Natural Catastrophe Models for Insurance Risk Management

PAVLA JINDROVÁ, VIERA PACÁKOVÁ
Institute of Mathematics and Quantitative methods
Faculty of Economics and Administration
University of Pardubice
Studentská 95, 532 10 Pardubice
CZECH REPUBLIC
Pavla.Jindrova@upce.cz;Viera.Pacakova@upce.cz http://www.upce.cz

Abstract: - Catastrophic events are characterized by three main points: there are relatively rareness, there are statistical unexpected and there have huge impact on the whole society. Insurance or reinsurance is one way of reducing the economic consequences of catastrophic events. Risk management of insurance and reinsurance companies have to have available relevant information for estimation and adjusting premium to cover these risks. The aim of this article is to present two of the useful methods – block maxima method and peaks over threshold method. These methods use information from historical data about insured losses of natural catastrophes and estimates future insured losses. These estimates are very important for actuaries and for risk managers as one of the bases for calculating and adjusting premiums of products covering these types of risks.

Key-Words: - Block maxima model, catastrophic events, insured losses, modelling, peaks over threshold, risk.

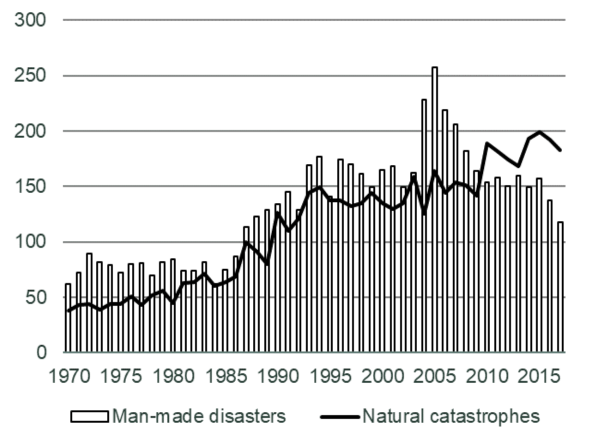
1 Introduction

The catastrophic events have a huge impact for a whole society. This impact is deep and long. Two next points characterize these events too. There are relatively rare and statistically unexpected.

The catastrophic events are divided into two groups according to their cause. By the Emergency Events Database (EM-DAT) [1] the first group are natural catastrophes caused by the natural influences such as the effects of natural influences such as geological disasters (i.e. earthquake, mass movement, volcanic activity), meteorological disasters (i.e. extreme temperature, fog, storm), hydrological disasters (i.e. flood, landslide, wave action), climatological disasters (i.e. drought, glacial lake outburst, wildfire), biological disasters (i.e. epidemic, insect infestation, animal accident) and extra-terrestrial disasters (i.e. impact, space weather).

The second group consists of catastrophic events caused by human activity, i.e. man-made disasters, such as industrial accident (i.e. chemical spill, collapse, explosion, fire, gas leak, poisoning, radiation, oil spill and other), transport accident (i.e. air, road, rail, water) and miscellaneous accident (i.e. collapse, explosion, fire and other).

Fig. 1 shows the number of catastrophic events in the period 1970-2017. We can see growing trends in number of man-made disasters and in number of natural catastrophes too. In terms of Sigma Swiss Re criteria (Table 1), there were 301 catastrophes worldwide in 2017, less than the number 329 in 2016. There were 183 natural catastrophes (compared with 192 in 2016), and 118 man-made disasters (compared with 137 in 2016). [2]

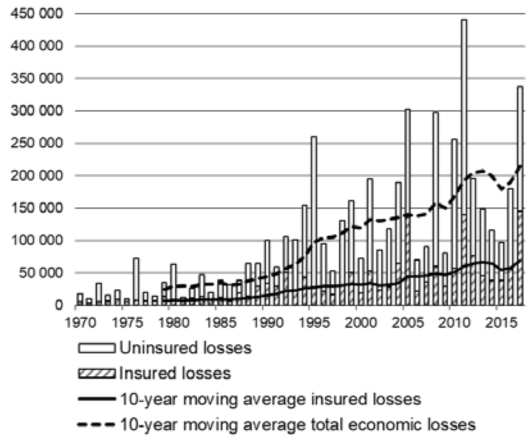


Source: Sigma Swiss Re, [2]

Fig. 1. Number of catastrophic events, 1970-2017

Fig. 2 shows the total economic losses by catastrophes. When we talk about total economic losses that means the sum of insured and uninsured loses.

Fig. 2 shows the difference between insured and economic losses during time period 1970-2017. This figure shows 10-year moving average of total economic losses and insurance losses too. We can see increasing trend in both cases but we can see also increasing differences in these trends.



Source: Sigma Swiss Re, [2]

Fig. 2. Insured vs uninsured losses, 1970 – 2017, in USD billion at 2017 prices

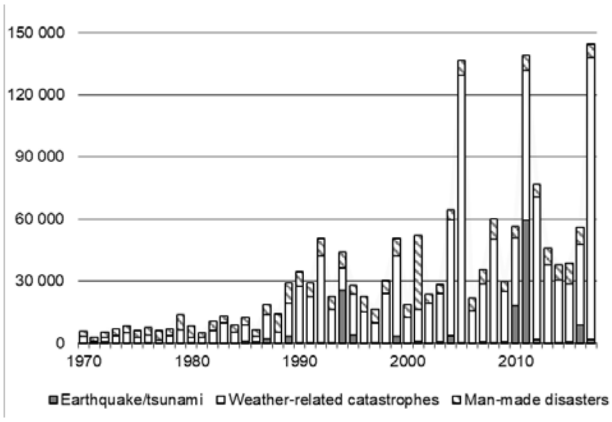
To classify event as a catastrophe according to Sigma criteria [2]-[5], the economic losses, insured claims or casualties associated with an event must exceed one of the thresholds, which are shown in Table 1 for year 2017.

Table 1. Sigma event selection criteria for catastrophic events for 2017

Insured losses (in USD million)	
Maritime disasters	20.3
Aviation	40.7
Other losses	50.5
or **Total economic losses** (in USD million)	101.0
or **Casualties**	
Dead or missing	20
Injured	50
Homeless	2000

Source: Sigma Swiss Re, [2]-[5]

In Fig. 3 we can see the major insured losses suffered mainly due to natural influences in the period 1970-2017. The highest insured catastrophe losses caused by hurricanes Katrina, Rita and Wilma were occurred in 2005, in 2011 there were especially the devastating consequences of the earthquakes in Japan and New Zealand and of the flood in Thailand and in 2017 catastrophic event caused by hurricanes Harvey, Irma and Maria. The highest man-made catastrophes in the period 1970-2017 were terrorist attacks on September 2001 in the USA. [2]



Source: Sigma Swiss Re, [2]

Fig. 3. Insured losses in the period 1970-2017

2 Problem Formulation

Catastrophe modelling helps insurers and reinsurers better assess the potential losses caused by natural and man-made catastrophes. Natural catastrophe models have been developed for a wide range of catastrophic risks and geographic territories worldwide. The Pareto model is very often used for description of the random behavior of extremal losses [6]. Especially quantile methods provide an appropriate and flexible approach to the probability modelling needed to obtain well-fitted tails [7]-[8]. Application of quantile modelling methods has its foundation in the Order statistics theory [9].

Extreme value theory (EVT) [10]-[12] is a promising class of approaches to modelling catastrophe losses. These methods although originally utilized in other fields such as hydrology or operational risk [13]. There are two main types of models in EVT: block maxima models and peaks over threshold (POT) models. The more traditional are Block maxima models, which are collected from the largest observations of large samples of historical data. The whole sample is divided into equal non-overlapping time intervals and the biggest value from each interval is used for modelling [14]-[20]. In the more modern approach using POT model (or the threshold exceedances model) the large enough threshold is determined and the observations above this threshold are considered for extremal losses modelling [21]-[24]. The Extreme value methods do not predict the future with certainty, but they do offer models for explaining the extreme events in the past. These models are not arbitrary but based on rigorous mathematical theory concerning the behavior of extrema [25]-[27].

For the objective of this article the Block maxima model and the Peaks over threshold model have been chosen for catastrophe modeling based on real data of insured losses of natural catastrophes published by Swiss Re Sigma [2]-[5].

2.1 Using methods for modelling

2.1.1 The Block maxima model

The block maxima models are models for the largest observations collected from large samples of identically distributed observations.

The Fisher-Tippett theorem [25] is the fundamental result of the Extreme Value Theory (EVT) and can be considered the same important as the central limit theorem for statistical inference. The theorem describes the limiting behavior of appropriately normalized sample maxima.

Suppose catastrophe losses are independent, identically distributed random variables denoted by X1, X2, …, whose common distribution function is is $F_X(x) = P(X \le x)$, where $x > 0$.

Extreme Value Theorem [10]: Suppose $X_1, X_2, \dots$ are independent, identically distributed with distribution function $F_X(x)$. If there exist constants $c_n > 0$ and $d_n \in R$ such that

$$\frac{M_n - d_n}{c_n} \to Y, \quad n \to \infty$$

where $M_n = \max(X_1, \dots, X_n)$, Y is non-degenerate with distribution function G. Then G is of one the following types:

Gumbel

$$\Lambda(x) = exp\{-exp^{-x}\}, \qquad x \in R$$

Frechet

$$\Phi_\alpha(x) = \begin{cases} 0 & x \le 0 \\ exp\{-x^{-\alpha}\} & x > 0 \end{cases}$$

Weibull

$$\Phi_\alpha(x) = \begin{cases} exp\{-x^{-\alpha}\} & x < 0 \\ 0 & x \ge 0 \end{cases}$$

These three types of limiting distribution there are in standard form. We can parameterize them within the location and scale families:

Gumbel

$$\Lambda(x) = exp\left\{-exp\left[-\left(\frac{x-d}{c}\right)\right]\right\}, \qquad x \in R$$

Frechet

$$\Phi_\alpha(x) = \begin{cases} 0 & x \le d \\ exp\left\{-\left(\frac{x-d}{c}\right)^{-\alpha}\right\} & x > d \end{cases}$$

Weibull

$$\Phi_\alpha(x) = \begin{cases} exp\left\{-\left(\frac{x-d}{c}\right)^{-\alpha}\right\} & x < d \\ 0 & x \ge d \end{cases}$$

The generalized Gumbel, Frechet and Weibull families can be combined into a single family of the Generalized extreme value distributions (GEV) in the form

$$G(x) = exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{1}$$

where

$$1 + \xi\left(\frac{x-\mu}{\sigma}\right) > 0$$

It is straightforward to check the result by letting:

$$\begin{aligned} &\alpha = \frac{1}{\xi} \\ &d = \mu - \frac{\sigma}{\xi} \\ &c = \begin{cases} \frac{\sigma}{\xi} & \text{if } \xi > 0 \\ -\frac{\sigma}{\xi} & \text{if } \xi < 0 \end{cases} \end{aligned} \tag{2}$$

2.1.2 The Peaks over threshold model

The modelling using the Peaks over threshold method follows the assumptions and conclusions in Generalized Pareto Distribution (GPD) Theorem.

Suppose $x_1, x_2, \dots, x_n$ are raw independent observations from a common distribution $F(x)$. Given a high threshold u, assume $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ are observations that exceed threshold u. Here we define the ascendances as $x_i = x_{(i)} - u$ for i = 1, 2, …, k.

Then for a large enough threshold u by the GPD Theorem [27] the generalized Pareto distribution (3) is the limiting distribution for the distribution of the excesses, as the threshold tends to the right endpoint.

The conditional distribution function of variable Y = ($X - u$ / $X > u$) is approximately

$$H(x) = 1 - \left(1 + \frac{\xi x}{\tilde{\sigma}}\right)^{-\frac{1}{\xi}} \quad (3)$$

defined on
$\{x: x > 0$ and $(1 + \xi x/\tilde{\sigma} > 0)\}$, where
$\tilde{\sigma} = +\xi(u - \mu)$

The family of distributions defined by equation (3) is called the generalized Pareto family (GPF). For a fixed high threshold u, the two parameters are the shape parameter ξ and the scale parameter $\tilde{\sigma}$. For simpler notation, we may just use σ for the scale parameter if there is no confusion.

By GPD Theorem x_i may be regarded as realization of independently random variable, which follow a generalized Pareto family with unknown parameters ξ and σ. In case $\xi \neq 0$ the likelihood function can be obtained directly in the from

$$L\left((\xi, \sigma/\mathbf{x})\right) = \prod_{i=1}^{k}\left[\frac{1}{\sigma}\left(1 + \frac{\xi x_i}{\sigma}\right)^{-1/\xi - 1}\right] \quad (4)$$

3 Problem Solution

3.1 Data and exploratory analysis

For modelling by Block maxima model and Peaks over threshold model were used the real data. The analysis focus on 479 insured losses (in USD million) of natural catastrophes in time period from January 2010 to December 2016, published in Swiss Re Sigma 2011-2017 [2]-[5]. The time series plot (Fig.4) allows us to identify the most extreme losses and their approximate times of occurrences.

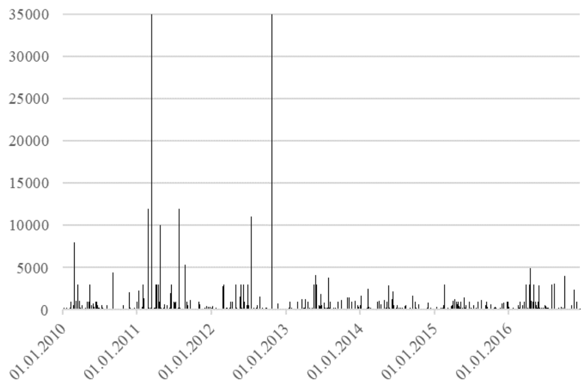


Fig. 4. Chronologically arranged the insured losses of natural catastrophes in USD million
Source: Own processing by Sigma Swiss Re, [2]-[5]

Table 2 shows the summary statistics of insured losses caused by natural catastrophes using our real data. In this table we can see that for example average, which is equal to 827.02, is higher than median which is equal to 300. The value of skewness is bigger than 10 and for example the value of kurtoses is really high – its value is equal 130.45.

Table 2. Summary statistics for insured losses of natural catastrophes

Count	479
Average	827.02
Median	300
Standard deviation	2 577.56
Coefficient of variation	311.67
Skewness	10.46
Kurtosis	130.45
Upper quartile	100
Lower quartile	649

Source: Own calculations

Box plot (Fig. 5) shows that there are many small losses and a few very large values of losses. The conclusion is that we need to find some long tail distribution that provides a suitable model for the variation amongst the catastrophe losses data.

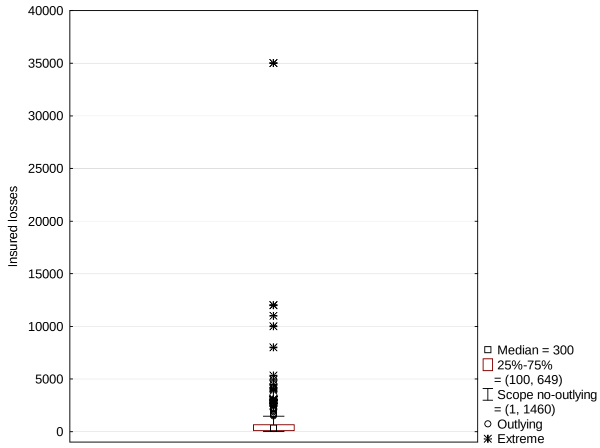


Fig. 5. Box plot of using data
Source: Own calculations

3.2 Block maxima model - results

The catastrophe insured losses data presented by Fig. 4 have been divided into *n* blocks (Table 3) of essentially equal size *N*. For this part of modelling have been used spreadsheet MS Excel.

Table 3. Number of blocks and values in the blocks

Number of blocks (*n*)	**Number of values in the block (*N*)**
5	95
10	47
15	31
20	23
25	19
30	15

Source: Own calculations

For modelling of these blocks of data has been used generalized extreme value distribution (GEV) in statistical package Statistica 12. Estimated values of parameters of GEV distribution by formulas (1) and (2) for different blocks of data shows Table 4.

In Table 4 we can see *p*-values of Kolmogorov-Smirnov tests. We can see the highest *p*-value for blocks for 25 values: $p = 0.9097$. This model represents the best fit of our data with the GEV model.

Table 4. Results of block maxima modelling

n	**Parameter ξ**	**Parameter μ**	**Parameter σ**	***p*-value of Kolmogorov-Smirnov test**
5	0.6806	963.42	843.67	0.2474
10	0.6685	1614.34	1348.23	0.5305
15	0.7188	2016.51	1681.28	0.5504
20	0.7887	2603.81	1915.22	0.4894
25	0.7121	3168.59	2432.64	0.9097
30	0.8909	3037.96	2405.69	0.8257

Source: Own calculations

The Figure 6 shows GEV distribution together with empirical distribution function with 95% confidence interval. We can see good fit of these distributions models of block maxima dataset.

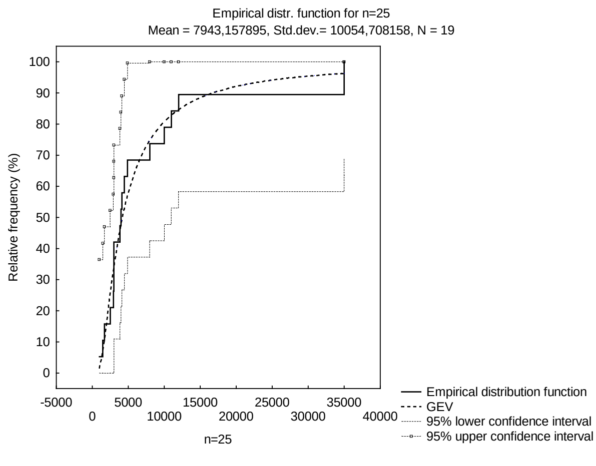


Fig. 6. GEV distribution fitted to block maxima for $n = 25$
Source: Own processing

To verify the quality of the GEV model we have used graphical analysis including Q-Q plot. Q-Q plot compares quantiles of theoretical and observed probability distribution. Fig. 7 shows Q-Q plot against the GEV distribution fitted to block maxima for n = 25. We can see the good fit of theoretical GEV model and the real data.

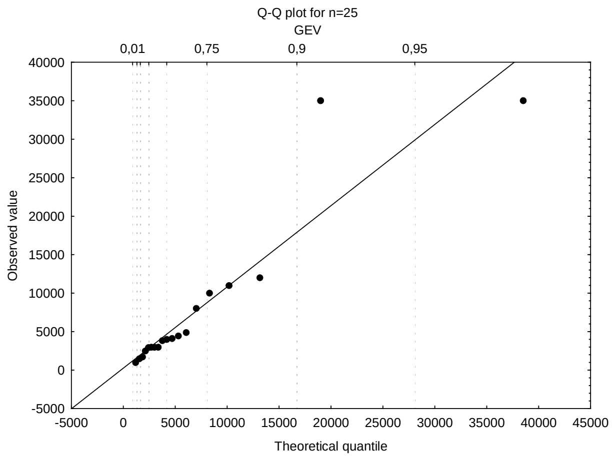


Fig. 7. Q-Q plot against the GEV distribution fitted to block maxima for n = 25

Source: Own processing

For the best estimated GEV model for n = 25 we can calculate selected quantiles. Some of these quantiles shows Table 5. By these value we can estimate, that for example 50% insurance extreme losses of natural catastrophes exceed value 4 632.03 million USD and 1% exceed value 39 104 million USD.

Table 5. Quantiles of fitted GEV of block maxima for n = 25

Quantiles	**GEV**
0.50	4 632.03
0.75	9 586.18
0.90	17 118.30
0.95	23 344.90
0.99	39 104.00

Source: Own processing

3.3 Peaks over threshold model - results

As the second method for modelling of insured losses of natural catastrophes from the period 2010-2016 according to the sub-chapter 2.1.2 have been used Peaks over threshold method. As a fitting distribution have been used a generalized Pareto distribution with maximum likelihood method of parameters estimation on the data above thresholds of u = 1 600, u = 2 600, u = 3 600 and u = 4 600. Table 6 contains the maximum likelihood estimated parameters of fitted generalized Pareto distributions on the data above four different thresholds.

Table 6. Parameters of GPD for different thresholds

Threshold *u*	**1600**	**2600**	**3600**	**4600**
Parameter ξ	1908.57	782.01	3741.66	8402.54
Parameter σ	-0.4546	-0.9479	-0.6534	-0.1849

Source: Own calculations

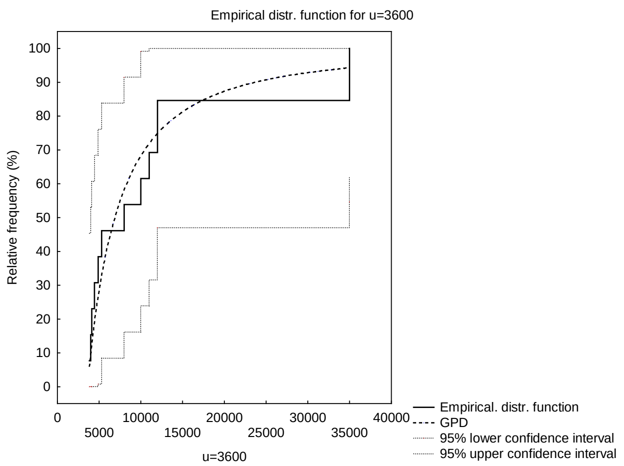
Table 7 contains the number of values exceed appropriate threshold and the results of Kolmogorov-Smirnov goodness of fit test for GPD for four different thresholds. Using p-values in Table 7 we can indicate the best fit model in the case of threshold u = 3 600 by because p = 0.9138 is the biggest one.

Table 7. Results of Kolmogorov-Smirnov goodness of fit test for GPD for different thresholds

Threshold *u*	**1600**	**2600**	**3600**	**4600**
n	46	37	13	9
***p*-value**	0.0138	0.0045	0.9138	0.6985

Source: Own calculations

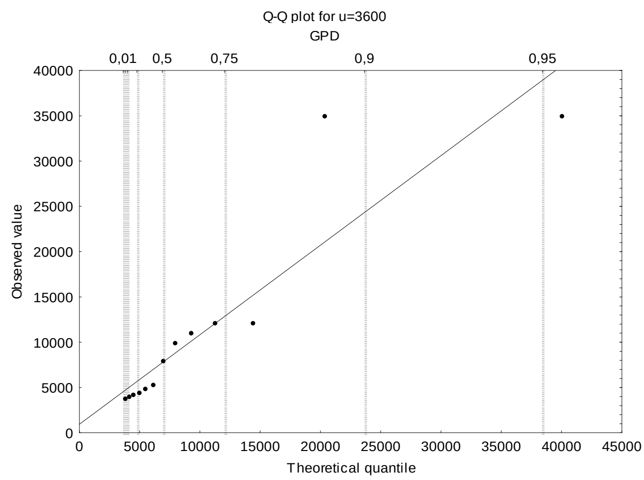
Fig. 8 shows the good fit of GEV distribution and empirical distribution function on 13 losses above the threshold u = 3 600. This visual form confirm good fit of these distribution and so relevance of the model.



Source: Own processing

Fig. 8. GPD distribution fitted by peaks over threshold u = 3 600

The Q-Q plot on Fig.9 also confirm good fit of estimated GPD distribution on the losses over threshold u = 3 600.



Source: Own processing

Fig. 9. QQ-plot of goodness of fit of the GPD on the losses over threshold u = 3 600

For the best fitted GPD model of the losses above threshold u = 3 600 we can calculate quantiles. Some of these quantiles are shown in Table 8. By these values we can estimate, that for example 50% insured extreme losses of natural catastrophes exceed 6 239.5 million USD and 1% exceed value 39 055 million USD.

Table 8. Quantiles of fitted GPD by Peaks over threshold method for u = 3 600

Quantiles	**GPD**
0.50	6 239.5
0.75	10 814.4
0.90	22 374.0
0.95	38 778.7
0.99	39 055.0

Source: Own processing

4 Conclusion

Catastrophic events have a huge impact on society as a whole. We can observe a growing trend in both the number of catastrophic events as well as in total and insured losses. Insurance and reinsurance undertakings must be prepared to pay for insured losses as a result of catastrophic events. A number of methods are used to help estimate and refine future claims cover. The block maxima method and peaks over threshold method are two of these methods.

This article presents application of both methods on real data of insured losses caused by natural catastrophes during time period 2010-2016. Have been found probability models with good fit on these data, by block maxima methods the generalized extreme value distribution - GEV model and by peaks over threshold method the generalized Pareto distribution - GPD model with parameters estimated by maximum likelihood method. This models have been used to calculate the selected quantiles.

By the values of quantiles of GEV model the insurance and reinsurance companies can expect 50% of insured catastrophe losses above 4 632.03 million USD, 10% of the insured losses above 17 118.3 million USD and 1% above 39 104 million USD. In case of the GPD model 50% of insured catastrophe losses exceed 6 239.5 million USD, 10% exceed 22 374.0 million USD and 1% exceed 39 055.0 million USD. This information is useful for actuaries and risk managers of insurance and reinsurance companies.

References:

[1] *The international Disaster Database*, 2018. [Online]. Available: https://www.emdat.be/classification

[2] *Sigma* No 1/2018: Natural catastrophes and man-made disasters in 2017: a year of record-breaking losses. Swiss Re [Online]. Available: http://institute.swissre.com/research/overview/sigma/1_2018.html

[3] *Sigma* 2/2017: Natural catastrophes and man-made disasters in 2016: a year of widespread damages. Swiss Re [Online]. Available: http://institute.swissre.com/research/overview/sigma/2_2017.html

[4] *Sigma* No 1/2016: Natural catastrophes and man-made disasters in 2015. Swiss Re [Online]. Available: http://institute.swissre.com/research/overview/sigma/1_2016.html

[5] *Sigma* No 2/2015: Natural catastrophes and man-made disasters in 2014, Swiss Re [Online]. Available: http://media.swissre.com/documents/sigma2_2015_en_final.pdf

[6] V. Pacáková, J. Gogola, Pareto Distribution in Insurance and Reinsurance. *Conference proceedings from 9th international scientific conference Financial Management of Firms and Financial Institutions*, VŠB Ostrava, 2013, pp. 298-306.

[7] P. Jindrová, Ľ. Sipková, Statistical Tools for Modeling Claim Severity. *Conference proceeding from 11th International Scientific Conference on European Financial Systems 2014*. Lednice, Masaryk University Brno, Czech Republic, 2014, pp. 288-294.

[8] W.G. Gilchrist, *Statistical modelling with quantile functions*. London: Chapman & Hall, 2000.

[9] H. A. David, H. N. Nagaraja, *Order Statistics*, 3rd ed. USA: John Wiley & Sons, 2003.

[10] A. J. McNeil, *Estimating the Tails of Loss Severity Distributions using Extreme Value Theory*. ETH Zentrum, Zürich (1996) [Online]. Available on: https://www.casact.org/library/astin/vol27no1/117.pdf

[11] A. McNeil, Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory. *ASTIN Bulletin*, 27(1), 1997, pp. 117-138.

[12] P. Embrechs, C. Kluppelberg, T. Mikosch, *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin, 1997.

[13] A. Brdar Turk, A Quantitative Operational Risk Management Model, *Transactions on Business and Economics*, Issue 5, Volume 6, 2009, pp. 241-253.

[14] P. Jindrová, V. Pacáková, Probability Models of Natural Catastrophe Losses. In *Proceedings of the 4th International Conference on Mathematical, Computational and Statistical Sciences* (MCSS'16), Barcelona, 2016, pp.

[15] P. Jindrová, V. Pacáková, Modelling of extreme losses in natural disasters. *International Journal of mathematical models and methods in applies sciences*, Volume 10, 2016, pp. 171-178.

[16] P. Jindrová, V. Pacáková, I. Laštovičková, Modelling of Insured Losses of Natural Catastrophes Using Block Maxima Model, In *Proceeding of the 3rd International Conference Applied Mathematics, Computational Science and Systems Engineering* (AMCSE 2018), Roma, 2018, submitted for publication.

[17] P. Jindrová, M. Papoušková, Modelling insured catastrophe losses. *Proceedings of the 10th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena*, Zakopane, 2016, pp. 53-61.

[18] P. Jindrová, R. Jakubínský, Significance and possibilities of major accident insurance. *E & M EKONOMIE A MANAGEMENT*, 18, Issue 4, 2015, pp. 121-131.

[19] V. Pacáková, L. Kubec, Modelling of catastrophic losses. *Scientific Papers of the University of Pardubice*, Series D, Vol. XIX, No. 25, 2012, pp. 125-134.

[20] V. Pacáková, P. Jindrová, T. Musil, Quantification of risks of natural catastrophes. *8th International Scientific Conference on Managing and Modelling of Financial Risks*, Ostrava, 2016, pp. 735-742.

[21] A. Davison, R. Smith, Models for exceedances over high thresholds, Journal of the Royal Statistical Society, Series B, No. 52, 1990.

[22] V. Skřivánková, A. Tartaľová, Catastrophic Risk Management in Non-life Insurance. In *E&M Economics and Management*, No 2, 2008, pp. 65-72.

[23] H. Zhongxian, Actuarial modelling of extremal events using transformed generalized extreme value distributions and generalized Pareto distributions, Doctoral thesis, The Ohio State University (2003) [Online]. Available:

http://www.math.ohio-state.edu/history/phds/abstracts/pdf/Han.Zhongxian.pdf.

[24] V. Pacáková, D. Brebera, Loss Distributions and Simulations in General Insurance and Reinsurance. *International Journal of Mathematics and Computers in Simulation*. NAUN, Volume 9, 2015, pp. 159-168.

[25] R. Fisher, L. Tippett, Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proceedings of the Cambridge Philosophical Society*, 24, 1928, pp. 180-190.

[26] A. Balkema, L. de Haan, Residual life time at great age, *Annals of Probability*, 2, 1974.

[27] J. Pickands, Statistical inference using extreme order statistics, *The Annals of Statistics*, 3, 1975, pp. 119-131.