ACTA MONTANA 1992
Series A, No. 2(88), 121-132

TRUNCATED PARETO DISTRIBUTION USED IN THE STATISTICAL ANALYSIS OF ROCKBURST HAZARD

[a] Stanisław Lasocki

[a] *Institute of Geophysics, University of Mining and Metallurgy, al. Mickiewicza 30, 30-059 Cracow, Poland*

Abstract: Non-stationarity of sequences of seismic events directly connected with mining works allows to construct a general time-dependent model of rockburst generation process. The model enables real time estimation of the probability of occurrence of a strong tremor from event series recorded during a certain period preceding the moment of estimate. To reach the goal a specific form of the distribution of event energy must be set.

The estimation procedure has been supplied with a weighing ability to account for the continuity of process dynamics and a truncated Pareto distribution has been tried to represent the distribution of event energy. The latter involves an assumption that in a given place and time there exists maximal energy which may be released in the form of a seismic event. A complete algorithm to estimate parameters of the truncated Pareto distribution and to evaluate time-dependent strong event hazard has been given in the paper. An example of the use of the method to process event sequence recorded during mining works has been presented.

Key words: mining seismicity; rockburst hazard; real time estimation; time-dependent tremor generation; strong tremor prediction

1. INTRODUCTION

The non-stationary character of a sequence of tremors directly connected with mining operations has been noticed by the author in all available data sets. The suggestion was proved by means of statistical inference methods. The results of analysis and their discussion are given elsewhere (Lasocki 1992). In consequence one may attempt to build methods for short term rockburst prediction in which seismicity data will play an important part. Following the studies of stochastic structure of seismicity occurring in a close vicinity of mining stopes (Lasocki 1990, 1992, 1992b) the below assumption can be accepted when constructing a time-dependent model of rockburst generation process:

- the events are mutually independent. The process is memoryless;
- the rate of event occurrence and the event energy are mutually independent;
- the process is non-stationary , but it is time-dependent only through a dependence of the parameters of its distribution;
- the process is segmental stationary, i.e. it is stationary for the time interval $[t,t+\Delta T]$, where t - any time moment and ΔT takes a specific value for local conditions. ΔT is supposed to be long enough to record a sequence of seismic events sufficiently numerous for a reliable estimation of process parameters;
- in the interval of stationarity the rate of tremor occurrence is distributed according to Poisson's distribution.

Under these assumptions the cumulative distribution function of tremor energy E for any time interval $(t,t+\Delta t)$, conditional upon event occurrence is given by:

$$F(E|N>0,\Delta t;t) = \{1-\exp[-\lambda(t)\cdot\Delta t]\}\cdot F_e(E;t) \quad (1)$$

where: $\lambda(t)$ is the mean number of events per unit time in the time interval $(t,t+\Delta t)$;
$F_e(E;t)$ is the cumulative distribution function of marginal distribution of tremor energy;
$\Delta t<\Delta T$.

The probability of occurrence of the event having energy greater than or equal to a given value E_p within the time interval $(t,t+\Delta t)$ is given by:

$$R(\Delta t,E_p;t) = \{1-\exp[-\lambda(t)\cdot\Delta t]\}\cdot\{1-F_e(E_p;t)\} \tag{2}$$

Since it is assumed that within the time interval of length ΔT, regardless its position on the time axis, the process is stationary, events recorded within this interval can be treated as an elementary statistical sample. Therefore if n events were recorded in the period, say, $(t-\Delta T,t)$ and their energies were respectively $E_1,\ldots,E_n$, then values for the constant for this period model parameters can be obtained by means of any well known statistical estimation procedure. For regularity purposes useful when evaluating errors of estimates, the maximum likelihood method is preferred whenever possible. The maximum likelihood estimator for λ is given by:

$$\hat{\lambda} = n/\Delta T \tag{3}$$

The other parameters are obtained solving the respective standard equations after the specific form of the energy distribution $F_e(E;t)$ is set.

The estimates are valid for the whole period $(t-\Delta T,t)$ thus, in particular, their values may be assigned to the upper limit t. Hence if Δt small compared to ΔT, the value achieved from formula (2), when substituting parameters with estimates, represents the probability of occurrence of the event having energy greater than or equal to E_p estimated for time moment t and predicted for time interval $(t,t+\Delta t)$.

The details of the model construction and estimation procedures have been presented elsewhere (Lasocki, 1990, 1992a).

Two problems of the above method of strong event hazard analysis have been considered in the presented work. The consequences of continuity of process dynamics have been discussed and in order to account for it the estimation procedure has been supplied with a weighing facility. Secondly the truncated Pareto distribution has been tried to represent in the model the event energy distribution.

2. WEIGHTS OF OBSERVATIONS - CONTINUITY OF PROCESS DYNAMICS PROBLEM

The time variation of model parameters results from the time changes of local geological and mining conditions controlling seismic energy discharging process in the mining rockmass. Since mining works in active stopes are carried on in more or less continuous manner, the local stress field is redistributed continuously and then the process dynamics is supposed to possess continuity property. It is assumed, however, that there exists a finite interval ΔT in which the process parameters remain unchanged and the observation are considered to build an elementary statistical sample. This assumption has been introduced because the true function representing process dynamics is unknown and means that it may be fruitfully, from the point of view of strong event hazard analysis, approximated by the staircase function of lag ΔT. Within the interval $(t-\Delta T, t)$ the observations recorded close to the upper limit contain more up-to-date information about the state of tremor generation process than the data about events occurring closer to the lower limit. The estimation procedure does not distinguish between the data hence a certain delay of the estimated probability with respect to the true probability may be expected due to incorporating older (i.e corresponding to the past states of the process) data into the estimation procedure.

The length of interval of stationarity ΔT is an external parameter in the model and is to be set a'priori with accordance to the event rate and the local variability of conditions controlling tremor generating process. In any case it must be large enough to ensure sufficient size of data set used in estimation procedure, otherwise the estimates would not be reliable. There are two ways to select the value of ΔT. It may represent absolute time that is it is a'priori chosen in time units and remains unchanged throughout the whole period of mining works in a stope which is the target of strong event hazard analysis. It does not exclude other trials with other ΔT values but one sequence of the probability estimates for all time moments is obtained with the use of constant value ΔT. The other method is probably more appropriate for stopes where mining operations cause movement of active front of any kind, e.g. longwall mining. In such cases the local stress variations might be related to the rate of face advance and to stabilize process variations corresponding to irregularities of the advance rate the ΔT should vary so that the face advance remains constant all the time during ΔT time periods.

Regardless the way in which the ΔT value is chosen the above mentioned delay of probability estimates may be significant in case the event rate forces long ΔT selection. The problem can be clearly seen in the constant advance method when mining has been finally stopped. Due to the way of ΔT selection the strong event probability estimate would remain constant while the real probability will be falling down with time due to rheological processes. Therefore it would be reasonable to attribute certain weights to the data used in estimation procedure so that the newest observations from the $(t-\Delta T,t)$ interval have the greatest importance on the resultant estimates at moment t while the oldest - the least.

As the first order approximation let us assume that during ΔT period information 'ages' exponentially. Then the weight of an event recorded at moment t_i from the point of view of moment t $(t \geq t_i\ ;\ t-t_i \leq \Delta T)$ is given by:

$$w_i = \beta \cdot \exp\{-\alpha(t-t_i)\} \tag{3}$$

where: α determines rapidity of 'ageing';
β is a normalization constant.
Let the number of events recorded during $(t-\Delta T,t)$ time period be n. We assume that this number is invariant of the weighing operation hence:

$$\sum_{t-\Delta T}^{t} w_i = n$$

then:

$$\beta = \frac{n}{\sum \exp\{-\alpha(t-t_i)\}} \tag{4}$$

Let the probability density function of tremor energy be $f_e(E;\xi_1,..,\xi_r)$, where $\xi_1,..,\xi_r$ are the distribution parameters. For the observed sequence $E_1,..,E_n$ the logarithmic maximum likelihood function is:

$$l(\xi_1,..,\xi_r) = \sum_{i=1}^{n} \ln[f(E_i)] = \sum_{i=1}^{n} \ln(p_i) \tag{5}$$

where p_i is the probability a'posteriori. If the data set can be split into say K classes each containing n_k observations of equal value E_k, $\sum_{k=1}^{K} n_k = n$, then formula (5) can be converted into:

$$l(\xi_1,..,\xi_r) = \sum_{k=1}^{K} \ln(p_k)^{n_k} \qquad (6)$$

Relation (6) represents relation (5) setting K=n and n_k=1 for all k.

When all observed events have equal weights (=1) each of them is treated as one observation. If the event i has the weight w_i we treat it formally as w_i observation. Then following relations (5), (6) the logarithmic maximum likelihood function of n weighed events $E_1,..,E_n$ is given by:

$$l(\xi_1,..,\xi_r) = \sum_{i=1}^{n} \ln(p_i)^{w_i} = \sum_{i=1}^{n} w_i \ln[f(E_i)] \qquad (7)$$

The respective standard equations are obtained as partial derivatives of l with respect to the parameters $\{\xi\}$.

3. TRUNCATED PARETO DISTRIBUTION AS THE ENERGY DISTRIBUTION MODEL

According to relations (1), (2) the method requires a specific form of energy distribution. In previous applications of the method the Pareto distribution has been used for this purpose. (Lasocki 1990, 1992a, Głowacka and Lasocki 1992). The Pareto distribution of energy is achieved when one assumes that above a certain energy threshold energies of events occurring in an area of mining works follow statistically the simple Gutenberg-Richter relation as follows:

$$\log N = a + b\cdot\log E \qquad E \geq E_0 \qquad (8)$$

where N denotes number of events of energies greater than or equal to E, and a, b - parameters.

The resultant density probability function of event energy is given by:

$$f_e(E) = b\cdot E_0^b / E^{b+1} \qquad E \geq E_0 \tag{9}$$

The model is identical with the seismological simple large earthquake model (Lomnitz 1974), with the only difference that in the latter magnitude is used instead of energy, and hence exhibits analogous properties. In particular it also leads to the so called energy catastrophe, which consists in the fact that the anticipated value of energy given by:

$$E[E] = \begin{cases} 1+E_0/(b-1) & b>1 \\ \longrightarrow \infty & b\leq 1 \end{cases} \tag{10}$$

is divergent for certain range of parameter values. The simple large earthquake model was also criticized due to the noticed deficit of events in large magnitude range with respect to the number predicted by the model. A similar phenomenon in mining induced seismicity was reported by several authors. (Kushnir et al. 1980, Bath 1984, Gibowicz 1990 and others) This feature of energy distribution has been sometimes noticed in sequences of events directly connected with mining operation but the evidence collected till now is not definitely convincing.

Both problems, the energy catastrophe and the discrepancy between theoretical and empirical energy distributions disappear when one assumes an existence of a certain upper magnitude/energy limit. (Cosentino et all. 1974) In case of the seismicity directly connected with mining this would mean that in the region of mining works there exists a maximum of energy which can be discharged in a form of seismic event. Since the tremor generation process is time-variant, this maximum is also supposed to change in time while mining. With an assumption like this the distribution of tremor energy takes a form of the truncated Pareto distribution. (Lasocki 1990) Its probability functions and the maximum likelihood standard equation for the shape parameter b with regard to the weighing of observation presented above are given by:

$$f_e^T(E) = [1-(E_0/E_m)^b]^{-1}\cdot bE_0{}^b/E^{b+1} \tag{11a}$$

$$F_e^T(E) = [1-(E_0/E_m)^b]^{-1}\cdot[1-(E_0/E)^b] \tag{11b}$$

$$n/b + \frac{n\cdot(E_0/E_m)^b}{1-(E_0/E_m)^b}\cdot \ln(E_0/E_m) - \sum_{i=1}^{n} w_i\cdot\ln(E_i/E_0) = 0 \tag{11c}$$

where: $E_0 \le E \le E_m$ and E_m is the upper limit of energy.
The likelihood function decreases monotonically with the increase of parameter E_m, therefore it cannot be estimated by means of maximum likelihood method. Because of the small size of samples used for estimation of parameters in the time-dependent strong event hazard analysis the estimation of this parameter cannot be made as it is carried on in similar seismological models. In fact the estimation of E_m is very troublesome and due to that the applications of this model in the real-time estimation of strong event occurrence is strongly limited. Kijko (1981) suggested a compromising way of E_m estimation which has been applied in the presented algorithm. An estimate of maximum of common logarithm of energy is given by:

$$(\log E_m)^{est} = 2\cdot\log E_{max}^{obs} - \log E_{max-1}^{obs} \tag{12}$$

where: E_{max}^{obs} is the largest and E_{max-1}^{obs} is the second largest energy in the sample. To obtain a reasonable estimate a considerable sample size is required which in turn usually forces long ΔT selection. Thus the method with the truncated Pareto distribution for tremor energy can be used only together with weighing procedure.

Fig. 1 presents a fragment of the day by day estimated probability record obtained by means of presented algorithm with the truncated Pareto distribution model for tremor energy. The data processed were acquired around an active longwall in the Pokoj coal mine in Upper Silesia, Poland. Mining works were carried on at the depth of 850-900 meters. The coal seam of average 3 meters thickness was overlaid by a complex of alternating rigid sandstone, shale and coal layers. The front was moving along a 40 meter wide coal pillar which was left to separate a ventilating cycle of the mine from the adjacent mine. This pillar was suspected to be the cause of the majority of strong seismic events occurring in the region of longwall. The recorded energies belonged to the range 10^2- 10^7J and those of order 10^3J and more satisfied the energy distribution model. The event rate was variable. More than 1300 events in all were recorded during two years mining of the slope. Due to disturbances in mining process caused by strong tremors, the face advance rate was irregular, in average 0.77 meter a day.

The analysis was done with the constant advance method. The period of stationarity ΔT was changing so that the face advance during the period was kept 30 meters at any moment of inference. The applied weighing caused that an influence of observation was dropping by a half every ten days. The procedure was used to estimate the probability of occurrence of event having energy greater than or equal to $5 \cdot 10^5$J.

Compared to the results obtained while using previous estimating algorithm with the normal Pareto distribution for tremor energy and without weighing function the above presented algorithm delivers records in which the probability build-ups prior to the strong events are more steep and last shorter time. Also the decreases of estimate are faster. These effects are probably due to weighing operation. An evident improvement in predictive efficiency of the algorithm due to the change of the model for energy distribution has not been observed. However, too small number of cases has been processed till now to conclude about the effectiveness of the new algorithm.

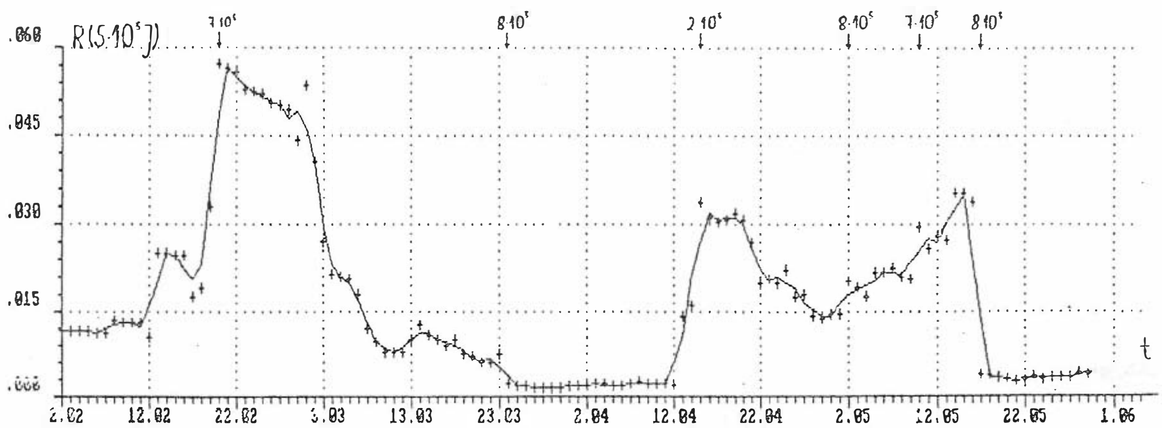

Fig. 1. The estimated probability of tremor occurrence of energy ≥ $5 \cdot 10^5$J in the longwall region.

3. CONCLUSIONS

The worked out method for weighing observations may be applied in any algorithm of real-time analysis of seismic event sequence in which information about the present and future states is inferred from the past data. In particular it is useful for the algorithm to estimate the time-varying probability of occurrence of strong tremor close to mining works. The method partially diminishes effects of the contradiction between the idea of variability in time of tremor generating process that is the demand of short duration of the stationarity period and the requirement of statistically significant size of sample used in estimation. It is particularly important when the analysis is carried on for low event rate sequences. The exponential form of weights should be understood as the first approximation only and should be reviewed by more detailed studies.

There has been developed the algorithm to estimate the probability of occurrence of strong tremor in the region of mining works in which the truncated Pareto distribution is used to model the event energy distribution. Although such a model eliminates difficulties which show up while interpreting the previously used simple Pareto distribution (the expected value of energy is always finite in this model) and also may give a better fit to the empirical energy distributions, due to the imperfect way of estimating the upper limit of tremor energy, the effectiveness of this algorithm is not visibly better than the effectiveness of the algorithm used previously. This conclusion is not, however, definite and further studies of field examples are necessary.

REFERENCES

Bath, M. (1984). Rockburst seismology, in *Proc. First International Congress on Rockburst and Seismicity in Mines, Johannesburg, 1982*, SAIMM, Johannesburg, 7-16

Cosentino, P., V. Ficarra and D. Luzio (1974). Truncated exponential frequency-magnitude relationship in earthquake statistics, *Bull. Seism. Soc. Am.* 67, 1615-1623

Gibowicz, S.J. (1990). The mechanism of seismic events induced by mining, in *Proc. 2nd International Symposium on Rockbursts and Seismicity in Mines, Minneapolis, June 1988*, A.A. Balkema, Rotterdam, 3-27

Głowacka, E. and S. Lasocki (1982). Probabilistic synthesis of the seismic hazard evaluation in mines, in *Proc. XX Czechoslovak-Polish Conference on Mining Geophysics, Sedlce-Prcic, May 1990, Acta Montana* (in print)

Kijko, A. (1991). Personal communications

Kushnir, N.J., D.P. Ashwin and A.G. Bradley (1980). Mining induced seismicity in the North Staffordshire Coalfield, England, *Int. J. Rock Mech. Min. Sci. and Geomech. Abstr.* 17, 45-55

Lasocki, S. (1990). Prediction of strong mining tremors, *Zesz. Nauk. AGH, s. Appl. Geophys.* 7, 110 pp. (in Polish with English abstract)

Lasocki, S. (1992). Non-Poissonian structure of sequence of mining tremors, in *Proc. XX Czechoslovak-Polish Conference on Mining Geophysics, Sedlce-Prcic, May 1990, Acta Montana* (in print)

Lasocki, S. (1992a). Time-dependent rockburst hazard analysis, in *Proc. Fifth Conference on Acoustic Emission/Microseismic Activity in Geologic Structures and Materials, The Pennsylvania State University July 1991, Trans Tech Publications* (in print)

Lasocki, S. (1992b). Weibull distribution for time intervals between mining tremors, in *Proc. XXI Polish-Czechoslovak Conference on Mining Geophysics, Lubiatow, May 1991, Publ. Inst. Geophys. Pol. Acad. Sci.* (in print)

Lomnitz, C. (1974). *Global Tectonics and Earthquake Risk,* Elsevier Scientific Publishing Co., Amsterdam, 320pp