

# Non-Parametric Characterisation of Mining Induced Seismic Sources

**B. Orlecka-Sikora, S. Lasocki** Department of Geophysics, Faculty of Geology, Geophysics and Environmental Protection, AGH University of Science and Technology, Poland

Statistical properties of seismicity are expressed by probabilistic distributions of seismic source parameters, out of which the magnitude distribution is the most important in the seismic hazard analysis. Owing to its complexity the magnitude distribution of mining-induced events cannot be accurately represented by any of the known parametric models. Hence, it was recently proposed that a kernel nonparametric estimator be used for this purpose. This paper describes a modified version of this estimator, which adapts to the uneven coverage of the magnitude range with sample data.

The probabilistic characterization is incomplete without confidence intervals of the probability function estimates. In order to estimate the confidence intervals for the cumulative distribution function of magnitude we have developed an algorithm based on the bootstrap and jackknife resampling methods. Monte Carlo simulations have shown that this technique provides satisfactory results for a reasonably wide range of sample sizes, regardless the actual shape of the magnitude distribution. We also provide an example of use of the nonparametric interval estimation method for characterizing the seismicity from an underground copper mine in the Legnica-Głogów Copper District in Poland.

## 1 INTRODUCTION

In recent years several authors have indicated breaks in scaling in the frequency magnitude distribution or, more generally, the complexity of the distribution for worldwide as well as regional earthquake data (e.g. Wesnousky, et al., 1983, Davison and Scholz, 1985, Pacheco and Sykes, 1992, Stirling, et al., 1996, Knopoff, 2000, Leonard et al., 2001). In mines, where rockmass rupturing process is controlled by various factors acting in different scales, the complexity and multimodality of the probability distribution of seismic event magnitude is more often and distinct. The observational evidences of the multicomponential structure of frequency-magnitude distribution in mining seismicity have been provided since the late seventies of the last century (e.g. Gibowicz and Lasocki, 2001 and the references therein). Using statistical inference methods Lasocki (2001) validated, with high probability, hypotheses on the complexity of magnitude distribution of mining events. Furthermore, he demonstrated that no parametric model of the distribution could fit such data so that the accuracy of representation required in the probabilistic seismic hazard analysis was ensured.

Since the presently known magnitude distribution models are not suitable for complex distributions Lasocki et al. (2000, 2002) and Kijko et al. (2001a,b) proposed the use of a kernel non-parametric density estimator (e.g. Silverman, 1986 and the references therein) as an alternative approach to magnitude distribution modelling. The estimator, modified for a finite range of event magnitude, takes the form of (e.g. Kijko et al., 2001a)

$$\hat{f}_m(m) = \frac{(\sqrt{2\pi})^{-1} \sum_{i=1}^n \frac{1}{h} \exp\left[-0.5\left(\frac{m-m_i}{h}\right)^2\right]}{\sum_{i=1}^n \left[ \Phi\left(\frac{m_{\max}-m_i}{h}\right) - \Phi\left(\frac{m_{\min}-m_i}{h}\right) \right]} \quad [1]$$

for  $m \in [m_{\min}, m_{\max}]$  and zero outside this range, where  $m_{\min}$  is the level of catalogue completeness,  $n$  is the number of events greater than or equal to  $m_{\min}$ ,  $m_i$  are the sizes of these

events,  $m_{\max}$  is the upper limit of source size,  $\Phi(\bullet)$  denotes the standard Gaussian cumulative distribution, and  $h$  is the smoothing factor. The smoothing factor is automatically selected upon the data by means of the least-squares cross-validation technique (e.g. Bowman et al., 1984), which in the case of estimator [1] means solving the equation,

$$\sum_{i,j} \left\{ 2^{-0.5} \left[ \frac{(m_i - m_j)^2}{2h^2} - 1 \right] \exp\left[-\frac{(m_i - m_j)^2}{4h^2}\right] - 2 \left[ \frac{(m_i - m_j)^2}{h^2} - 1 \right] \exp\left[-\frac{(m_i - m_j)^2}{2h^2}\right] \right\} - 2n = 0 \quad [2]$$

for the unknown  $h$  (Kijko et al., 2001a). The upper limit of magnitude range,  $m_{\max}$ , is estimated using the generic formula (Cooke, 1979, Kijko and Graham, 1998)

$$\hat{m}_{\max} = m_{\max}^{obs} + \int_{m_{\min}}^{m_{\max}} [F_m(m)]^n dm \quad [3]$$

where  $m_{\max}^{obs}$  is the largest observed magnitude and  $F_m(m)$  is the magnitude cumulative distribution. As can be seen, the estimator [1] is model-free and is completely determined by the data. Monte Carlo simulation experiments showed that down to sample sizes of about 100 this approach provides reasonable estimates of probability functions of magnitude, regardless the actual shape of the underlying distribution. In this paper we present a modified version of the kernel estimator of magnitude, which adapts to the considerably uneven coverage of the magnitude range with sample data.

The probabilistic characterization of a random variable is incomplete without confidence intervals of the probability function estimates. In the standard statistical approach the interval estimation is possible if either the distributions of probability function estimators are known or many samples are available to approximately reconstruct these distributions. Neither of the above occurs in the non-parametric approach to magnitude distribution estimation. We propose then to

solve this problem using resampling methods: the bootstrap and jackknife, since this is the only way to evaluate interval estimates of non-parametric magnitude distribution functions on the basis of one data set. In brief, the resampling techniques applied to a one-sample data provide many replicas of this sample, which differ from the original data but preserve its probabilistic properties (Hall, 1992, Efron and Tibshirani, 1998). These newly constructed samples, in particular sample percentiles, are used to construct confidence intervals of the magnitude probability density (PDF) or cumulative distribution (CDF) functions. The resampling approach was successfully tried on Monte Carlo generated data of given exponential and complex population distributions (Orlecka-Sikora, 2004a,b).

**2 THE ADAPTIVE KERNEL ESTIMATOR OF MAGNITUDE DENSITY**

Simply speaking, the non-parametric density estimate is constructed by summing up Gaussian densities of the same standard deviation,  $h$ , centred at each data point. The estimate is regular in densely populated ranges of the random variable under study but can exhibit spurious irregularities in those ranges where the data are sparse. The latter are low probability intervals and are often unimportant. In probabilistic seismic hazard studies, however, one is mainly interested in magnitude values of low exceedance probability, hence in proper modelling of the low probability, larger magnitude range. Moreover, the magnitude distribution is exponential-like, which means that there is a sharp decrease in the number of events when moving from smaller to greater magnitude values. In such a problem the classic, constant kernel estimates of magnitude distribution functions may not be accurate enough.

The problem of a locally non-optimal estimate of the far right-hand side wing of the magnitude distribution can be partially overcome by replacing in [1] the constant smoothing factor,  $h$ , with a variable one that grows when the data become sparse. Such an adaptive density estimator of magnitude is

$$\hat{f}_m^a(m) = \frac{(\sqrt{2\pi})^{-1} \sum_{i=1}^n \frac{1}{\omega_i h} \exp\left[-0.5\left(\frac{m-m_i}{\omega_i h}\right)^2\right]}{\sum_{i=1}^n \left[ \Phi\left(\frac{m_{\max}-m_i}{\omega_i h}\right) - \Phi\left(\frac{m_{\min}-m_i}{\omega_i h}\right) \right]} \quad [4]$$

where  $\omega_i, i=1, n$ , are the local bandwidth factors. The corresponding cumulative distribution estimator is

$$\hat{F}_m^a(m) = \frac{\sum_{i=1}^n \left[ \Phi\left(\frac{m-m_i}{\omega_i h}\right) - \Phi\left(\frac{m_{\min}-m_i}{\omega_i h}\right) \right]}{\sum_{i=1}^n \left[ \Phi\left(\frac{m_{\max}-m_i}{\omega_i h}\right) - \Phi\left(\frac{m_{\min}-m_i}{\omega_i h}\right) \right]} \quad [5]$$

for  $m \in [m_{\min}, m_{\max}]$ , zero for  $m < m_{\min}$  and one for  $m > m_{\max}$ . The local bandwidth factors,  $\omega_i$ , can be, for instance

$$\omega_i = \left[ \frac{\tilde{f}(m_i)}{g} \right]^{-0.5} \quad [6]$$

where  $\tilde{f}(\bullet)$  is the pilot, constant kernel estimator in the unbounded magnitude range,

$$\tilde{f}(m) = \frac{1}{\sqrt{2\pi}nh} \sum_{i=1}^n \exp\left[-\frac{1}{2}\left(\frac{m-m_i}{h}\right)^2\right] \quad [7]$$

and  $g = \left[ \prod_{i=1}^n \tilde{f}(m_i) \right]^{\frac{1}{n}}$  is the geometric mean of all constant kernel estimates (Silverman, 1986).

In order to study advantages of the adaptive approach we drew 50 samples, of 200 elements each, from the both sides truncated exponential distribution of magnitude, of the form

$$f_m(m) = \frac{\beta \exp[-\beta(m-m_{\min})]}{1 - \exp[-\beta(m_{\max}-m_{\min})]} \quad [8]$$

for  $m \in [m_{\min}, m_{\max}]$  and zero outside this range, which is just the probabilistic representation of the Gutenberg-Richters' relation with an upper bound for event magnitude.  $\beta = b \ln 10$ , where  $b$  is the Gutenberg-Richters' constant. The parameters for the simulations were:  $b=0.8$  ( $\beta=1.842$ ),  $m_{\min}=1.0$  and  $m_{\max}=5.2$ . Every generated sample was used to estimate the cumulative distribution,  $F_m(m)$  and the probability that the specified value of seismic event magnitude  $m_p$  will be exceeded in  $D$  time units

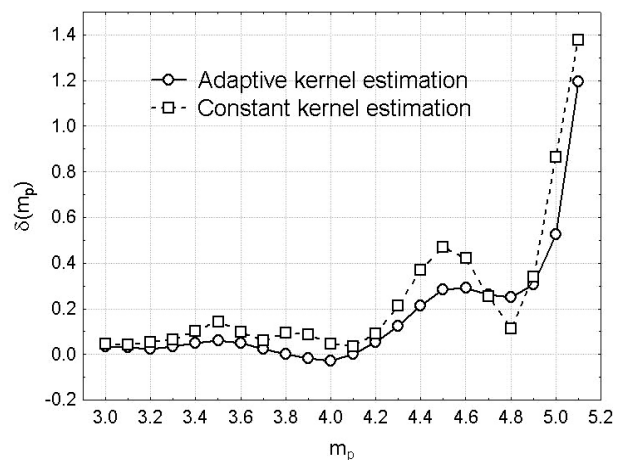
$$R(m_p, D) = 1 - \exp\{-\lambda D [1 - F_m(m_p)]\} \quad [9]$$

where  $\lambda$  is the mean event rate. The exceedance probability, [9], is a principal parameter of the probabilistic seismic hazard analysis. The estimation was performed by means of both the constant kernel and adaptive kernel estimators. We used  $\lambda D=25$  and  $m_p$  from 3 to 5.1 with the step 0.1.

Figure 1 shows the relative error of mean exceedance probability for both the constant kernel and adaptive kernel approaches. The error was evaluated by

$$\delta(m_p) = \frac{\bar{R}(m_p, D) - R^T(m_p, D)}{R^T(m_p, D)} \quad [10]$$

where  $\bar{R}$  was the mean of exceedance probability estimates from all samples, and  $R^T$  was the exact exceedance probability, calculated from [8]. It can be seen that in all but two cases the error of the constant kernel estimate is greater than the error of the adaptive kernel estimate. In the most important in practice, larger magnitude range the difference in favour of the adaptive kernel method can reach 10-30 percent.



**FIG. 1** Relative errors of mean exceedance probability estimates from the constant kernel and adaptive kernel non-parametric estimators of magnitude density

### 3 NONPARAMETRIC INTERVAL ESTIMATION OF THE CUMULATIVE DISTRIBUTION FUNCTION OF MAGNITUDE

#### 3.1 Bootstrap and Jackknife Methods

The bootstrap and jackknife resampling techniques replicate an original data sample in different ways. Let the original data set,  $x$ , consist of  $n$  points,  $(x_1, x_2, \dots, x_n)$ . A bootstrap sample,  $x^*=(x_1^*, x_2^*, \dots, x_n^*)$ , is obtained by random sampling  $n$  times, with replacement, from  $x$ . The  $i$ -th jackknife sample,  $x_{(i\text{jack})}$ , is defined to be  $x$  with the  $i$ -th data point removed from the original data set,  $x_{(i\text{jack})}=(x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ , where  $i=1, 2, \dots, n$  (Efron and Tibshirani, 1998). If  $\theta$  is an unknown parameter and  $\hat{\theta}=s(x_1, x_2, \dots, x_n)$  is its sample estimate, then on the basis of each bootstrap sample a bootstrap replication  $\hat{\theta}^*$  of parameter  $\theta$  is obtained by:

$$\hat{\theta}^* = s(x^*) \tag{11}$$

and  $i$ -th jackknife replication  $\hat{\theta}_{(i\text{jack})}$  of parameter  $\theta$  is defined as  $s(\bullet)$  evaluated for  $i$ -th jackknife sample,  $x_{(i\text{jack})}$ , say:

$$\hat{\theta}_{(i\text{jack})} = s(x_{(i\text{jack})}) \tag{12}$$

The bootstrap technique can provide as many data samples as demanded, but the jackknife method supplies only the number of jackknife sets equal to the number of points in the original data sample,  $n$ . Both these methods are used in the magnitude CDF interval estimation algorithm.

#### 3.2 The Algorithm for Constructing Magnitude CDF Confidence Intervals

The magnitude CDF confidence interval estimation algorithm is based on the so-called *bias corrected and accelerated method* ( $BC_a$  method, Efron, 1987). According to the  $BC_a$  method, for any magnitude value,  $m$ , the  $(1-2\cdot\alpha)$  probable confidence interval of the non-parametric magnitude CDF,  $F_m$ , is

$$\left( \hat{F}_{m,\alpha 1}^*, \hat{F}_{m,\alpha 2}^* \right) \tag{13}$$

where  $\hat{F}_{m,\alpha 1}^*$  and  $\hat{F}_{m,\alpha 2}^*$  are bootstrap estimated percentiles of the distribution of non-parametric CDF estimator,  $\hat{F}_m$ . The orders of percentiles,  $\alpha 1$  and  $\alpha 2$ , are calculated from the equations

$$\alpha 1 = \Phi \left( z_0 + \frac{z_0 + z_\alpha}{1 - \hat{a}(z_0 + z_\alpha)} \right) \tag{14}$$

$$\alpha 2 = \Phi \left( z_0 + \frac{z_0 + z_{1-\alpha}}{1 - \hat{a}(z_0 + z_{1-\alpha})} \right) \tag{15}$$

where  $\Phi(\bullet)$  is the standard Gaussian cumulative distribution,  $z_\alpha$  and  $z_{1-\alpha}$  are percentiles of the standard Gaussian distribution,  $z_0$  is the value of bias-correction and  $\hat{a}$  is the acceleration constant. In order to evaluate  $\hat{F}_{m,\alpha 1}^*$ ,  $\hat{F}_{m,\alpha 2}^*$  and  $\hat{z}_{0'}$ , a number (say  $k$ ) of bootstrap samples is drawn from the original magnitude data set. Each of these bootstrap samples is used to estimate  $F_m$  according to [5]. The estimates,  $\{\hat{F}_m^*\}$ , are sorted in the ascending order, leading to a series  $\hat{F}_{m,(1)}^*, \dots, \hat{F}_{m,(k)}^*$  such that  $\hat{F}_{m,(i)}^* < \hat{F}_{m,(i+1)}^*$  for every  $i$ . The percentile estimate,  $\hat{F}_{m,\alpha}^*$  is the  $k\cdot\alpha$ -th element of the series  $\hat{F}_{m,(1)}^*, \dots, \hat{F}_{m,(k)}^*$ . The bias correction estimate,  $\hat{z}_{0'}$ , is obtained from the proportion of bootstrap replications leading to CDF estimates,  $\hat{F}_{m,i}^*$ , smaller than the estimate  $\hat{F}_m$  obtained from the original data sample,

$$\hat{z}_0 = \Phi^{-1} \left( \frac{\text{number of } \hat{F}_{m,i}^* < \hat{F}_m}{k} \right) \tag{16}$$

where  $\Phi^{-1}(\bullet)$  indicates the inverse function of the standard Gaussian CDF. The acceleration constant,  $\hat{a}$ , can be evaluated in different ways, for instance from

$$\hat{a} = \frac{\sum_{i=1}^n (\hat{F}_{m,(\bullet)} - \hat{F}_{m,(i\text{jack})})^3}{6 \left\{ \sum_{i=1}^n (\hat{F}_{m,(\bullet)} - \hat{F}_{m,(i\text{jack})})^2 \right\}^{3/2}} \tag{17}$$

where  $\hat{F}_{m,(i\text{jack})}$  denotes  $i$ -th jackknife non-parametric estimate of magnitude CDF, and  $\hat{F}_{m,(\bullet)}$  is the arithmetic mean of all jackknife estimates.

In order to improve the accuracy of results we have proposed to modify the presented method and to use an iterated bootstrap for estimating the bias-correction parameter,  $z_0$ . The iterated bootstrap is based on the principle of resampling from a bootstrap sample. It means that from every bootstrap sample a set of second order bootstrap samples is drawn and used to estimate  $z_0$ . In this way one obtains  $k$  estimates of  $z_0$ , where  $k$  is the number of the first order bootstrap samples. Finally, the mean value of these estimates,  $\hat{z}_{0'}$ , is calculated and used to evaluate  $\alpha 1$  and  $\alpha 2$ . We named this procedure *the iterated  $BC_a$  method*. Monte Carlo simulation studies indicated that the iterated  $BC_a$  method, applied for nonparametric interval estimation of magnitude CDF and related parameters of seismic hazard, can significantly improve the results of estimation. (Orlecka-Sikora, 2004a,b). A flow chart of the iterated  $BC_a$  algorithm is shown in Figure 2.

#### 3.3 Number of Bootstrap Replications

The number of bootstrap replications is another important aspect of the iterated  $BC_a$  method of nonparametric interval estimation. It is obvious that from the point of view of computer time that the number of bootstrap samples,  $k$ , should be kept as small as possible. The error of Monte Carlo bootstrap approximation to the distribution of an unknown parameter, is the sum of two independent errors: the statistical and the simulation error. The statistical error is unavoidable and is independent of  $k$ , but the simulation error depends on  $k$  (Wehrens et al., 2000). Thus, from this standpoint  $k$  should be chosen such that the simulation errors were smaller than the statistical error. The more observations in a dataset, the more accurate the bootstrap approximations. But at the same time, to make the simulation error accordingly small, a higher  $k$  should be taken. A rule of thumb says that the number of bootstrap samples  $k \geq 1000$  lead to acceptably small simulation errors (e.g. Efron and Tibshirani, 1998; Wehrens et al., 2000). In the nonparametric magnitude CDF interval estimation, Monte Carlo simulation experiments indicated that  $k = 1000$  provides good results (Orlecka-Sikora, 2004a,b).

### 4 INTERVAL ESTIMATION OF THE SEISMIC HAZARD PARAMETERS

The magnitude CDF confidence intervals can be used to find the confidence intervals of other probabilistic descriptors of seismic sources, viz. the exceedance probability of magnitude  $m_p$  in  $D$  time units, [9], and the mean return period of events of size  $m_p$ , defined by

$$T(m_p) = \left\{ \lambda [1 - F_m(m_p)] \right\}^{-1} \tag{18}$$

Both,  $R(m_p, D)$  and  $T(m_p)$  functions are monotonic with  $F_m(\bullet)$ , then the confidence intervals of magnitude CDF can be readily converted into the confidence intervals of  $R$  and  $T$ . The confidence interval of exceedance probability is

$$\left( \hat{R}(m, D)_{\alpha 1}^*; \hat{R}(m, D)_{\alpha 2}^* \right) \tag{19}$$

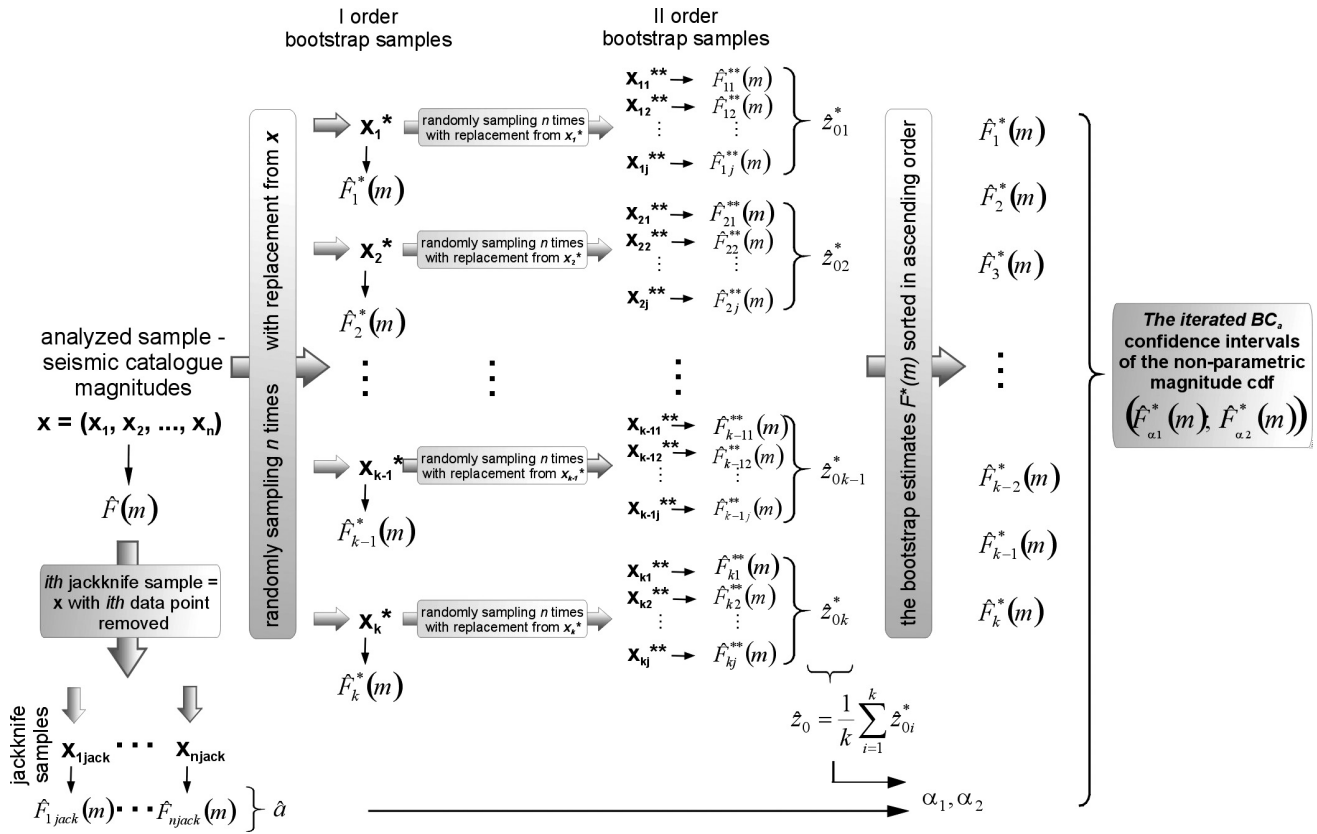


FIG. 2 Flowchart of the iterated BC<sub>a</sub> algorithm for the nonparametric interval estimation of magnitude cumulative distribution function

where

$$\hat{R}(m, D)_{\alpha_1}^* = 1 - \exp\{-\lambda D [1 - \hat{F}_m^*(m)_{\alpha_2}^*]\} \quad [20]$$

$$\hat{R}(m, D)_{\alpha_2}^* = 1 - \exp\{-\lambda D [1 - \hat{F}_m^*(m)_{\alpha_1}^*]\} \quad [21]$$

and the confidence interval of the mean return period is

$$(\hat{T}(m)_{\alpha_1}^*; \hat{T}(m)_{\alpha_2}^*) \quad [22]$$

where

$$\hat{T}(m)_{\alpha_1}^* = \{\lambda [1 - \hat{F}_m^*(m)_{\alpha_2}^*]\}^{-1} \quad [23]$$

$$\hat{T}(m)_{\alpha_2}^* = \{\lambda [1 - \hat{F}_m^*(m)_{\alpha_1}^*]\}^{-1} \quad [24]$$

## 5 A PRACTICAL EXAMPLE

Legnica-Głogów Copper District (LGOM) in Poland's Lower Silesia region is an area of intense seismic activity induced by underground copper ore exploitations in three mines: Lubin, Rudna and Polkowice-Sieroszowice. The seismicity of LGOM has a clustered structure. The seismic events form distinct time-space active zones, usually associated with active stopes (Gibowicz and Lasocki, 2000; Orlecka-Sikora and Lasocki, 2001). Studying mining seismic catalogues from the last 20 years, Lasocki et al. (2003) identified 99 such zones of seismicity. We present here an example of use of the presented techniques, viz. the non-parametric, adaptive magnitude<sup>1</sup> CDF estimation and the iterated BC<sub>a</sub> procedure for nonparametric confidence interval estimation of magnitude CDF, to assess probabilistic properties of the seismic source size in one of the identified zones, Z36b. The zone Z36b, located in the Lubin mine, comprised 276 events that occurred between October 1999 and August 2001. The

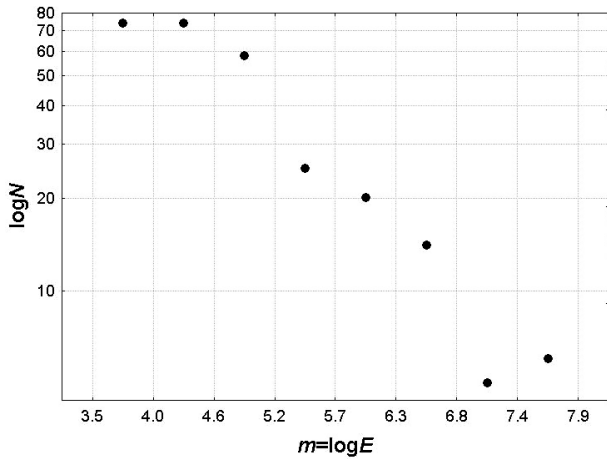
mean activity rate,  $\lambda$ , was 13.3 events/month. The empirical recurrence relation, shown in Figure 3, distinctly suggests a complex event size distribution.

To construct the CDF confidence intervals of seismic event size,  $m$ , for the zone Z36b a set of 1000 bootstrap samples was drawn from the 276-elements data sample. Each of the bootstrap samples was used to estimate the events CDF, according to equation [5]. Each of 1000 bootstrap samples was the source of a set of 200 second order bootstrap samples. These were used to estimate the bias-correction parameter,  $z_0$ . In this way 1000  $\hat{z}_0$  values were obtained and their mean value  $\hat{z}_0$  was used to calculate  $\alpha_1$  and  $\alpha_2$  values ([14], [15]). Finally the confidence intervals were calculated with a confidence level equal 95 per-cent. Figure 4 presents the result of the analysis.

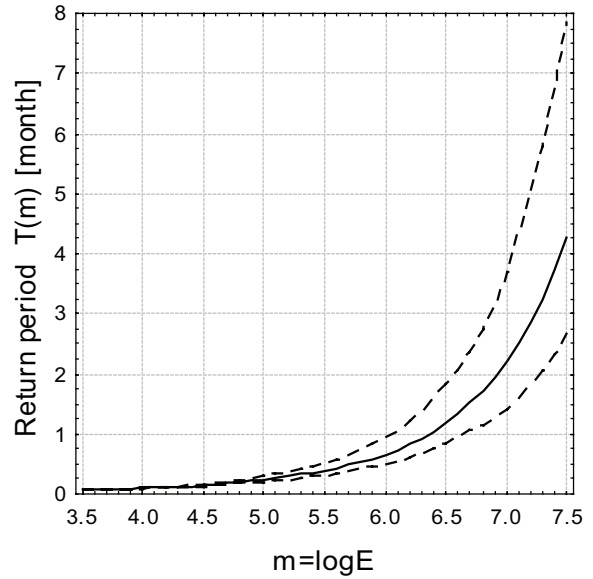
On the basis of CDF confidence intervals, the exceedance probability,  $R(m_p, D)$  ([9]), and mean return period,  $T(m_p)$  ([18]), confidence intervals were evaluated, according to equations [19] – [24]. The results are shown in Figures 5 and 6, respectively. The exceedance probability presented there was estimated for  $m_p = \log E = 7$ .

<sup>1</sup> In Polish mining catalogues the event size is parameterised by the seismic expressed in joules. As well known, statistically magnitude is linearly related to the logarithm of energy. Since our considerations in previous sections referred to magnitude, we represented, in this example, the event size with the common logarithm of energy,  $m = \log E$ . Although the parameter  $m$  is not any commonly used magnitude it has the same shape of probabilistic distribution as magnitude and, being such, can be studied with the techniques from Sections 2-4.

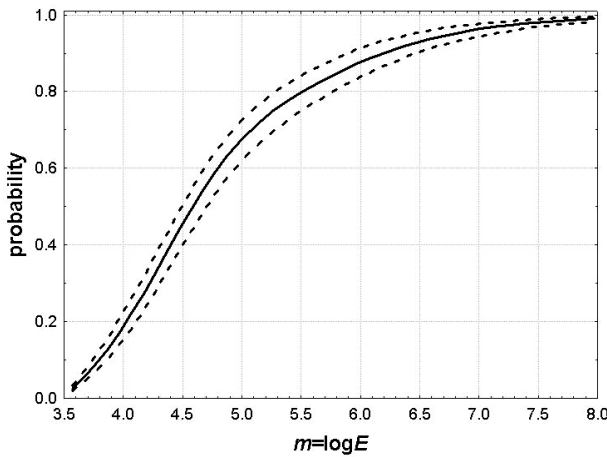




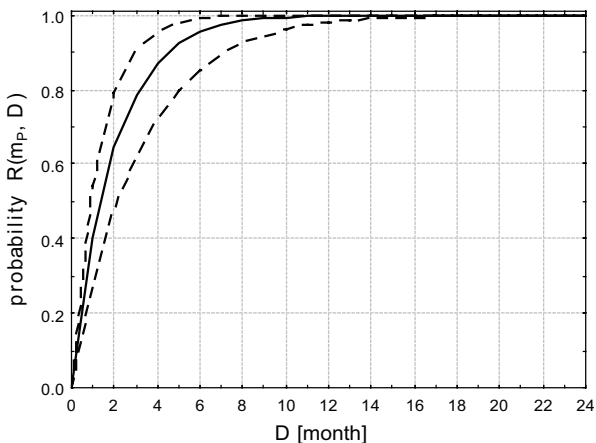
**FIG. 3** The empirical cumulative recurrence relation for the studied dataset of zone Z36b



**FIG. 6** Nonparametric estimates of the mean return period for the studied dataset from zone Z36b. The solid line represents the point estimate while the 95 per-cent confidence intervals are represented by dashed lines



**FIG. 4** Nonparametric estimates of  $m = \log E$  CDF for the studied dataset from zone Z36b. The solid line represents the point estimate while the 95 per-cent confidence intervals are represented by dashed lines



**FIG. 5** Nonparametric estimates of the exceedance probability of  $m = \log E = 7$  for the studied dataset from zone Z36b. The solid line represents the point estimate while the 95 per-cent confidence intervals are represented by dashed lines

**6 CONCLUSIONS**

The variety and time-variability of factors influencing seismicity generation in mines cause the probability distribution of mining-induced event magnitude to be both complex and multimodal. Since none of the presently known parametric models of distribution is accurate enough from the point of view of probabilistic seismic hazard analysis, the required accuracy is offered by the non-parametric, kernel estimation methods.

The specific, exponential-like form of magnitude distribution, in connection with the interest focused at larger magnitudes exceedance probabilities, make the adaptive kernel estimate more suitable than the constant kernel one in terms of representing the magnitude distribution.

The Monte Carlo simulation experiment showed that the use of the adaptive kernel estimator could reduce the relative error of mean exceedance probability by up to 30 percent.

In practical hazard studies the magnitude CDF estimation is carried out from only one data sample, i.e. a seismic catalogue. The differences between the actual CDF and its one-sample estimate result in hazard errors of unknown percentage. In order to improve the nonparametric estimation of magnitude, CDF one-sample estimate should be completed with an interval estimate. In the case of the adaptive kernel estimator it is only possible through the use of the bootstrap and jackknife resampling methods.

Monte Carlo simulation studies showed that the iterated bias corrected and accelerated method is an effective technique of nonparametric interval estimation of magnitude CDF. It provides satisfactory results for a reasonably wide range of sample sizes, regardless the actual shape of the magnitude distribution. The confidence intervals of CDF can be readily transformed into the confidence intervals of the exceedance probability of event size and mean return period. The presented practical example illustrated how this method, applied to real datasets, can provide useful information on dispersion of hazard parameter estimates.

## ACKNOWLEDGMENTS

This work was financed by the Polish State Committee for Scientific Research under contract No. 5 T12A 046 25 during the period 2003-2005.

## REFERENCES

- Bowman, A.W., Hall, P. and Titterton, D.M. (1984) Cross-validation in nonparametric estimation of probabilities and probability densities. *Biometrika* 71: pp. 341-351.
- Cooke, P. (1979) Statistical inference for bounds of random variables. *Biometrika* 66: pp. 367-374.
- Davison, F.C.Jr. and Scholz, C.H. (1985) Frequency-moment distribution of earthquakes in the Aleutian Arc: A test of the characteristic earthquake model. *Bull. Seismol. Soc. Am.* 75: pp. 1349-1361.
- Efron, B. (1987) Better bootstrap confidence intervals. *J. Am. Stat. Assoc.* 82: pp. 171-200.
- Efron, B. and Tibshirani, R.J. (1998) *An Introduction to the Bootstrap*. Chapman and Hall, New York.
- Gibowicz, S.J. and Lasocki, S. (2001) Seismicity induced by mining: Ten years later. *Advances in Geophysics* 44: pp. 39-181.
- Hall, P. (1992) *The Bootstrap and Edgeworth Expansion*. Springer – Verlag, New York.
- Kijko, A. and Graham, G. (1978) "Parametric-Historic" procedure for probabilistic seismic hazard analysis. Part I: Assessment of maximum regional magnitude  $m_{max}$ . *Pure Appl. Geophys.* 152: pp. 413-442.
- Kijko, A., Lasocki, S. and Graham, G. (2001a) Nonparametric seismic hazard analysis in mines. *Pure Appl. Geophys.* 158: pp. 1655-1676.
- Kijko, A., Lasocki, S., Graham, G. and Retief S.J.P. (2001b) Non-parametric seismic hazard analysis in mines. In 5th Int. Symp. Rockbursts and Seismicity in Mines "Dynamic rockmass response to mining", Magalisberg, 17-20 September 2001 (eds. G. van Aswegen, R.J. Durrheim, W.D. Ortlepp) SAIMM S27, Johannesburg, South Africa, pp. 493-500.
- Knopoff, L. (2000) The magnitude distribution of declustered earthquakes in Southern California. *Proc. Natl. Acad. Sci. USA* 95: pp. 11880-11884.
- Lasocki, S. (2001) Quantitative Evidences of Complexity of Magnitude Distribution in Mining-induced Seismicity: Implications for Hazard Evaluation. In 5th Int. Symp. Rockbursts and Seismicity in Mines "Dynamic rockmass response to mining", Magalisberg, 17-20 September 2001 (eds. G. van Aswegen, R.J. Durrheim, W.D. Ortlepp) SAIMM S27, Johannesburg, South Africa, pp. 543-550.
- Lasocki, S., Kijko, A. and Graham, G. (2000) Model-free Seismic Hazard Estimation. In *Proc. Int. Conf. Earthquake Hazard and Risk in the Mediterranean Region, EHRMR'99, Lefkosa, 18-22 November 1999* (ed. H. Gokcekus) Educational Foundation of Near East University, Lefkosa, T. R. N. Cyprus, pp. 503-508.
- Lasocki, S., Kijko, A. and Graham, G. (2002) Model-free Seismic Hazard Analysis. In *Seismogenic Process Monitoring*. (eds. H. Ogasawara, T. Yanagidani, and M. Ando). Balkema, Lisse, pp. 327-339.
- Lasocki, S., Kleczek, Z., Mirek, J., Olszewska D., Orlecka-Sikora, B. and Gibowicz, S.J. (2003) Prediction of Seismic Impact on „elazny Most” Repository With a Use of Acceleration Ground Motion Data. KGHM „Polska Mied ” S.A., Hydrotechnical Division – Final Project Report (unpublished).
- Leonard, T., Papanoulotis, O. and Main, I.G. (2001) A Poisson model for identifying characteristic size effects in frequency data: Application to frequency-size distributions for global earthquakes, "starquakes", and fault lengths. *J. Geophys. Res.* 106 (B7): pp. 13473-13484.
- Orlecka-Sikora, B. (2004a) Resampling methods in nonparametric seismic hazard estimation. *Acta Geophys. Pol.* 52: pp. 15-27.
- Orlecka-Sikora, B. (2004b) Bootstrap and jackknife resampling for improving in the nonparametric seismic hazard estimation. *The IUGG 2003 Proceedings Volume*.
- Orlecka-Sikora, B. and Lasocki, L. (2001) Clustered structure of seismicity from the Lubin-Glogow copper district. *Publ. Inst. Geophys. Pol. Acad. Sci. M-24 (340)*: pp. 105-119.
- Pacheco, J.F. and Sykes, L.R. (1992) Seismic moment catalog of large shallow earthquakes, 1900 to 1989. *Bull. Seismol. Soc. Am.* 82: pp. 1306-1349.
- Silverman, B.W. (1986) *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.
- Stirling, M.W., Wesnousky, S.G. and Shimazaki, K. (1996) Fault trace complexity, cumulative slip, and the shape of the magnitude-frequency distribution for strike-slip faults: A global survey. *Geophys. J. Int.* 124: pp. 833-868.
- Wehrens, R., Putter, H. and Buydens, L.M.C. (2000) The bootstrap: a tutorial. *Chemometrics and Intelligent Laboratory Systems* 54: pp. 35-52.