\ &= 1 - \left(1 + \gamma \frac{x}{\sigma}\right)_{+}^{-1/\gamma} \end{aligned} \tag{11}$$

when $1 + \gamma x / \sigma > 0$ *is said to have a Generalised Pareto distribution (GPD).*

Here $\sigma > 0$ is a scale parameter, which means that using (11), σ has to be adjusted according to the unit of measurement in use. The γ parameter is a shape parameter, which can take any value. The "+" signifies "positive part", which means that for γ negative, $H(x) = 1$ for $x \geq \sigma / \gamma$, i.e. the distribution has the positive finite endpoint σ / γ. The special cases $\gamma = -1$ and $\gamma = 0$ yield, respectively the uniform distribution on interval $(0, \sigma)$ and the exponential distribution with distribution function

$$H(x) = 1 - e^{-x/\sigma}, \text{ for } x > 0 \tag{12}$$

If $\gamma > 0$, which is to be expected in insurance applications modelling large claims, then (11) is close to the, so called, Pareto distribution of the first kind, defined by the distribution function

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= 1 - \left(\frac{c}{x}\right)^{\alpha} \end{aligned} \tag{13}$$

when $\alpha, c > 0$, and $x \geq c$. In Rytgaard [14] this is called the "European" Pareto distribution. It is extensively used in insurance and reinsurance rate making. In (13) c is the scale parameter, which typically is supposed to be controlled by the researcher and typically used as the "data capture limit" (Hesselager [7]) and α the shape parameter to be estimated. This is not the case in (11), where σ and γ both are supposed to be estimated simultaneously. However, we will see later that when GPD is used in a POT model, it is used to model outcome $Y = X - u$, which is an observed outcome X minus a fix "high level", u, i.e. a new parameter to be introduced, with an interpretation similar to c in (13).

In the case where $\gamma > 0$, suppose $Y = X - u$. Then, as demonstrated in Rootzén et al [13], the relationship between (11) and (13) is:

i) If X has a Pareto distribution of the first kind, given by (13), then Y has an exact GPD, given by (11), with $\sigma = u\alpha$ and $\gamma = 1/\alpha$
ii) If Y has a GPD, given by (11), then the tail of X is approximately equal to that of a Pareto distribution of the first kind, given by (13), since when $x \to \infty$ we have

$$\begin{aligned} P(X \leq x) &= 1 - P(X > u) \\ &\quad \times \left(1 + \gamma \frac{x - u}{\sigma}\right)_{+}^{-1/\gamma} \\ &\sim 1 - constant \times x^{-1/\gamma} \end{aligned}$$

which is (13) with $constant = c^{1/\gamma}$ and $\alpha = 1/\gamma$

Taking nothing else into account, GPD seems to dominate Pareto, as a more flexible distribution with a wider range of application. This is the case also in the context of insurance. In the LF case, this extra flexibility was useful in [13], since there is a good fit of the wind storm losses to a GPD, but deviations from a Pareto distribution were detected.

7.3.3. The POT model

We are now well prepared to introduce the POT model in the context of company losses caused by wind storm insurance. Interested in large claims only, we consider modelling losses in excess of a suitably chosen "high level" u. We assume that the number of wind storms causing losses in excess of u is generated by a Poisson process, so that the random number of such event per year, N, can be modelled with a Poisson distribution, given by (10). Next we assume that for each such occurrence, the severity of the loss is generated by a process, such that the loss amount exceeding u, i.e. $Y_i = X_i - u$, $i = 1, 2, ..., N$, has a GPD, given by (11). At last we make the usual independency assumptions, i.e.

- Y_1, Y_2, ..., Y_N are independent of each other, and
- the distribution of each of Y_1, Y_2, ..., Y_N is also independent of N.

If these assumptions hold (approximately) true, then, looking T years ahead, it is a logical necessity that the probability that the maximum loss, M_T, in that period, will be at most $u + v$, is (approximately)

$$P(M_T \leq u + v) = e^{-(1 + \gamma A(v, \lambda T, \sigma, \gamma))^{-1/\gamma}} \quad (14)$$

where

$$A(v, \lambda T, \sigma, \gamma) = \frac{v - ((\lambda T)^{\gamma} - 1)\,\sigma/\gamma}{\sigma\,(\lambda T)^{\gamma}}$$

when $v > 0$, which is the GEV defined by (9), with GEV parameters (μ, σ', γ') equal to $\{((\lambda T)^{\gamma} - 1)\ \sigma/\gamma,\ \sigma(\lambda T)^{\gamma}, \gamma$ are the parameters in the compound Poisson/GPD model for the maximum, M_T, T years ahead. However, since the Fisher-Tippett theorem holds asymptotically, formula (9) is only approximate for finite sample sizes, while formula (14) is exact, provided that the peaks over tresholds model is true.

If the compound Poisson/GPD model holds true, formula (14) can be used to compute the probability of any event involving M_T.

Example 5. *Using LF parameter estimates used in Section 5 and 6 again, letting* $u + v =$ 136 *, i.e. current LF record loss, and neglecting impact of statistical error, then the probability of a new LF wind storm loss record next year is, with*

A (135.1, 3.83, 3.87, 0.71) = 12.589

inserted into (14)

$P(M_1 > 136 = 1 - P(M_1 \leq 136) =$

$= 1 - e^{(-1 + 0.71 \times 12.589)^{-1/0.71}} = 0.039$

or 3.9%. *Thus, with current exposure and constant money value, the expected waiting time for a new record loss seems to be* 1/0.039 ≈ 26 *years.*

Example 6*. Given a per event XL-cover ending at r + l = u + v =* 850 *and LF parameter estimates, again neglecting the impact of statistical error, then the probabilty of a next year spill over loss becomes, with*

A (849.1, 3.83, 3.87, 0.71) = 83.697

inserted into (14)

$P(M_1 > 850 = 1 - P(M_1 \leq 850) =$

$= 1 - e^{(-1 + 0.71 \times 83.697)^{-1/0.71}} = 0.003$

or 0.3 %. *The expected waiting time for the recurrence of the event seems to be* 1/0.003=333 *years.*

7.3.4. Some properties of the POT model

With reference to some important properties which holds for the POT model given by (14), Rootzén et al [13] argues that this statistical model is almost the only possible one. Two such properties given in the following two propositions characterize GPD. This means that given a certain property, this property uniquely determines a certain family of distribution functions which have the property and all others do not have the property.

Proposition 1*. The POT model is characterised by being stable under an increase of the level.*

Proposition 2*. The Fisher-Tippett Limit Law holds if and only if the POT model holds.*

The first proposition says: If excesses over level u come as a Poisson process and the sizes of the excesses are GPD and independent, then the excesses over a higher level $u+v$ (for $u>0$) belong to the same families of distributions; the excesses over $u+v$ have a Poisson distribution and the sizes of the excesses are again independent GPD. No other distributions have this property. Thus, suppose you insist on the same model for any high level u, then the POT model is unavoidable. All other choices means that if u is altered, then the families of distributions in your model also have to be altered.

The second proposition brings a similar message: Accepting the existence of a limiting distribution for the maximum is equivalent to accepting that observations are generated by the POT model.

8. Statistical estimation in the peaks over tresholds model

8.1. Inference of unknown true model behavior from sample information

Here we will recall the problem going from e.g. formula (1) to (2). When theoretical models and reasoning are applied in practical decision making, one critical and often neglected step in the decision process is the substituting of unknown parameters in the theoretical models with actual numbers considered for the numerical calculations. In the context of mathematical statistics the art of doing this is a most important area of research, alternatively known as Theoretical Statistics, Statistical Analysis, Statistical Inference or Statistical Estimation. Scientists specializing in this area are concerned about the fundamentals and methods of scientific inference. From a statistical point of view the problem posed is the following: ”What can be said about a population of measurements, when all we got is a sample of measurement from that population?” Or adopted to risk theory context: ”What can be said about the unknown parameters in a hypothesized claim generating process, which over time will generate an unlimited sequence of claims, but of which only a subset has been reported up to today?”

In Sections 5 and 6 we argued that limited model knowledge force decision makers to apply estimating procedures in decision making, and that this is making risky decision making even riskier. There is, thus, a need to reduce the added estimation risk as much as possible by arriving at estimates with a minimal error. Statistical modelling makes estimation error observable and measurable. To the general public this approach is standard today in the context of professionally conducted surveys or polls, where estimated numbers always come together with estimates of margin of error.

8.2. Estimation principles and efficiency

Statistics is about analysis of data. Consequently statistical estimation assumes that there are some data available carrying information on the distribution function to be estimated. How to do it might at first glance seem totally ambiguous. However, there are principles to rely on. Applying such principles results in quantities which are functions of observations. Such functions are essential to statistical estimation and such a function is most often called an estimator. Since in analysis the observations themselves are considered to be random, estimators also become random variables.

An estimator which in some well defined sense tends to produce estimates closer to the unknown true value than any other proposed estimator usually does, can be considered optimal or efficient in this sense. The analysis of the behavior of estimators is therefore an important field for research.

One estimation principle often used is just an advice to adjust the unknown distribution function by replacing the unknown parameters in such a way that the first low moments in the theoretical and empirical distribution function coincide. This technique is known as the Method of Moments (MOM). The related method of Probability Weighted Moments (PWM) has proved useful when looking for estimators for GEV and GPD. The corresponding estimators are called method of moment estimators and probability weighted moment estimators, respectively.

However, the most popular approach today is to apply the maximum likelihood (ML) principle, originally suggested in a 1922 paper by Ronald A. Fisher. His principle turns an estimation problem into an optimization problem. According to the likelihood principle a set of unknown parameters ought to be replaced by values such that the probability of obtaining the observations obtained is maximized. The popularity of the likelihood approach is due to the fact that when there is an optimal estimator in a given application, then it often turns out, that this is the estimator derived according to the maximum likelihood principle. This is particularly true in large sample context. In small samples, however, the maximum likelihood estimators can sometimes be inferior to others.

8.2. Estimation in the POT model

8.2.1. Introduction

Estimation in the POT involves estimation of λ in the Poisson distribution, σ and γ in GPD and u in the compound POT model.

Monte Carlo simulation experiments reported in Rootzén and Tajvidi [13] suggest that LF wind storm quantile PML estimates by formula (2) can deviate considerably from the true value. It also turns out that properies of present standard methods are not yet fully explored.

8.2.2. Estimating Poisson parameter λ

Given $N_S = N_1 + N_2 + ... + N_s$ wind storm loss events in each of year 1, 2, ..., s, the most natural estimator of λ is

$$\hat{\lambda} = \frac{N_s}{s} = \frac{\sum_{i=1}^{s} N_i}{s}, \quad \text{for } i = 1, 2, \ldots, s \tag{15}$$

which also is optimal in the sense of statistical efficiency. Putting $s = 12$ and observed N_s into (15) gives

$$\hat{\lambda} = \frac{46}{12} = 3.83,$$

which is used throughout in the examples.

The estimation in GPD is less obvious, but has been considered by Hosking et al [8].

8.2.3. Estimating (σ, γ) in GPD

Hosking et al [8] report on the ML, MOM and PWM estimators of GPD parameters (σ, γ), respectively, the asymptotic properties of the different estimators, and on asymptotic methods for estimating statistical error. Small sample properties are examined with Monte Carlo simulation.

The estimating functions will not be given here, but can be found in the Hosking and Wallis paper, from which we summarize:

1 the second moment of GPD doesn't exist for $\gamma \geq 0.5$, which means that in this case MOM can be expected to perform less well,

2. in the context of PWM two different, asymptotically equivalent, estimators are proposed, one using the sample ordered by size, the other using the empirical distribution function,

3. the ML estimators can not be given in closed form and have to be obtained by numerical methods,

4. Monte Carlo simulations suggest that unless sample is 500 or more, estimators derived by MOM or PWM are more reliable than those derived from ML principle, for $1/2 < \gamma < 1/2$.

In estimating a LF wind storm loss GPD on losses as reported in Figure (1), PWM and ML estimators performed similar, but Rootzén and Tajvidi suggest the ML estimates $(\hat{\sigma}, \hat{\gamma}) = (3.87, 0.71)$ to be used in the examples, since the PWM estimators seem to systematically produce heavy underestimates on estimating quantiles in the POT model.

8.2.4. Estimating u in the POT model

For the choice of u in the POT model it is reported that there are theoretical suggestions, but they don't seem to solve the practical problem. Instead the choice of level has to be made from subject matter knowledge, from looking at different diagnostic plots, e.g. QQ-plots, mean excess or median excess plots, and on experimenting with different levels. If the model produces very different results for different choices, the result of course should be viewed with more caution.

For LF wind storm data Rootzén and Tajvidi used $\hat{u} = 0.9$, which has also been used in the numerical examples.

8.2.5. Estimating $x_{T,\,p}$ in the POT model

Using proposed estimators and reported estimates finally gives formula (2) as the used estimate of LF wind storm quantile PML, $x_{T,\,p}$, in (1). Monte Carlo simulation experiments suggest that this estimator produces less biased estimates than present small sample standard suggesting PWM estimators. Still, the quantile PML seems to be systematically underestimated. Depending on γ, Monte Carlo simulation experiments by Rootzén and Tajvidi suggest that the median of such estimates underestimate true value in the range of 10-15%, when sample size $n = 46$. The downward deviation from true value can be considerable, with first quartile in the simulated distributions in the range of only 55-70% of the true value depending on γ. That is, 25% of simulated estimates are even worse off, meaning that there is a non negligible risk that the point estimates in table 5.1 are too optimistic. Also, standard small samples methods for interval estimates are reported to have much lower coverage probabilities than the nominal.

9. Summary and needs of further research

It has been shown that formulas (1), (3) and (4) are extremely easy to apply, given estimates on the unknown model parameters. Also several theoretical arguments have been put forward justifying them from a theoretical point of view. Thus, in summary the proposed methods turn out to be

- far more powerful, versatile and easier to apply from a practical point of view than tools traditionally used
- well motivated and – as it seems – the only possible ones from a theoretical point of view.

The approach proposed seems also to be the most parsimonious one, since the calculations just depend on total losses in the observed events, as in Figure 1. However, a more elaborate extreme value model, which takes individual claims, spatial and climatological information into account, might possibly reduce the statistical error and give improved predictions (Tajvidi [15]).

Below, finally, follows a list of areas where further research is needed:

1. bias reduction and small sample interval estimates for the estimators in the POT model
2. use of information on individual claim siz-

es and on the distribution of the number of claims in the estimation of accumulated loss sizes

(a) in order to facilitate and improve on exposure adjustments due to changes in policy wording, social inflation and market shares over time

(b) as a means to reduce sampling error, if possible

3. use of meteorological information to improve parameter estimates

4. Bayes and credibility methods for the POT model

5. accumulation control methods for a reinsurance company

6. developing decision making rules, as a support to management to decide on company optimal trade off between spill over risk and other company goals.

References

[1] Beirlant, C.D., Tengels, J.L. and Vyuckier, P., *Practical Analysis of Extreme Values,* Leuven: Leuven University Press, 1996.

[2] Benktander, G., *A note on the most "dangerous" and skewest class of distributions,* ASTIN Bulletin, 2, 1962, 387-390.

[3] Daykin, C.D., Pentikäinen, T. and Pesonen, M., *Practical Risk Theory for Actuaries*, London: Chapman & Hall, 1994.

[4] DYP, "The World's Top 20 Catastrophic Losses", *Catastrophe Reinsurance Newsletter,* No 30, 1995, 149-151.

[5] Embrechts, P., Klüppelberg, C. and Mikosh, H., *Modelling of extremal events in insurance and finance*, Zürich: book manuscript, Department of Mathematics, ETH.

[6] Fisher, R.A. and Tippett, L.H.C., *Limiting forms of the frequency distribution of the largest or smallest member of a sample*, Journal of Royal Statistical Society B, 24, 1928, 180-190.

[7] Hesselager, O., *A Class of Conjugate Priors with Application to Excess-of-Loss Reinsurance*, ASTIN Bulletin, 23, 1993, 77-93.

[8] Hosking, J.R.M. and Wallis, J.R., *Parameter and Quantile Estimation for the Generalized Pareto Distribution*, Technometrics, 29, 1987, 339-349.

[9] Jung, J., *On the use of extreme values to estimate the premium for an excess of loss reinsurance.* ASTIN Bulletin, 4, 1964, 178-184.

[10] Kremer, E., *On the Probable Maximum Loss*, Blätter der Deutschen Gesellshaft für Versicherungsmathematik, 1990.

[11] Kremer, E., *A Probable Maximal Loss Theory*, Research Report, Verain zur Förderung der Angewandten Mathematischen Statistik und Riskotheorie, Hamburg: 1991/93.

[12] Kremer, E., *More on the Probable Maximum Loss*, Blätter der Deutschen Gesellshaft für Versicherungsmathematik 11, 1994, 319-326.

[13] Rootzén H. and Tajvidi, N., *Extreme value statistics and wind storm losses: a case study*, Research report 1995:5, Gothenburg: Department of Mathematics, Chalmers University of Technology, 1995 (To appear in Scandinavian Actuarial Journal).

[14] Rytgaard, M., *Estimation in the Pareto Distribution*, ASTIN Bulletin, 20, 1990, 201-220.

[15] Tajvidi, N., *Multivariate generalised Pareto distributions*, Research report 1995:17, Gothenburg: Department of Mathematics, Chalmers University of Technology, 1995.