CrossMark

RESEARCH ARTICLE - SPECIAL ISSUE

Time-dependent seismic hazard in Bobrek coal mine, Poland, assuming different magnitude distribution estimations

Konstantinos Leptokaropoulos[1] · **Monika Staszek**[1] · **Szymon Cielesta**[1] · **Paweł Urban**[1] · **Dorota Olszewska**[1] · **Grzegorz Lizurek**[1]

Abstract The purpose of this study is to evaluate seismic hazard parameters in connection with the evolution of mining operations and seismic activity. The time-dependent hazard parameters to be estimated are activity rate, Gutenberg–Richter b-value, mean return period and exceedance probability of a prescribed magnitude for selected time windows related with the advance of the mining front. Four magnitude distribution estimation methods are applied and the results obtained from each one are compared with each other. Those approaches are maximum likelihood using the unbounded and upper bounded Gutenberg–Richter law and the non-parametric unbounded and non-parametric upper-bounded kernel estimation of magnitude distribution. The method is applied for seismicity occurred in the longwall mining of panel 3 in coal seam 503 in Bobrek colliery in Upper Silesia Coal Basin, Poland, during 2009–2010. Applications are performed in the recently established Web-Platform for Anthropogenic Seismicity Research, available at https://tcs.ah-epos.eu/.

✉ Konstantinos Leptokaropoulos
kleptoka@igf.edu.pl

[1] Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

Introduction

Earthquake catalogs constitute a robust and beneficial tool for a variety of seismic analyses. Since seismicity is directly associated with physical quantities and mechanical properties of the crust such as strain accumulation, pore–fluid interactions and frictional response of the rupture zones, earthquakes provide a major source of information that cannot be usually obtained by direct measurements. Spatial and temporal seismicity rate anomalies are essentially reported as the most frequent intermediate-term precursory phenomenon in timescales varying from a few days to several years. The use of the well-established Gutenberg–Richter (G–R) law has been routinely incorporated into many seismic hazard assessment studies (e.g., Cornell 1968; Convertito et al. 2012). Alternatively, non-parametric approaches can be performed for hazard assessment evaluation under certain conditions (e.g., Kijko et al. 2001; Lasocki and Orlecka-Sikora 2008).

Seismic events may be controlled by either natural or anthropogenic factors. During the last decades, the rising demands for energy and minerals have sharpened the problem of hazards induced by exploration and exploitation of georesources (Davis et al. 2013). Among the diverse technologies capable of inducing earthquakes, one of the most well studied origins of anthropogenic hazard is underground mining. The undesirable rockmass response to mining operations was firstly observed back in the eighteenth century and during the last years it is being constantly reviewed and documented (e.g., Gibowicz and Lasocki 2001; Li et al. 2007; Gibowicz 2009, and references therein). A variety of factors control the rockmass fracturing and nucleation process in mines such as tectonic stresses accumulation, removal of material from the mine, explosions and the interaction among seismic events and

rockbursts. The instability of mining activities, especially when they are extended, may result in the residual subsidence and localized or generalized collapsing, potentially with significant societal and economic impacts. There are numerous documented cases where mining-induced seismicity has caused personnel injuries, production losses, extensive damage to infrastructure, collapse of drifts and stopes, and occasionally, fatalities. For all these reasons, seismic hazard assessment in mines is a task of paramount importance.

The purpose of this study is to evaluate seismic hazard parameters in connection with the evolution of mining operations and therefore to detect a causative relationship between seismic events and mining operations. The time-dependent hazard parameters to be estimated are the activity rate, the Gutenberg–Richter b-value, the mean return period and the exceedance probability of a prescribed magnitude for selected time windows related with the advance of the mining front. Four magnitude distribution estimation methods are applied and the results obtain from each one are compared with each other. Those approaches are maximum likelihood using the unbounded Gutenberg–Richter relation-based model (GRU), maximum likelihood using the upper-bounded Gutenberg–Richter relation-based model (GRT), unbounded non-parametric kernel estimation (NPU) and upper-bounded non-parametric kernel estimation (NPT). In addition, three different ways to construct subsequent datasets are applied: Time windows of constant duration, time windows with constant event number and time windows corresponding to constant front advance position. The spatial constraints are set in terms of the distance perpendicular and normal to the mining front at each time point (beginning and ending of time windows). The method is applied and results are discussed for the longwall mining of panel 3 in coal seam 503 in Bobrek colliery in Upper Silesia Coal Basin (USCB) in Poland, during 2009–2010. As shown in Fig. 1, this is a large area where coal mining has been carried out since the

eighteenth century and continues till nowadays in more that thirty mines in which coal is exploited by applying the longwall method.

Data

Bobrek Mine is a hard coal mine located in Bytom city in the area of USCB, Poland (Fig. 1). USCB constitutes one of the most seismically active mining areas worldwide, with almost 56,000 mining tremors of energy $E > 10^5$ J recorded between 1974 and 2005 (Stec 2007). The analysis of our study was based on selected set of data connected with exploitation in panel No. 3 in coal seam 503. This dataset is retrieved due to the specially organized virtual Laboratory for Monitoring Mining Induced Seismicity (LMMIS) where seismic data and technological data, such as mining front advance, were gathered. The seismic data had been prepared based on integration of two seismological networks operating at different scales (registration of events in the near and far seismic field). Additionally, information about the geology and the tectonics of the area was available. This set of data is integrated as an episode, which comprehensively describes a geophysical process induced or triggered by human technological activity, posing hazard for populations, infrastructure and the environment. All the data from this episode are available on TCS-AH platform as BOBREK MINE: local seismicity linked to longwall mining (https://tcs.ah-epos.eu/#episodes:BOBREK).

During underground mining works of longwall panel 3/503, a total of 2996 seismic events were recorded and analyzed, with a seismic energy greater than 10^2 J (local magnitude $M_L > 0.1$), occurred from April 12th, 2009 until July 8th, 2010. The strongest observed events with local magnitudes equal to 2.9, 3.7, 3.0 and 2.8 (seismic energy greater than 10^7 J) took place on May 20th 2009, December 16th 2009, February 5th 2010 and March 3rd 2010,

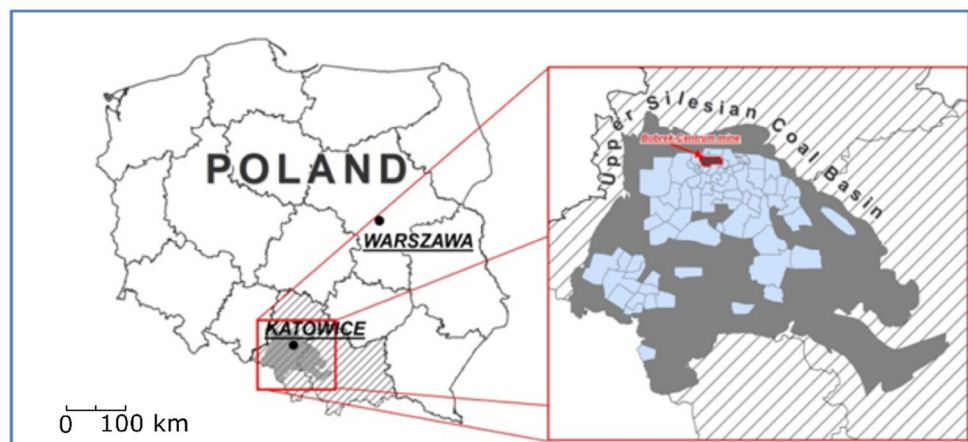**Fig. 1** The location of the Bobrek mine in Upper Silesia Coal Basin, Poland. Figure comes from TCS-AH platform https://tcs.ah-epos.eu/

Table 1 Events with $M_L \geq 2.8$ which occurred in analyzed period of time

ID	Occurrence time	M_L
1	20-May-2009, 14:28:48	2.9
2	16-Dec-2009, 02:06:37	3.7
3	05-Feb-2010, 10:59:18	3.0
4	11-Mar-2010, 00:07:21	2.8

respectively (Table 1). Figure 2 shows the local magnitude histogram. The information about positions of the longwall excavation front advance in Bobrek mine is given as polygon coordinates in different time moments. The distance between subsequent positions of the mining front is approximately 50 m and time interval is one month. The foci of the seismic events caused by underground mining operations in panel No. 3, the locations of underground seismic stations and the position of the mining front advance during excavation of that longwall are demonstrated in Fig. 3 (figure is a snapshot of 3D visualization available via TCS-AH platform). The average seismic activity of the panel 3/503 was 6.6 events per day (Fig. 4). However, this rate is far from being considered as stable since there are significant seismic activity changes in time. In Fig. 4 it is shown that seismicity rates start growing from the beginning of November 2009. Then, the largest event occurred ($M_L = 3.7$). The highest seismic activity equal to 18.4 events per day was observed from middle of January up to the middle of April 2010 and during that period two other strongest events occurred.

Methodology

Four different magnitude distribution estimation methods are applied presented in the following sub-sections. Hazard parameters are calculated and plotted for each one of the time windows for which sufficient data are available in order to perform the necessary calculation. A different minimum number of events in each window is considered for the calculations to be performed, according to the method selected: for unbounded Gutenberg–Richter method it is 7; for upper-bounded Gutenberg–Richter method it is 15; and for non-parametric kernel-based methods it is 50 events.

Note that for the Unbounded models (GRU and NPU) an infinite upper magnitude limit, M_{max}, is considered whereas in the Truncated (upper-bounded) approaches (GRT and NPT) M_{max} is evaluated using the Kijko–Sellevoll generic formula (Kijko and Sellevoll 1989; Kijko 2004; Lasocki and Urban 2011). If convergence is not reached the Robson and Whitlock (1964) simplified formula is used:
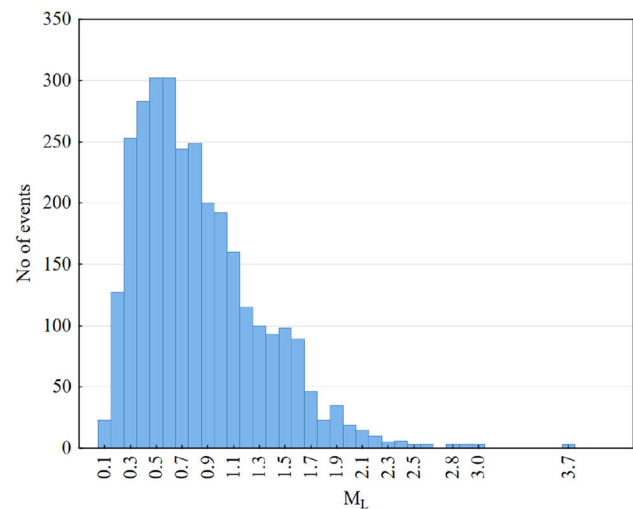
Fig. 2 Histogram of local magnitude of seismic events

$M_{max} = 2M_{maxobs} - M_{max2obs}$, where M_{maxobs} and $M_{max2obs}$ stand for the largest and second largest magnitudes in a given catalog, respectively.

The hazard parameters evaluated in this study are the mean return period for a specified magnitude and the exceedance probability of a specified magnitude to be exceeded within a certain time period. The mean return period of magnitude M is the average elapsed time between the consecutive earthquakes of magnitude M. Given the mean activity rate for events with $M \geq M_{min}$, λ, in a specified time period and the corresponding cumulative magnitude distribution function, $F(M)$, the mean return period is estimated as:

$$T = \frac{1}{\lambda(1 - F(M))} \tag{1}$$

The exceedance probability of a specified magnitude, M, during a predefined time period, Δt, is given as

$$P = 1 - e^{-\lambda \Delta t (1 - F(M))}. \tag{2}$$

The cumulative magnitude distribution function, $F(M)$, is calculated with respect to the selected method:

Unbounded GR law (GRU)

Assume that the unlimited Gutenberg–Richter relation leads to the exponential model of distribution for events with magnitude above the catalog completeness level M_{min}. Under this assumption, an infinitely large maximum magnitude is possible. The shape parameter of this distribution and consequently the G–R b-value is estimated by maximum likelihood method (Aki 1965; Utsu 1999) with the Probability Density Function (PDF) of magnitudes given as:

$$f(M) = \beta e^{-\beta \left(M - M_{min} + \frac{\Delta M}{2}\right)} \tag{3}$$

Fig. 3 Distribution of seismic events during excavation of longwall panel no. 3 in Bobrek Mine. *Black triangles* represent the nearest seismic stations, and *black lines* show the subsequent positions of the longwall excavation front advance
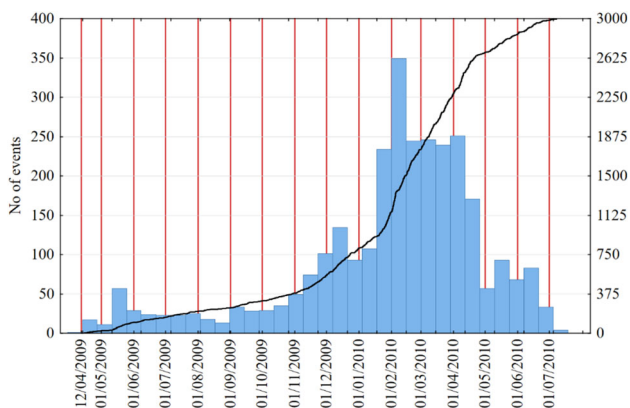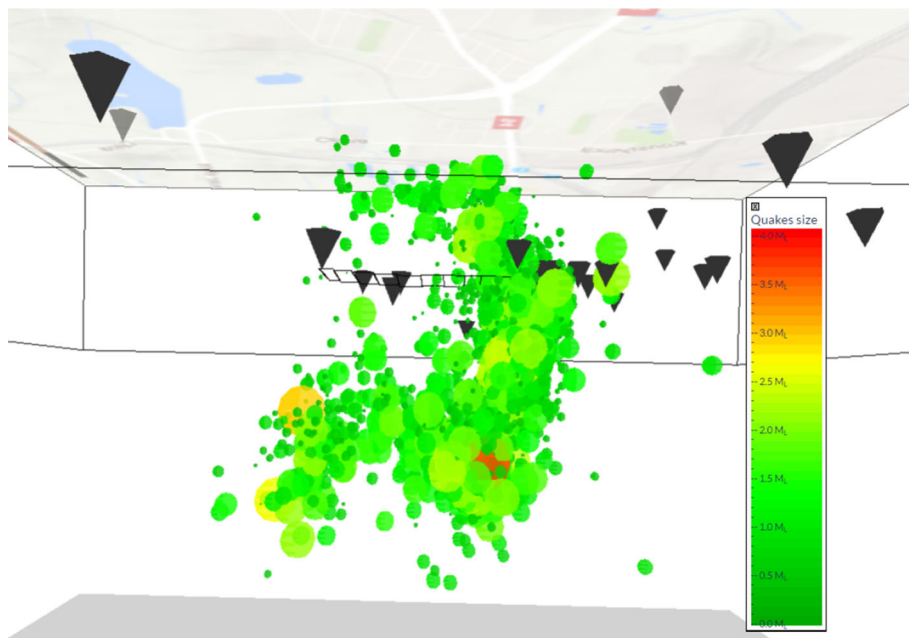

Fig. 4 Seismic activity plot. *Blue bars* show the number of events per 2 weeks. *Black line* is cumulative number of events. *Red lines* represent the date of mining front localization

with parameter $\beta = \ln 10 \cdot b$, where b stands for the well-known G–R b-value, which is estimated by the well-known Aki (1965) maximum likelihood estimator as:

$$b = \frac{1}{\ln(10)[\langle M \rangle - (M_{\min} - \Delta M/2)]},$$ (4)

where $\langle M \rangle$ is the sample mean of the considered events. The introduction of term $\Delta M/2$ accounts for a correction for the finite binning width of the catalog, ΔM (Utsu 1966; Bender 1983), which is equal to the minimum non-zero difference among data. The corresponding Cumulative Distribution Function (CDF) of (3) is:

$$F(M) = 1 - e^{-\beta\left(M - M_{\min} + \frac{\Delta M}{2}\right)} \quad \text{for} \quad M \geq M_{\min}, \quad 0$$
$$\text{otherwise}$$ (5)

Truncated GR law (GRT)

The assumption on the upper-bounded Gutenberg–Richter relation leads to the upper truncated exponential model of distribution for events with magnitude above the catalog completeness level $M_{\min}$. The PDF of magnitudes is given as (Page 1968; Kijko and Sellevoll 1989, also for b-value evaluation):

$$F(M) = \frac{\beta e^{-\beta\left(M - M_{\min} + \frac{\Delta M}{2}\right)}}{1 - e^{-\beta\left(M_{\max} - M_{\min} + \frac{\Delta M}{2}\right)}} \quad \text{for} \quad M_{\min} \leq M \leq M_{\max},$$
$$0 \quad \text{otherwise}$$ (6)

With β and $\Delta M/2$ as explained in Eqs. (3) and (4). The corresponding CDF of (6) is:

$$F(M) = 1 \begin{cases} 0 & \text{for} \quad M < M_{\min} \\ \dfrac{1 - e^{-\beta\left(M - M_{\min} + \frac{\Delta M}{2}\right)}}{1 - e^{-\beta\left(M_{\max} - M_{\min} + \frac{\Delta M}{2}\right)}} & \text{for} \quad M_{\min} \leq M \leq M_{\max} \\ 1 & \text{for} \quad M > M_{\max} \end{cases}$$ (7)

Non-parametric approaches (NPU and NPT)

The kernel estimator approach proposed by Kijko et al. (2001) is a non-parametric (model free or data driven) alternative for estimating the magnitude distribution functions. This non-parametric approach is based on the kernel density estimator that totals the symmetric probability densities (kernels), individually associated with data points as (Parzen 1962; Silverman 1986):

$$\hat{f}(M|\{M_i\},h) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{M-M_i}{h}\right) \qquad (8)$$

where h is a non negative smoothing parameter (bandwidth), M_i are the magnitudes and $K(x)$ is a kernel function. The Kernel estimations chosen here for probability density (9) and cumulative distribution (10) have the forms of those adopted by Lasocki and Orlecka-Sikora (2008):

$$\hat{f}_M(M) = \frac{\sum_{i=1}^{n}\frac{1}{a_i h}e^{-\frac{1}{2}\left(\frac{M-M_i}{a_i h}\right)^2}}{\left(\sqrt{2\pi}\right)\sum_{i=1}^{n}\left[\Phi\left(\frac{M_{\max}-M_i}{a_i h}\right) - \Phi\left(\frac{M_{\min}-M_i}{a_i h}\right)\right]} \qquad (9)$$

$$\hat{F}_M(M) = \frac{\sum_{i=1}^{n}\left[\Phi\left(\frac{M-M_i}{a_i h}\right) - \Phi\left(\frac{M_{\min}-M_i}{a_i h}\right)\right]}{\sum_{i=1}^{n}\left[\Phi\left(\frac{M_{\max}-M_i}{a_i h}\right) - \Phi\left(\frac{M_{\min}-M_i}{a_i h}\right)\right]}, \qquad (10)$$

where n is the sample size, $\Phi(x)$ is the standard Gaussian cumulative distribution, a_i ($i = 1,2, \dots n$) are the local bandwidth factors and m_i, are the magnitudes with $M_{\min} \le M \le M_{\max}$. Note that $M_{\min}$ is equal to the completeness threshold of a given catalog. It is assumed that the magnitude distribution is unlimited from the right hand side (i.e., infinite maximum magnitude). The shape of the kernel estimates depends primarily on the value of h. From the point of view of the use of estimators (9), (10) in the hazard analysis, a global, integrant agreement between the actual density and its estimates is of the utmost importance. Therefore, we select the smoothing factor applying the least squares cross-validation technique that requires minimizing the integral of the squared difference between the actual density, $f(\xi)$, and the estimate (e.g. Bowman et al. 1984):

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} \left[\hat{f}(\xi) - f(\xi)\right]^2 \mathrm{d}\xi \qquad (11)$$

It has been shown (Kijko et al. 2001) that in the case of the Gaussian kernel this criterion is fulfilled if h is the root of the equation:

$$\sum_{i,j}\left\{\left[\frac{(M_i - M_j)^2}{2h^2} - 1\right]\exp\left[-\frac{(M_i - M_j)^2}{4h^2}\right]\right.$$
$$\left. -2\left[\frac{(M_i - M_j)^2}{h^2} - 1\right]\exp\left[-\frac{(M_i - M_j)^2}{2h^2}\right]\right\} = 2n$$
$$(12)$$

The local bandwidth factors, $\{\alpha_i\}$ can modify the width of the kernels at certain data points. Due to the fact, that the most important for the hazard analysis range of magnitudes, is that of the larger values, where the data are very sparse, the present version of the estimators uses the bandwidth factors that widen the kernels associated with

data points from this range (Lasocki and Orlecka-Sikora 2008):

$$a_1 = \frac{1}{\left[\frac{\hat{f}(M_i|\{M_i\},h)}{g}\right]^{1/2}}, \qquad (13)$$

where $\hat{f}$, is the constant kernel estimator in the unbounded magnitude range, and

$$g = \left[\prod_{i=1}^{n} \hat{f}(M_i|\{M_i\},h)\right]^{1/n} \qquad (14)$$

is the geometric mean of all constant kernel estimates (Silverman 1986).

Parameter values defined

Mining tremors occurrence is strongly associated with the excavation operations and therefore, seismicity properties and hazard parameters are estimated as a function of time and front advance. For this purpose, only the fraction of events which satisfies predefined properties is considered for such calculations. These events are selected after applying the following constrains: The completeness magnitude of the catalog was identified by considering the stations distances from the focal areas and the signal to noise ratio (Mutke et al. 2016). In such way $M_{\min}$ was set equal to 0.6. The margin along front strike (equal for both directions of the front strike) and distance along the front advance (equal for both in front of and behind the front) was set equal to 50 and 100 m, respectively, following Kozłowska (2013). Calculations were performed for overlapping time windows generated in three different modes: According the "time mode", events included in subsequent 30-day time windows were considered, overlapping per 1 day. In this approach time windows of the same span are used for hazard parameter evolution estimation. In the "events mode", the time windows are selected in such a way that they include equal number of events (also overlapping per one day) and therefore the parameter estimation errors are comparable in all the datasets. Finally, in the "front mode", a subsequent time window starts at the time point when the front advance is moved from the position it was in the beginning of the previous time window, to a distance indicated by a predefined value. In this case, the time windows correspond to approximately equal material mass removal from the mine (those windows are also overlapping per 1 day). The parameters set for our analysis in those three modes were 30 days in the "time mode", 80 events in the "event mode" and 50 meters in the "front mode". Finally, the event magnitude for the mean return period and the exceedance probability evaluation was set to 3.0, equal to the second largest magnitude in the dataset.
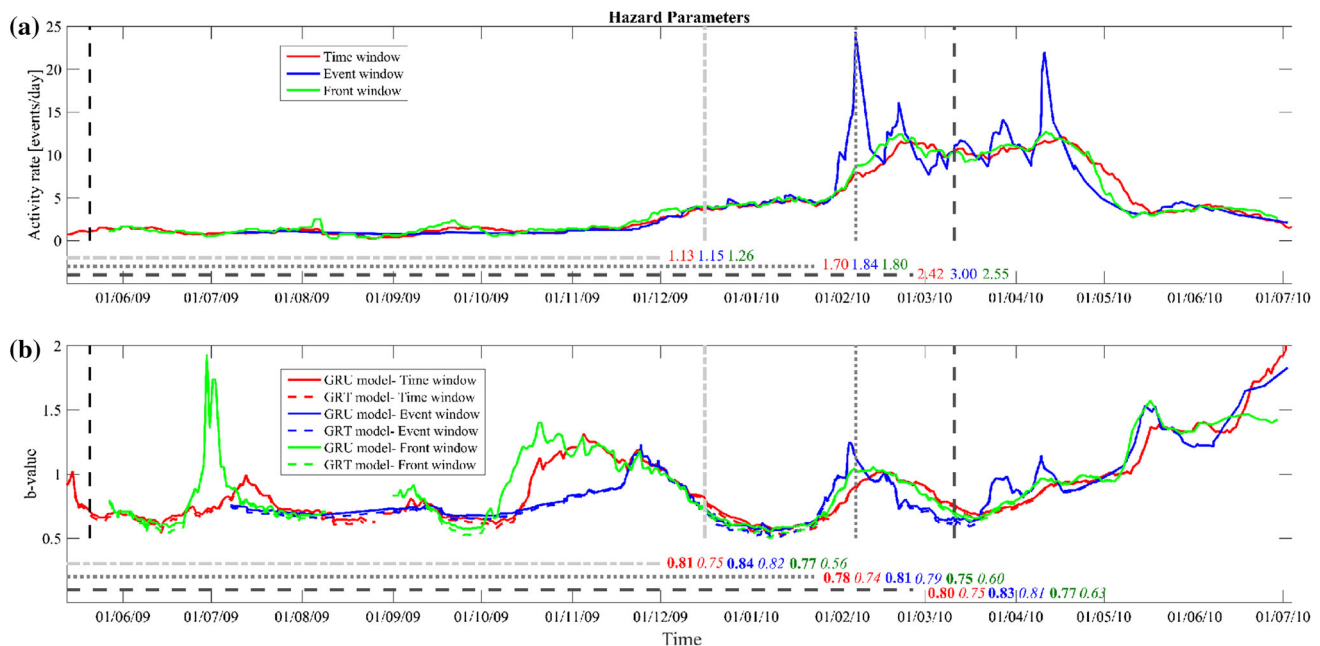
Fig. 5 Changes of seismic activity rate (**a**) and *b*-value (**b**) in time calculated using: time windows of 30 days, event windows of 80 events, front windows of 50 m (see text for details). Strong events are plotted with *gray dashed vertical lines*. *Gray horizontal lines* indicate time periods chosen to calculate mean values of parameters (indicated by the numbers in *colored fonts*)

Results

We analyzed changes of seismic activity rate, *b*-value, return period of $M_L = 3.0$ and exceedance probability of $M_L = 3.0$ (within 30 days) as a function of time before the occurrence of strong events with $M_L \geq 2.8$ (Figs. 5, 6, 7). Four events with such magnitudes occurred in analyzed period of time (Table 1). In further analyses, we considered only events with IDs 2–4, because of the too short time period between the initiation of registrations and occurrence of the event with ID 1.

It can be observed that in the case of events (2) and (3) activity rate increases, although in different manner, before the occurrence of considered events (Fig. 5a). The opposite statement can be done for event (4), where for both time and front windows activity rate decreases before the event. However, slight increase shortly before the event is observable in the case of event window mode which due to the constant amount of events in each window is more sensitive in detecting sudden changes in occurrence rates. Additionally, it is worth mentioning that only the curve resulting from event windows shows a evident peak in activity rate before event (3). In the case of other windows the increase is slight.

Changes in *b*-values before the occurrence of strong events are also not completely consistent with each other. Before events (2) and (4) decrease of *b*-value can be observed. However, before the occurrence of event (3), *b*-value increases in the case of all window types. Again,

changes in the case of curve obtained using event window mode are the largest (Fig. 5b).

To describe quantitatively how distribution of hazard parameters changes before high magnitude events, we compared the last value of given parameter calculated before the occurrence of the event with the mean value of the same parameter calculated in time window from the beginning of observations to 14 days before considered event (Fig. 5). The time period of 14 days was chosen on the basis of observations of parameter changes before the big events (Fig. 5). The results reveal that seismic activity rate ratio is >1 for all events and window types, with a maximum value of 13.23 for event window for event (3) (Fig. 8a). *b*-value ratios of events (2) and (4) have generally values <1, what is a result of *b*-value decrease before big events mentioned earlier, however, for the event (3) the ratio of *b*-values are highly above 1 (Fig. 8b).

The values of exceedance probability of $M_L = 3.0$ (within 30 days) and return period of $M_L = 3.0$ just before the occurrence of events (2), (3) and (4) are plotted in Fig. 9. It can be observed that exceedance probability values obtained using Gutenberg–Richter methods are much higher than those estimated with non-parametric kernel methods (Fig. 9). Only event windows used in kernel methods give significantly high probabilities in case of events (3) and (4). Additionally, the increase of exceedance probability calculated with kernel methods occurs just before high magnitude events (Fig. 6), which makes it difficult to use for prediction purposes. On the contrary,
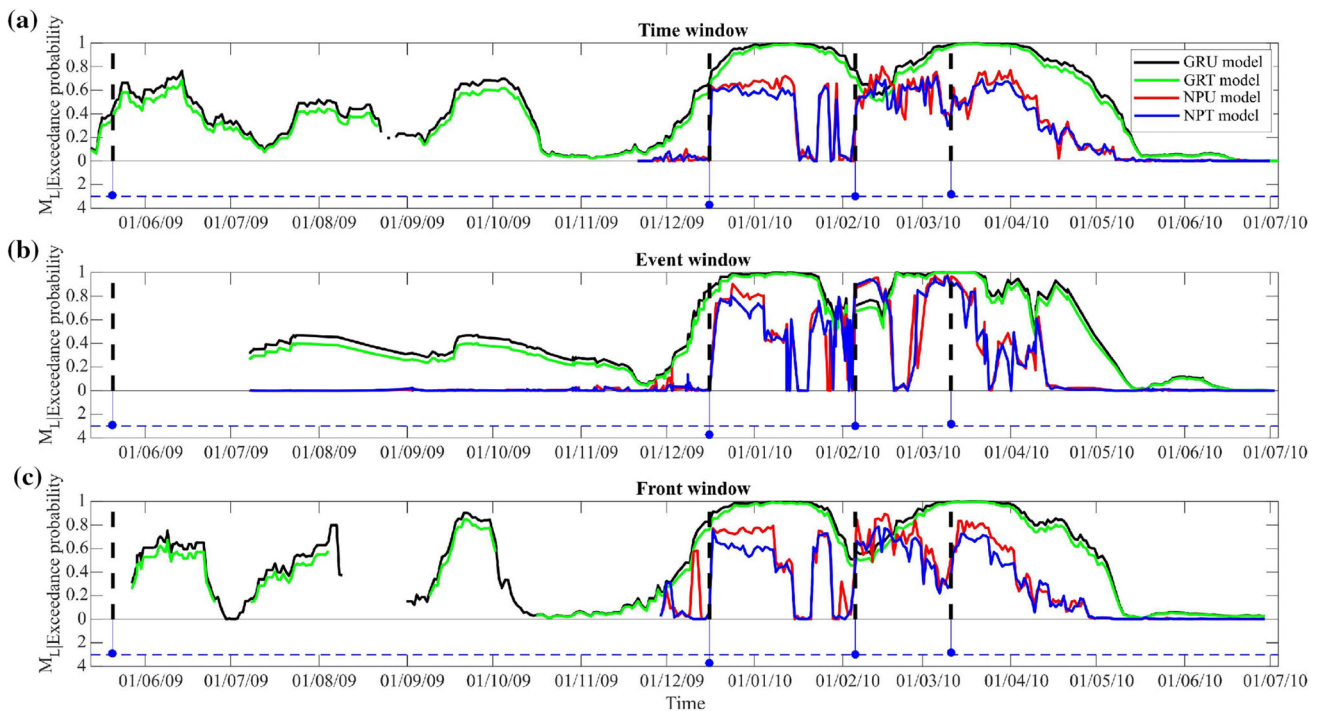
Fig. 6 Exceedance probability of $M_L = 3$ calculated using: time windows of 30 days (**a**), event windows of 80 events (**b**), front windows of 50 m (**c**)
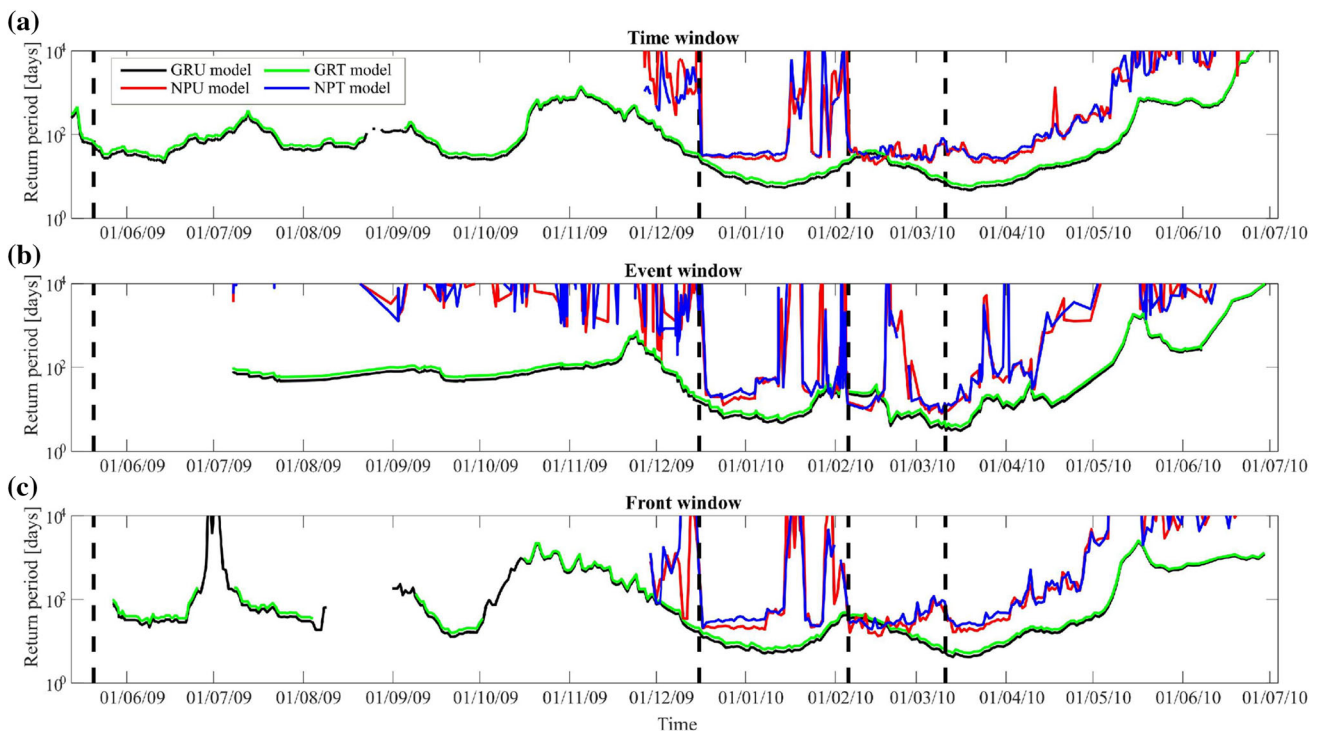

Fig. 7 Return period of $M_L = 3$ calculated using: time windows of 30 days (**a**), event windows of 80 events (**b**), front windows of 50 m (**c**)

changes of exceedance probability obtained using Gutenberg–Richter approaches are much more gradual, and thereby it is easier to follow the observed trend.

Return periods of $M_L = 3.0$ calculated with non-parametric kernel methods are comparable to those calculated with Gutenberg–Richter methods only in case of event (3)

(Fig. 10). In all other cases, the values of kernel estimations of return periods are much higher. It is worth mentioning that in the case of event (3), kernel-based return periods for event

Fig. 8 Ratios of the last values of activity rate (**a**) and *b*-value (**b**) before events (2), (3) and (4) to the mean values of these parameters calculated in time periods marked in Fig. 5. *Black solid lines* indicate ratio value equal 1

Fig. 9 Exceedance probability values before events (2), (3) and (4) calculated with unbounded Gutenberg–Richter (**a**), upper-bounded Gutenberg–Richter (**b**), unbounded kernel (**c**) and upper-bounded kernel (**d**) methods. *Black solid lines* indicate exceedance probability equal 0.5

windows are even lower than Gutenberg–Richter ones. In case of event (4), the lowest kernel return periods are also calculated on the basis of event windows (Fig. 10). These and previous observations can lead us to a conclusion that nonparametric kernel-based methods give compatible hazard results only if event window is used for calculations (i.e., meaning constant number of events in variable size window).

Previous analysis (Fig. 8, 9, 10) of parameters' changes before events (2), (3) and (4) leads us to the following conclusions concerning expectance of events on the basis of time-dependent hazard calculations. First, on the basis of temporal changes of activity rate, *b*-value, exceedance probability and return period, event (4) can be considered as the most expected one (activity rate increase, *b*-value decrease, high exceedance probability according to all methods, relatively short return period). Second, estimations of exceedance probability and return period with kernel-based methods do not give any possibility to predict the event (2) despite the fact that activity rate and *b*-value changes can suggest the occurrence of an impending strong event.
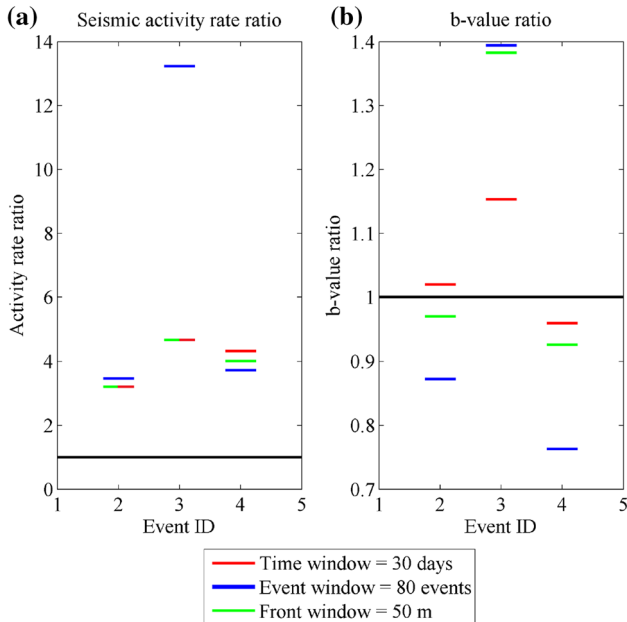
Discussion and conclusions

The use of fundamental observational and empirical parameters such as activity rate and G–R *b*-value may prove to be helpful in evaluation of seismic hazard for
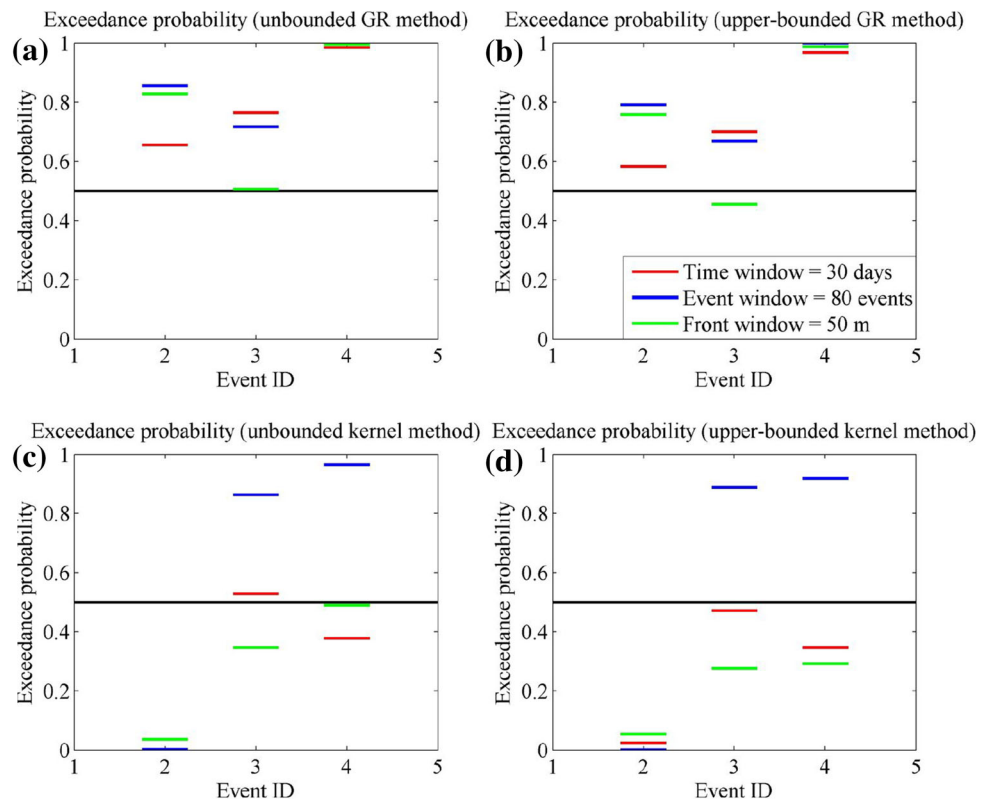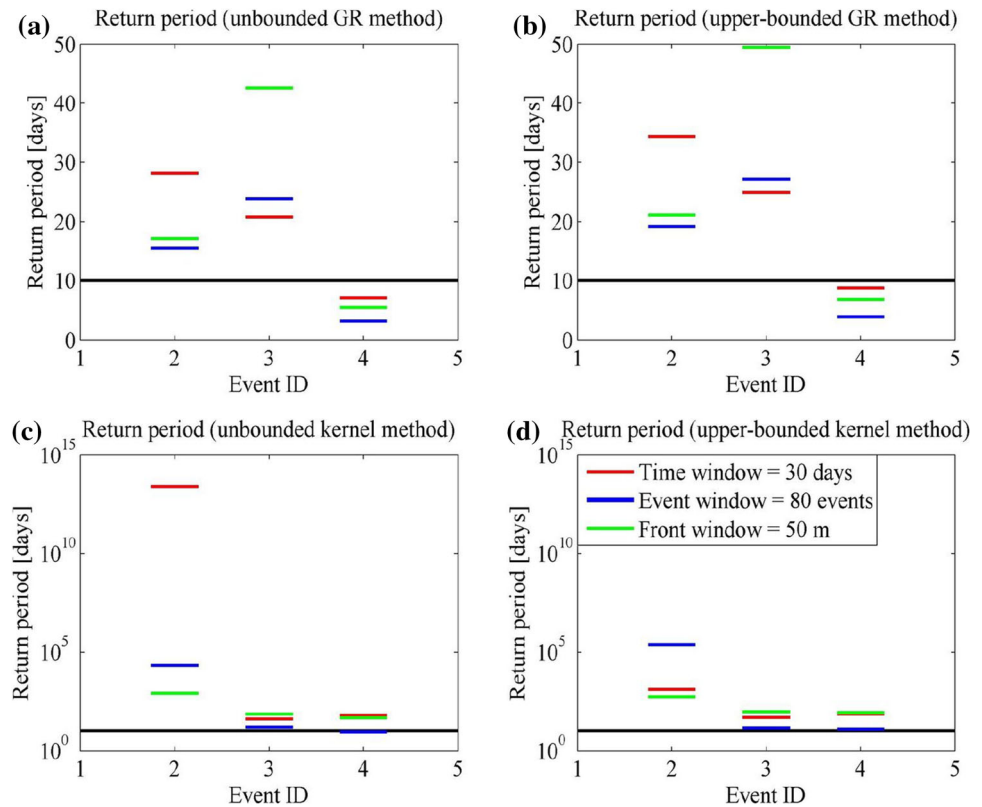
Fig. 10 Return period values before events (2), (3) and (4) calculated with unbounded Gutenberg–Richter (**a**), upper-bounded Gutenberg–Richter (**b**), unbounded kernel (**c**) and upper-bounded kernel (**d**) methods. *Black solid lines* indicate return period equal 10

specific study areas such as longwall mining. The G–R *b*-value has a clear physical interpretation, defining the relative proportion of the number of large events to small events. Anomalies in seismicity rates and *b*-value can be considered as an indicator of the stress state in rock mass (e.g., low *b*-values may indicate growing of stress, Scholz 1968). These anomalies are capable to lead to strong tremors and thus provide high seismic risk. Information of such kind can, therefore, be a premise for the decision to use specific measures to prevent bumps and adjust mining operations. Probability of the strong events occurring is higher when *b*-value is decreasing and simultaneously activity rate is increasing. In the present study, this case is clearly observed for events 2 and 4 for all window types (with the exception of *b*-value for event 2, which slightly increases when time window is considered). However, the activity rate is increased in all cases studied (Fig. 8a).

The need of introducing a non-parametric estimation for magnitude distribution arose from observed deviations of specific catalogs from G–R law in both natural and induced seismicity. Especially, induced seismicity exhibits diverse seismogenic processes in comparison with natural seismicity and it strongly depends on the human technological and production activities (Kijko et al. 1987; Trifu et al. 1993; Fritschen 2010; Maghsoudi et al. 2014). Two causes had been considered responsible for inconsistency of G–R law with the observed events distribution. The coupling

between tectonic stresses and mining activities in Bobrek mine has been studied by Marcak and Mutke (2013). Specifically for event 2, Kozłowska et al. (2016) showed that it was a tectonic event triggered by the ongoing exploitation and that the subsequent seismicity was a result of combined tectonic, coseismic and mining-induced stresses.

First, this inconsistency was attributed to the presence of broadly defined types of mining events (e.g., Gibowicz and Kijko 1994): events directly related to the mining operations and events resulting from the release of residual tectonic stresses accumulated in the rock mass. These two types of events are characterized by extensively different properties and features. Second, the non-linearity was attributed to the local geology and tectonics, i.e., heterogeneity and discontinuity of the rock structures with different thickness, strength, different in various parameters. For example, Naoi et al. (2014) showed that the distribution of magnitudes is consistent with the G–R relationship only when events closely linked to blasting are not taken into account.

As a result, magnitude distribution complexity may arise, characterized by number of modes or bumps in magnitude density. Evidences for the existence of multimodal earthquake size distribution were presented in many papers (e.g., Gibowicz and Kijko 1994; Gibowicz and Lasocki 2001; Lasocki and Orlecka-Sikora 2008).

Magnitude distribution may statistically significantly differ from the exponential distribution but the differences may be small (Urban et al. 2016). G–R model seems often inadequate to describe the earthquake size distribution.

To evaluate which estimation method would be more appropriate in our dataset, we performed statistical tests to examine the compatibility of earthquake size distribution with exponential distribution. First, we used Kolmogorov–Smirnov (K–S) goodness-of-fit one-sample test to check whether earthquake size distribution statistically differs from the exponential distribution. Details of this procedure and the corresponding findings in several cases of induced seismicity are described in the paper of Urban et al. (2016). Second, we examined the earthquake size distribution for its complexity using smoothed bootstrap test for multi-modality. Multimodality test allows us to examine the presence of irregularities in the distribution of given magnitude values by demonstrating the presence of more than one mode or more than one bump in the probability density function (Lasocki and Papadimitriou 2006; Lasocki and Orlecka-Sikora 2008). These features indicate the presence and mixing of different processes generating the events.

To perform the multimodality and K–S tests, first the randomization of any observed magnitude has been done within the rounding error. Randomization of magnitude was carried out by transformation (Lasocki 2001; Lasocki and Papadimitriou 2006)

$$\hat{M} = F^{-1}\{u[F(M + 0.5\delta M) - F(M - 0.5\delta M)] + F(M - 0.5\delta M)\},$$

(15)

where $\hat{M}$ and M are randomized and observed magnitude values, respectively, Δ is accuracy of magnitude estimation, u is the value from uniform distribution [0, 1], $F(\cdot)$ is magnitude CDF, and $F^{-1}(\cdot)$ is its inverse function.

In the K–S test case, we test the null hypothesis H_0: the distribution of magnitude is exponential. To verify H_0, we estimate p values and compare them with the adopted significance level $\alpha = 0.05$. A small p value suggests that the null hypothesis may be false. In the test for multi-modality case, we test two null hypotheses: H_0^1: density function of earthquake size distribution has no more when 1 mode; H_0^2: density function of earthquake size distribution has no more than 1 bump. Due to the fact that the magnitude randomization gives slightly different p-values and significances of the considered null hypotheses, we assumed as a final value its mean value (from 1000 magnitude randomized catalogs) considering their standard deviation as well.

The results from K to S test strongly suggest that earthquake size distribution is not exponential with mean p equal to 4.6×10^{-7} and standard deviation equal 3.9×10^{-7}. Because we estimate G–R b-value from the sample it requires calculation of additional statistics (D_m) (Urban et al. 2016). The value of the (D_m) is 2.82 which confirms previous results with the significance at least equal to 0.01.

The mean significance of considered null hypotheses from test for multimodality is 0.11 for H_0^1 with standard deviation equal 0.02 and 0.04 for H_0^2 with standard deviation equal 0.02. The results indicated that with 89–96% probability magnitude distribution is more complex than linear model.

Based on the results of K–S test and smoothed bootstrap test for multimodality, the non-parametric method for estimate of hazard parameters would be more appropriate. On the other hand, non-parametric technique starts to be effective for sample sizes starting from 200 events (Kijko et al. 2001). However, the hazard analysis performed in the present study considers time windows which contain diverse and mainly smaller number of events. Taking into account the size of the dataset we may conclude that segmentally, the G–R model is a good approximation for this type of analysis, providing also more comprehensive and stable results.

As a summary, we conclude to the following points:

- The obtained results strongly depend on the data number contained in the analyzed windows. According to that, event window approach may be preferable because it leads to identical or at least comparable uncertainties of the estimated parameters (equal size data samples are tested—see "Appendix A").
- Unbounded and upper-bounded GR approaches lead to similar results (see also "Appendix A"). In the same way, unbounded and upper-bounded non-parametric methods also lead to similar results. On the other hand, there are distinct differences between parametric and non-parametric estimation techniques.
- Kernel-based methods exhibit sharp fluctuations of estimated parameter values, which are essentially, not practical in terms of prediction implications for datasets of such size. However, they seem to provide consistent results when event windows are considered.
- On the basis of all methods' results, the last event (4) can be considered as the most expected one. It should be noted that there are several other windows for which exceedance probability is practically 1, when the G–R approaches are considered.
- The analysis of the catalog as a whole strongly indicates that earthquake size distribution does not obey G–R relation. However, smaller datasets in the analyzed windows show that G–R relation is an adequate approximation providing more stable results.

Appendix A

b-Values uncertainties

The *b*-values together with their uncertainties are demonstrated in this appendix (Fig. 11). Unbounded and upper bounded G–R model provide identical results in the vast majority of cases. The upper and lower bounds in the figure correspond to one standard deviation of the Aki's maximum likelihood estimation, given by:

$$\sigma_b = \frac{b}{\sqrt{N}} \tag{A1}$$

where, N, is the sample size. It is shown that these boundaries are almost constant in the event window case (middle frame) but they exhibit significant fluctuations in the other two cases (time window and front window).

Apparently, it is impossible to provide a straightforward comparison regarding the efficiency of each window selection method, since there is no direct correspondence among those windows (each method produces unique windows with different time range and data included). However, we may draw some conclusions concerning the average behavior of these methods. As illustrated in Figs. 12 and 13, in both variations of Gutenberg–Richter law, the event window method leads to the smallest errors in comparison with the time and front windows. Almost two-thirds of the error values when event window is considered are below 0.1 and the maximum error values do not exceed 0.2. Time and front windows on the other hand reproduce larger error values (both average and maximum error values).
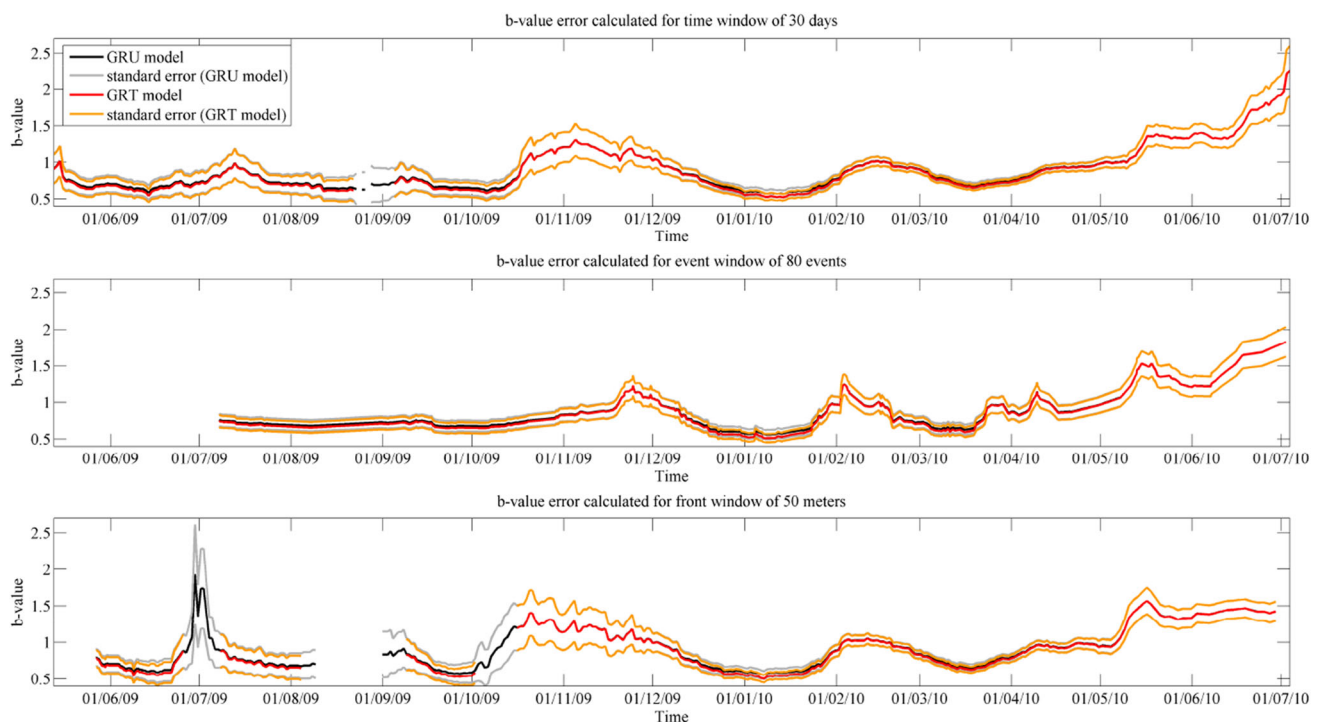

Fig. 11 *b*-values estimated by the unbounded Gutenberg–Richter (*black curve*) and upper-bounded Gutenberg–Richter (*red curve*), together with their respective standard errors (*gray* and *yellow curves*, respectively) for all datasets corresponding to time windows (*upper frame*), event windows (*middle frame*) and front windows (*lower frame*)
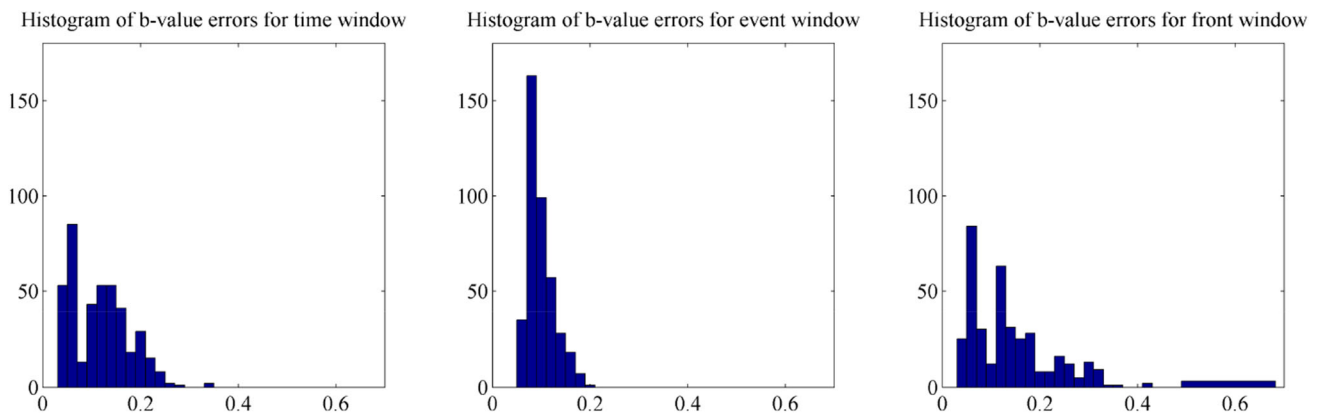
Fig. 12 Histograms with b-value standard errors for the 3 methods of window selection, assuming the unbounded Gutenberg–Richter approach
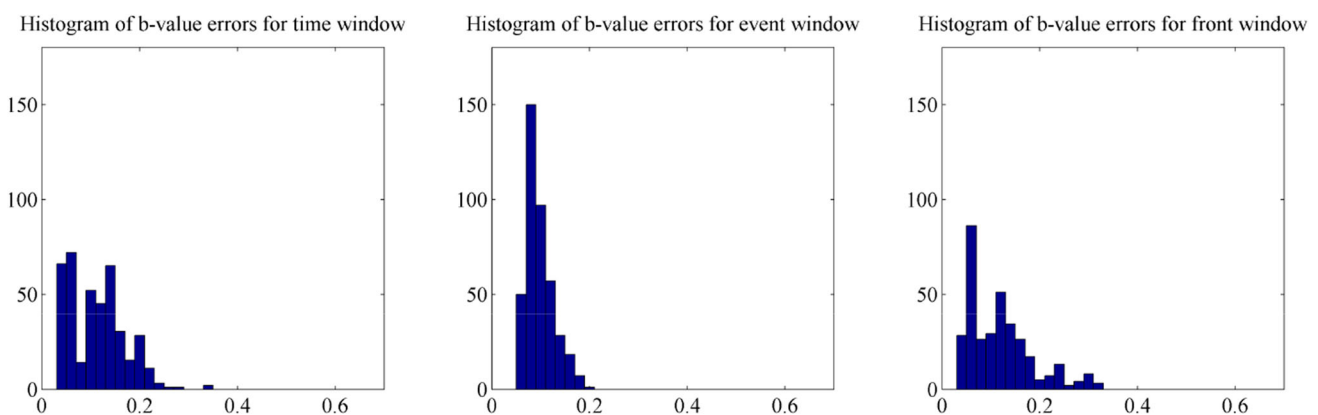

Fig. 13 Histograms with b-value standard errors for the 3 methods of window selection, assuming the upper-bounded Gutenberg–Richter approach

Following these results we may conclude that although the number of the events determines the individual estimation accuracy for each data set, the constant event window is in general more accurate in calculation of b-values and consequently of hazard parameters. This is also valid for the non-parametric (purely data driven) approaches which are even more sensitive to the sample size. In addition, constant event approach not only ensures a lower average error, but also a comparable error in all windows created, since all datasets exhibit exactly the same sample size.

Appendix B

TCS-AH web platform (https://tcs.ah-epos.eu/)

The IS-EPOS IT-platform is an open virtual access point for researchers studying anthropogenic seismicity and related hazards into European Plate Observing System—Anthropogenic Hazards Thematic Core Services (EPOS

AH-TCS). IS-EPOS platform constitutes a digital research space for providing a permanent and reliable access to advanced Research Infrastructures (RI) to the Induced Seismicity (IS) Community. This objective is implemented as a prototype which offers access to various datasets related to selected anthropogenic seismicity cases, specialized software for elementary and advanced data analysis and document repository. The relevant seismic and non-seismic data are gathered in the so-called episodes of induced seismicity. The IS-EPOS platform integrates presently seven episodes of anthropogenic seismicity respectively linked to underground hard rock and coal mining in Poland, shale gas extraction in UK, hydroelectric energy production (Poland and Vietnam) and geothermal energy production experiment in Germany. The researcher accessing the platform can make use of low level software services for data browsing, selecting and visualizing and a number of high level services for advanced data processing out of which the probabilistic seismic analysis service group is particular rich. The IS-EPOS platform is a working prototype of AH-TCS belonging to pan-European

multidisciplinary research platform created within EPOS long term plan for integration of national and transnational research infrastructures for solid earth science in Europe. Platform is available for registered users for free (https://tcs.ah-epos.eu). For purpose of this work two services were used: Time-Dependent Seismic Hazard (in mining front surroundings) and Time-Dependent Seismic Hazard (in mining a selected area). The episodes data, access to services and document repository and come of the source codes are available for registered users.

References

Aki K (1965) Maximum likelihood estimate of b in the formula logN = a − bM and its confidence limits. Bull Earthq Res Inst Tokyo Univ 43:237–239

Bender B (1983) Maximum likelihood estimation of b-values for magnitude grouped data. Bull Seismol Soc Am 73:831–851

Bowman AW, Hall P, Titterington DM (1984) Cross-validation in non-parametric estimation of probabilities and probability densities. Biometrika 71:341–351

Convertito V, Maercklin N, Sharma N, Zollo A (2012) From induced seismicity to time-dependent seismic hazard. Bull Seismol Soc Am 102:2563–2573

Cornell AC (1968) Engineering seismic risk analysis. Bull Seismol Soc Am 58:1583–1606

Davis R, Foulger G, Bindley A, Styles P (2013) Induced seismicity and hydraulic fracturing for the recovery of hydrocarbons. Mar Petrol Geol 45:171–185

Fritschen R (2010) Mining-induced seismicity in the Saarland, Germany. Pure Appl Geophys 167:77–89

Gibowicz S (2009) Seismicity induced by mining: recent research. Adv Geophys 51:1–53

Gibowicz SJ, Kijko A (1994) An introduction to mining seismology. Academic Press, San Diego

Gibowicz SJ, Lasocki S (2001) Seismicity induced by mining: ten years later. Adv Geophys 44:39–181

Kijko A (2004) Estimation of the maximum earthquake magnitude, mmax. Pure Appl Geophys 161:1655–1681. doi:10.1007/s00024-004-2531-4

Kijko A, Sellevoll MA (1989) Estimation of earthquake hazard parameters from incomplete data files. Part I. Utilization of extreme and complete catalogs with different threshold magnitudes. Bull Seismol Soc Am 79:645–654

Kijko A, Drzezla B, Stankiewicz T (1987) Bimodal character of the distribution of extreme seismic events in Polish mines. Acta Geophys Pol 35:157–166

Kijko A, Lasocki S, Graham G (2001) Nonparametric seismic hazard analysis in mines. Pure Appl Geophys 158:1655–1676

Kozłowska M (2013) Analysis of spatial distribution of mining tremors occurring in rudna copper mine (Poland). Acta Geophys 61:1156–1169

Kozłowska M, Orlecka-Sikora B, Rudziński Ł, Cielesta S, Mutke G (2016) Atypical evolution of seismicity patterns resulting from the coupled natural, human-induced and coseismic stresses in a longwall coal mining environment. Int J Rock Mech Min Sci 68:5–15

Lasocki S (2001) Quantitative evidences of complexity of magnitude distribution in mining-induced seismicity: Implications for hazard evaluation. In: van Aswegen G, Durrheim RJ, Ortlepp WD (eds), Proceedings of the fifth international symposium on rockbursts and seismicity in mines (RaSiM 5) 'Dynamic rock mass response to mining'. South African Institute of Mining and Metallurgy, Johannesburg, 543–550

Lasocki S, Orlecka-Sikora B (2008) Seismic hazard assessment under complex source size distribution of mining-induced seismicity. Tectonophysics 456:28–37

Lasocki S, Papadimitriou EE (2006) Magnitude distribution complexity revealed from seismicity in Greece. J Geophys Res. doi:10.1029/2005JB003794

Lasocki S, Urban P (2011) Bias, variance and computational properties of Kijko's estimators of the upper limit of magnitude distribution, Mmax. Acta Geophys 59:659–673

Li T, Cai MF, Cai M (2007) A review of mining-induced seismicity in China. Int J Rock Mech Min 44:1149–1171

Maghsoudi S, Hainzl S, Cesca S, Dahm T, Kaiser D (2014) Identification and characterization of growing large-scale en-echelon fractures in a salt mine. Geophys J Int 196:1092–1105. doi:10.1093/gji/ggt443

Marcak H, Mutke G (2013) Seismic activation of tectonic stresses by mining. J Seismol 17:1139–1148

Mutke G, Pierzyna A, Barański A (2016) b-Value as a criterion for the evaluation of rockburst hazard in coal mines. In: Mitri HS, Shnorhokian S, Kumral MK, Sasmito A, Sainoki A (eds), Proceedings of 3rd international symposium on mine safety, science and engineering, August 13–19 2016, Montreal, Canada, 1–5

Naoi M, Nakatani M, Horiuchi S, Yabe Y, Philipp J, Kgarume T, Morema G, Khambule S, Masakale T, Ribeiro L, Miyakawa K, Watanabe A, Otsuki K, Moriya H, Murakami O, Kawakata H, Yoshimitsu N, Ward A, Durrheim R, Ogasawara H (2014) Frequency-magnitude distribution of −3.7 ≤ M W ≤ 1 mining-induced earthquakes around a mining front and b value invariance with post-blast time. Pure Appl Geophys 171:2665–2684. doi:10.1007/s00024-013-0721-7

Page R (1968) Aftershocks and microaftershocks. Bull Seismol Soc Am 58:1131–1168

Parzen E (1962) On estimation of a probability density function and mode. Ann Math Stat 33:1065–1076

Robson DS, Whitlock JH (1964) Estimation of a truncation point. Biometrika 51(1–2):33–39. doi:10.1093/biomet/51.1-2.33

Scholz C (1968) The frequency-magnitude relation of microfractiring in rock and its relation to earthquakes. Bull Seism Soc Am 58:399–415

Silverman BW (1986) Density estimation for statistic and data analysis. Chapman and Hall, London, p 175

Stec K (2007) Characteristics of seismic activity of the Upper Silesian Coal Basin in Poland. Geophys J Int 168(2):757–768

Trifu CI, Urbancic TI, Young RP (1993) Non-similar frequency-magnitude distribution for m < 1 seismicity. Geophys Res Lett 20:427–430

Urban P, Lasocki S, Blascheck P, do Nascimento AF, Giang NV, Kwiatek G (2016) Violations of Gutenberg–Richter in anthropogenic seismicity. Pure Appl Geophys 173:1517–1537

Utsu T (1966) A statistical significance test of the difference in b-value between two earthquake groups. J Phys Earth 14:37–40

Utsu T (1999) Representation and analysis of the earthquake size distribution: a historical review and some new approaches. Pure Appl Geophys 155:509–535