Archives of Mining Sciences **48**, 3 (2003) 315–337

JERZY KORNOWSKI*

LINEAR PREDICTION OF HOURLY AGGREGATED AE AND TREMORS ENERGY EMITTED FROM A LONGWALL AND ITS PERFORMANCE IN PRACTICE

LINIOWA PROGNOZA EMITOWANEJ ZE ŚCIANY, GODZINOWEJ SUMARYCZNEJ ENERGII ZDARZEŃ AE I WSTRZĄSÓW ORAZ JAKOŚĆ PROGNOZY W PRAKTYCE

Possibilities and achievable results of some variants of the linear prediction method have been studied with four long time series of hourly, logarithmic, total (AE + tremors) energy as observed with AE and seismic networks at two mining longwalls. Energies of AE and seismic events from the same space region and time interval (of one hour) have been aggregated resulting in evidently autocorrelated, predictable (to a degree) time series. Two of the series have been observed at the not very hazardous longwall and the two others — at a very hazardous one. It is argued that, contrary to difficulties encountered while predicting parameters (eg. time, place and magnitude) of seismic events, prediction of time series of total (AE + tremors) energy emitted from the observed longwall at constant time intervals is well-defined, simple to predict in the probabilistic sense and can be useful in practice. The linear prediction method allows to predict at any discrete time moment t_i, the mean value $\overline{E}(t_{i+1})$ and variance $\sigma^2_{\overline{E}}(t_{i+1})$ of the energy which will be emitted at the observation region during the nearest time unit (e.g. during the nearest hour). Since the prediction error distribution can be (for logarithmic energy data) approximated with the normal distribution, confidence intervals for prediction and probabilities of any ("safety" or "alarm") threshold exceedance can be easily estimated allowing for the formal hazard assessment.

Up to (approximately) ten-fold reduction of prediction error variance and threefold reduction of confidence intervals for prediction have been obtained, comparing to prediction which takes benefit of tremors only (i.e. neglecting AE) data.

Key words: mining-induced seismic hazard, prediction, time series, seismoacoustics/acoustic emission

W polskim górnictwie węgla kamiennego kopalnie — do oceny indukowanego działalnością górniczą zagrożenia sejsmicznego — obligatoryjnie stosują (m.in.) metody: sejsmologiczną i sejsmo-akustyczną (dalej: AE) i wyposażone są w odpowiednie sieci czujników, lecz otrzymywane wyniki bywają niezadowalające. Co gorsza — jeśli poszukiwana jest prognoza w postaci „miejsca, czasu i energii najbliższego wstrząsu" — to zagadnienie, w kontekście teorii procesów stochastycznych,

* GŁÓWNY INSTYTUT GÓRNICTWA, 40-166 KATOWICE, PLAC GWARKÓW 1

okazuje się źle zdefiniowane. Dobrze zdefiniowane i niezbyt trudne jest jednak zagadnienie zbliżone, lecz skromniejsze: jeśli zainteresowanie ograniczymy do otoczonego wyrobiskami, obserwowanego fragmentu pokładu oraz do ustalonej, najbliższej jednostki Δt czasu (np. godziny) — zatem gdy miejsce i czas są ustalone — to zadanie sprowadza się do prognozy wartości szeregu czasowego energii $E(i \cdot \Delta t)$, $i = 1, 2, \ldots$. Zagadnienie liniowej prognozy ma w tym przypadku znane, optymalne (tzn. minimalizujące średni kwadratowy błąd prognozy) rozwiązanie, w postaci ścisłej zwane prognozą wienerowską. Jakość prognozy zależy od wartości przyjmowanych przez funkcję autokorelacji szeregu czasowego i bywa, że nawet optymalna prognoza (np. szumu białego) nie jest wystarczająco dobra dla zastosowań. Tak właśnie zwykle dzieje się w przypadku (zarówno indukowanych, jak i naturalnych) ciągów wstrząsów: autokorelacja szeregu utworzonego z energii wstrząsów (pomijając AE) najczęściej jest znikoma, a prognoza — niezbyt użyteczna.

W pracy tej, argumentując że podział emisji na AE i wstrząsy jest nieuzasadniony — wprowadza się pojęcie ciągłego procesu (całkowitej, bez podziału na AE i wstrząsy) emisji (dalej: CPE) i jego energię E_T (lub energię logarytmiczną, E_{LT}) w ustalonych przedziałach Δt wykorzystuje się do zdefiniowania i prognozowania zagrożenia sejsmicznego.

W literaturze przedmiotu na licznych przykładach wykazano, że funkcja autokorelacji szeregów E_{LT} przyjmuje zwykle wartości rokujące użyteczną liniową prognozę, oraz że błędy tej prognozy mogą być w pierwszym przybliżeniu aproksymowane rozkładem normalnym. Można więc stosować znane algorytmy liniowej prognozy, estymując co Δt wartość średnią $\overline{E}_{LT}(t_{i+1})$ oraz wariancję $\sigma^2_E(t_{i+1})$ energii, która wyemitowana będzie w najbliższym okresie Δt — co umożliwia też obliczenie formalnie zdefiniowanego zagrożenia sejsmicznego oraz przedziałów ufności dla prognozowanej energii, jest więc formalnie poprawnym rozwiązaniem zadania prognozy (i energii i zagrożenia). W praktyce, co okres LS (zwany krokiem okna, np. $LS = \Delta t$ lub $LS = 24\ \Delta t$), korzystając z obserwacji w oknie czasowym o szerokości $NW \cdot \Delta t$, estymuje się parametry $(a_i, i = 1, \ldots, K$, gdzie K jest rzędem predyktora liniowego) predyktora i przez kolejne LS jednostek Δt ustalony predyktor stosowany jest do prognozy zgodnie z równaniem (4). Po upływie okresu LS okno ulega przesunięciu i procedura jest powtarzana. Ponieważ prognoza liniowa jest metodą statystyczną, jej jakość określa się parametrami V_N (5a) lub Q_N (5b), zdefiniowanymi jako wariancja błędu predykcji znormalizowana względem wariancji — obserwacji w przypadku V_N oraz ciągu wstrząsów w przypadku Q_N (zatem Q_N porównuje jakość prognozy CPE z jakością prognozy „sejsmologicznej"). Ponieważ istnieją różne sposoby estymacji parametrów predyktora (którego rząd K może być narzucony apriori lub estymowany) i uzasadnione mogą być różne wartości parametrów LS i NW określających okno i jego przemieszczenie — w rozdziale 4 porównano jakość wyników prognozy (określoną wskaźnikiem jakości V_N), gdy stosowane są różne warianty metody i różne okna czasowe.

Do prognozy wykorzystano 4 siedmiotygodniowe szeregi $E_{LT}(i)$ $i = 1, \ldots, 1176$, po dwa ze słabo zagrożonej ściany 331 w ZG „Piekary" i z silnie zagrożonej ściany 37 w kopalni „Wesoła". Wyniki (czyli wartości V_N) podano w tablicach 1–4, a graficzne przykłady rezultatów prognozy — na rysunkach 3–6.

Analiza porównawcza tych wyników wskazuje, że:

- nie istnieje „zawsze najlepszy" wariant metody ani okna,
- predyktory rzędu $3 \leq K \leq 5$ oraz predyktory, których rząd określano adaptacyjnie stosując odpowiednie kryterium, prowadzą do wyników zbliżonych, bliskich najlepszym,
- stosowanie okien czasowych dłuższych od 3 tygodni nie jest celowe,
- we wszystkich (czterech) przypadkach możliwe było osiągnięcie wariancji błędu prognozy około 2 razy mniejszej od wariancji obserwacji i od dwu (w przypadku ściany 331, ZG Piekary) do dziesięciu (w przypadku ściany 37, KWK „Wesoła") razy mniejszej od wariancji szeregu czasowego wstrząsów: oznacza to, że energia CPE jest znacznie lepiej prognozowalna od energii ciągu wstrząsów.

Można więc podsumować pracę wnioskiem, że liniowa prognoza parametrów rozkładu energii CPE (zatem i zagrożenia sejsmicznego) jest zawsze dobrze zdefiniowana i łatwa, oraz że jej praktyczne wyniki są lepsze — od wyników prognoz formułowanych na podstawie energii samych tylko wstrząsów — w stopniu uzasadniającym stosowanie w górnictwie.

Słowa kluczowe: indukowane zagrożenie sejsmiczne, prognoza, szeregi czasowe, sejsmoakustyka

1. Introduction

Prediction of mining-induced seismic ("rockburst") hazard is an important, so far unsolved problem. At the Polish coal mines, methods of (mining) seismology — as defined by Dubiński et al. (1996) — and seismoacoustics [known also as microseismic or acoustic emission (AE) method], defined by Kornowski et al. (1996), are used for rockburst hazard assessment and all the endangered coal mines are equipped with continuously working, mutually independent seismic and AE networks.

Seismic and AE events are observed in different frequency bands (seismic events inside the 0.5–50 Hz band, AE inside the 100–1200 Hz band) and classified according to their approximate energy ($\boldsymbol{E}$): if $E > 10^2$ J they are named "seismic", if $E < 10^0$ J, they are "AE events", the intermediate energy band, $10^0 < E < 10^2$ J is partly observed by the both networks. Only energies of seismic events are estimated to any degree of accuracy (it is believed that they are correct "to the order of magnitude"). Observing AE, only the spatial density of the AE wavefield energy nearby sensor is measured (and then rescaled taking into account the distance from the sensor to the observed longwall) and this is called "the energy of AE events". Due to abundance of AE events, we hope that their energy aggregated in 1-hour intervals is not very erroneous, and this — hourly aggregated — AE energy makes time series $\{\boldsymbol{E}_{AE}(t_i), i = 1, 2, ...\}$ called "AE (hourly) energy".

Sequences of seismic events are usually (e.g. Vere-Jones 1970; Lasocki 1995) treated as marked point processes, but they can always be transformed into time series of seismic-only energy $\boldsymbol{E}_s(t_i)$.

Usually miners, like all the people exposed to seismic hazard, would like to know, some time before the event, its simplest parameters (e.g. the time, place and energy) and such an information would be readily called "the (classical or proper) prediction". But no rational method of prediction in this sense has been found. Lomnitz (1994, p. 3) has noticed: "Earthquake prediction (in the sense of forecasting the date, location and magnitude) is not feasible today" and this diagnosis can be extended to rockbursts. Even worse, there are fundamental problems with ordering in multidimensional spaces [Kagan (1997, p. 506), invoking Vere-Jones (1995), clarifies: "...the multidimensional nature of earthquake process makes the definition of the next event impossible ...unless the process is made 1-D"] and with the non-additivity of different physical quantities (e.g. time and energy), so that there are serious doubts if the classical prediction of seismic events is possible at all, inside the stochastic point process paradigm (Geller 1997; Kagan 1997, 1999; Wyss and Dmowska 1997; Kornowski 2002a).

But there are no fundamental problems with prediction of 1-D time series: quality of results depends, of course, on various factors (e.g. autocorrelation function values), the results can be more or less useful in practice, but they are always well-defined and estimable.

Method of linear prediction is known since the time of Kolmogoroff and Wiener (see e.g. Robinson 1967) and has been put into trial in mining seismology too (e.g. Rudajev

et al. 1985) – with different, than used here, sort of data – but the results have not found miners acceptance.

The purpose of this paper is:

- **to introduce the concept of time series of the total (as opposed to seismic-only or AE-only) emission energy as a carrier of information concerning the well-defined and partly predictable seismic hazard;**
- **to discuss — and illustrate with examples — the achievable prediction quality, when using the optimal linear predictors of various orders.**

The data used here are seismic and AE hourly observations (always lasting 1176 hours = 7 weeks) from the longwall 37/501 in the Wesoła Coal Mine (henceforth: WCM and 37/WCM) and from the longwall 331/510 in the Piekary Coal Mine (henceforth: ZGP and 331/ZGP abbreviating the genuine name of Zakład Górniczy Piekary).

When the notion of **seismic hazard** is used, we mean always the mining-induced seismic hazard in coal mines formally defined in the next chapter. **Energy** means here the energy of the far field of seismic waves, like in seismology. **Linear predictor and autoregressive model** are treated as synonyms and used exchangeably.

This paper is a continuation and widening of (Kornowski 2002b) article, where observations from sensors mounted in seam and in the rock strata above it have been compared and their usefulness for various variants of linear prediction analysed.

2. Useful notions and definitions

Trying to predict the seismic hazard with the method of linear prediction, it seems natural to build a time series of seismic (e.g. daily — see Rudajev et al. 1985) energies $E_s(t_i)$ and then look — in it — for information allowing prediction. The well known in seismology problem is then, the negligible — or the lack of — autocorrelation [Lasocki (1995 p. 187) observed: "Distribution of tremors... is the Poisson distribution" (transl. J.K.); see also Kornowski (2002a ch. 2.5)], so that even the optimal prediction can be not very useful.

Here a different point of view is presented. There is no physical reason, in this author's opinion, to discriminate events into seismic ones and AE. There is no physical border between them. The discrimination is caused by psychological ("who is afraid of AE?") and technical (independent networks) reasons. The gaps — in the frequency and energy domains — mentioned earlier, deepen the suggestion of the two independent processes, but this is an appearance only. **Given the sufficiently small space segment V, there is only one stochastic emission process,** driven by the local stress to local critical stress ratio. This opinion Kornowski (2002a) has called **the unity of emission postulate.** One realization of the process is usually observable; its sharp maxima in the energy domain are called events, sometimes — tremors.

The process may be seen as discrete-time or continuous-time depending on the observation mode and the detection level. It is known (e.g. Leadbetter et al. 1983) that the maxima of exceedances (above a threshold) of a stochastic process tend,

with growing threshold level, to a Poisson process — and this happens almost exactly in case of AE and seismic observations.

Just like in case of any other stochastic process, **the subject of prediction is the process, not a particular its realization, so that "prediction" means always the prediction of the probability density of the "instantaneous" amplitude of the process.** The process we are interested in is called (Kornowski 2002a) **the continuous process of emission (CPE)**. It can be sampled at any time intervals, but in the mining industry AE data are interpreted hourly — this is arbitrary but commonly applied at the Polish coal mines — and this will be the our time unit.

Discrimination into AE and tremors is not only artificial and physically unfounded: in strict sense, dividing the total emission into continuous AE process and the Poisson point process of tremors undermines the very possibility of linear prediction of tremors from their own past observations and — creating the two, by definition mutually uncorrelated processes (point process of tremors and continuous process of AE) excludes any linear relations between AE and tremors. Being harmful and unfounded, the discrimination should be abandoned (and suitable — wideband in the frequency and energy domains — observing system should be implemented).

From now on the name of **total emission** and abbreviation **CPE** [of "continuous process of (total) emission"] will be used, meaning all the events above the noise level and its energy will be denoted $E_T(t)$ with the index "T" meaning "total". In the current practice, with the two independent networks (AE and seismic ones), $E_T(t)$ cannot be observed: it can only be estimated as the sum of the two, (seismic) $E_s(t)$ and (seismoacoustic) $E_{AE}(t)$ energies at the t_i $i = 1, 2, \ldots$ moments

$$E^*_T(t_i) = E_s(t_i) + E_{AE}(t_i) \tag{1}$$

where asterisk, *, denotes estimate.

Due to mentioned earlier gaps in — and partial doubling of — AE and seismic observations, serious differences between $E_T(t_i)$ and $E^*_T(t_i)$ are possible, most importantly the correlational relations inside the so estimated $E^*_T(t_i)$ process are distorted: it should be remembered, that despite the our way of estimation of the $E^*_T(t_i)$, **the genuine $E_T(t)$ process is not a sum of two independent processes:** it is just the energy of the CPE.

In this work input time series for prediction algorithms are obtained aggregating AE and seismic energies from the same space region (V) in hourly intervals according to (1) and then taking logarithm:

$$E_{LT}(t_i) = \log\{E^*_T(t_i) + 1\} \tag{2}$$

where L means "logarithmic" and 1 is added to avoid problems of zero energy. **Values of $E_{LT}(t_i)$ are used as the input data for linear prediction algorithms. It has been empirically found** (e.g. Kornowski 2002a; Kornowski and Kurzeja 2000) **that time series of $E_{LT}(t_i)$ are significantly autocorrelated** making linear prediction useful — and a few examples are shown in the chapter 4.

Note — what is important and obvious — that the autocorrelation function (acf) of $E^*_T(t_i)$ tends to the acf $[E_{AE}(t_i)]$ if the acf $[E_s(t_i)] \to 0$, but this is a feature of the $E^*_T(t_i)$ only, not of the $E_T(t_i)$: inside the genuine $E_T(t_i)$ process, any discrimination into AE and tremors is physically meaningless.

Partly due to logarithmic transformation, distribution of prediction errors [when predicting $E_{LT}(t_{i+1})$] — not being strictly normal — can be usefully approximated by the normal distribution. **By prediction at a time t_i, given observations at t_i, t_{i-1}, ..., we mean here the estimation of the most probable (mean) value $\overline{E}_{LT}(t_{i+1})$ and variance $\sigma^2_E(t_{i+1})$.** These two parameters fully characterize a normal random variable (and approximately characterize prediction errors in our case), enabling us to estimate (e.g.) confidence intervals for prediction.

Now let us define the seismic hazard:

The seismic hazard, $Z_{TV}(t_i, E_g)$ is the probability of exceedance — by the total energy $E_T(t_i)$ emitted from a time and space region $[(t_i, t_{i+1}), V] \equiv (T, V)$ — of a predetermined "safety threshold" E_g

$$0 \le Z_{TV}(t_i, E_g) \equiv P[E_T(t_i) > E_g]_V \le 1 \tag{3}$$

(nothing changes if E_T and E_g are logarithmed).

This, or similar definition is known in the literature (eg. Gibowicz and Kijko 1994; Marcak and Zuberek 1994) but using seismic-only energy, E_s. Threshold E_g may be a constant scalar (e.g. $E_g = 1 \cdot 10^5$ J) or a vector (e.g. E^a_g, E^b_g, ...) dividing the hazard space (0, 1) into a few "hazard states" or "alarm levels". It should be stressed that:

a) the total energy, $E_T(t)$ is "the carrier" of so defined seismic hazard. When (T, V) decreases to a point, $E_T(t)$ degenerates to zero or to the energy of an event if it happened inside the (T, V);
b) under normal approximation, prediction of $\overline{E}(t_{i+1})$ and $\sigma^2_E(t_{i+1})$ allows to estimate the hazard easily;
c) prediction should never be limited to the $\overline{E}_T(t_i)$ only — this would prove false almost surely. Much better is a cautious sentence, e.g. **"with the probability** (e.g.) **P = 90%, in the region (T, V), $E_{5\%} < E_T(t_{i+1}) < E_{95\%}$ and the probability that $E_T(t_{i+1}) > E_g$, equals Z"** (with numerical values of $E_{5\%}$, $E_{95\%}$, E_g and Z);
d) hazard prediction is not sufficient to take decisions. It should be complemented with the decision rule(s), e.g.

$$\text{if } \{Z > \text{threshold}(j)\}, \text{ then } \{\text{action } (j)\}$$

but the details are beyond the scope of this paper.

3. Linear prediction

The linear prediction method [also known as "autoregressive (AR) modelling" or "(approximate) Wiener filtering"] with its many variants, has long been known (e.g. Robinson 1967; Robinson and Treitel 1980; Rudajev et al. 1985; simple FORTRAN

programs can be found in Silvia and Robinson 1979) and widely used not only for prediction (e.g. Markel and Gray 1976), so that there is no need to discuss its foundations here. It seems enough to remind, that the method allows to estimate, optimally in the mean square sense, the values of autoregressive model, AR(K) — equivalently: of linear predictor of order K — coefficients $(a_j, j = 1, ..., K)$

$$\bar{x}(t+1) = a_1 x(t) + ... + a_k x(t - K + 1) \tag{4}$$

where $\bar{x}(t + 1)$ is the most probable (mean) value of the time series at $(t + 1)$. The same linear prediction algorithm usually allows to estimate the variance $\sigma^2_E(t_{i+1})$ of prediction errors.

From the equation (4) it follows that $\bar{x}(t + 1)$ is a simple linear function of some previous observations, justifying the name of the method. What is more important, it can be inferred that **any errors or inaccuracies of observations are immediately transmitted to predictions so that, on the average, predictions can not be more accurate than observations are. This is really important for seismic hazard prediction!**

If the best model order (K) is unknown (as usually is), the method should be equipped with order-selecting criterion (e.g. Akaike 1974; Broersen and Wensink 1998).

Because there are many apriori reasonable model orders (K), many order-selecting criteria and many algorithms estimating prediction parameters (e.g. Kalouptsidis et al. 1985; Strobach 1990), there exist many competing variants of linear prediction, generating more or less different results. What is more, time series of $E_{LT}(t_i)$ are not strictly stationary so that neither fixed nor steadily growing time window (where the predictor parameters are estimated) is the best in all conditions — and many different data windows can be used. Some of these possibilities have been investigated in our experiment and the results are reported in chapter 4.

In actual calculations:

- the Levinson-Durbin algorithm (e.g. Strobach 1990) has been used to estimate the predictor parameters $(a_j, j = 1, ..., K)$ from the autocorrelation function;
- the Kalouptsidis et al. (1985) [known also as a (modified) forward–backward least–squares] algorithm has been used to estimate predictor coefficients from observations $E_{LT}(t_i)$;
- in one special case (with $K = 168$, but with only a_1, a_2, a_{168} not equal to zero) elementary "substitution" method has been used to solve (3×3) normal equations;
- estimating value of K, the best order of predictor, criterion of Broersen and Wensink (1998) has been used;
- rectangular data windows of widths: 1, 2 or 3 weeks (168, 336 or 504 hours), moving with steps of 0 (i.e. unmovable), 1, 24 or 168 hours have been used.

In all the cases observations $E_{LT}(t_i)$ have been incoming sequentially, hour after hour, and in the same manner predictions $[\bar{E}_{LT}(t_{i+1}), \sigma^2_E(t_{i+1})]$ have been estimated, so that predictor — or AR(K) model — always has been estimated "inside the observed data set" (or: inside the current data window) **and then used to predict future observation parameters "outside the data set"** (or: outside the current data window).

It has been checked, at an earlier stage of the experiment, that:

- prediction results, obtained using the data with the mean value removed from the current data window, are usually slightly (0–1%) better than results obtained with unmodified data,
- the so called biased autocorrelation estimator (e.g. Jenkins and Watts 1968 ch. 5.3) allows to obtain slightly (0–1%) better results than the unbiased one.

These facts are known in the estimation theory and have been expected. Only the results obtained using the biased autocorrelation estimator and the data with the mean value removed from the current data window are reported in the chapter 4.

Given a time series x_t, [e.g. $E_{LT}(t_i)$], of observations, two **quality of prediction measures** (or: **quality indices**) — a particular measure V_N and a more general one, Q_N — are comonly accepted (it is a convention only: in fact measures $1 - V_N$ and $1 - Q_N$ better deserve the name):

a) observation-normalized variance of prediction errors

$$V_N = \sigma_p^2 / \sigma_x^2 \tag{5a}$$

where σ_p^2 is the prediction errors variance and σ_x^2 is the observations (x_t) variance;

b) benchmark–normalized variance of prediction errors

$$Q_N = \sigma_p^2 / \sigma_b^2 \tag{5b}$$

where σ_b^2 is a benchmark prediction error variance. If (e.g.) seismic-only time series is used as the benchmark (ignoring AE data), σ_b^2 is the variance of the tremors sequence from the same time and space region (T, V) as the x_t data.

From (5a) and (5b) we note also that

$$Q_N = V_N \, \sigma_x^2 / \sigma_b^2 \tag{5c}$$

For uncorrelated x_t data, $V_N \equiv 1$ and linear prediction gives us no information above this contained in the ("unconditional") distribution of observations: prediction is useless in the case. Prediction is called "usefull" if $0 < V_N < 1$, the smaller V_N is, the better. It should be remembered especially in applications, that **the variance is a statistical parameter, to be calculated with "sufficiently large" data set, so that evaluating the prediction performance with a single, or a few — even very important, like a rockburst in a mine — observations is a plain error.** This is the price to be paid for prediction in a stochastic environement (e.g. in actual rock strata), where only statistical data description is useful and possible.

4. Input data and prediction results

Two longwalls, at two — of very different mining and geological conditions — coal mines have been observed with seismic and AE networks: not too seismically hazardous 331/ZGP and heavily hazardous 37/WCM.

At 331/ZGP two geophones, designated as G11 and G23 — and at 37/WCM two sensors (accelerometers with an integrating circuit): A9 and A11 — have been, apart from seismic networks, the sources of data during the 7 weeks, or 1176 hours, of continuous observations. Four time series: $E_{LT}^{G11}(t_i)$, $E_{LT}^{G23}(t_i)$, $E_{LT}^{A9}(t_i)$ and $E_{LT}^{A11}(t_i)$ (with upper index pointing the sensor) for $i = 1, ..., 1176$, have been observed and calculated according to equation (2) and used as the input data for prediction. Mean values and variances of the time series respectively, have been: (1.97, 0.59), (1.92, 0.70), (3.83, 0.51), (3.44, 0.64). Some of the data and their simplest statistical characteristics (henceforth: "statistics") are shown at Fig. 1 and Fig. 2. Figures are divided into "lines": the uppermost line is always called "the first line" and so on, from the top to the bottom of a figure.

Fig. 1, with seismic and G11-331/ZGP data, and Fig. 2 with seismic and A9-37/WCM data, are identically constructed: in the first line the seismic-only hourly energies,

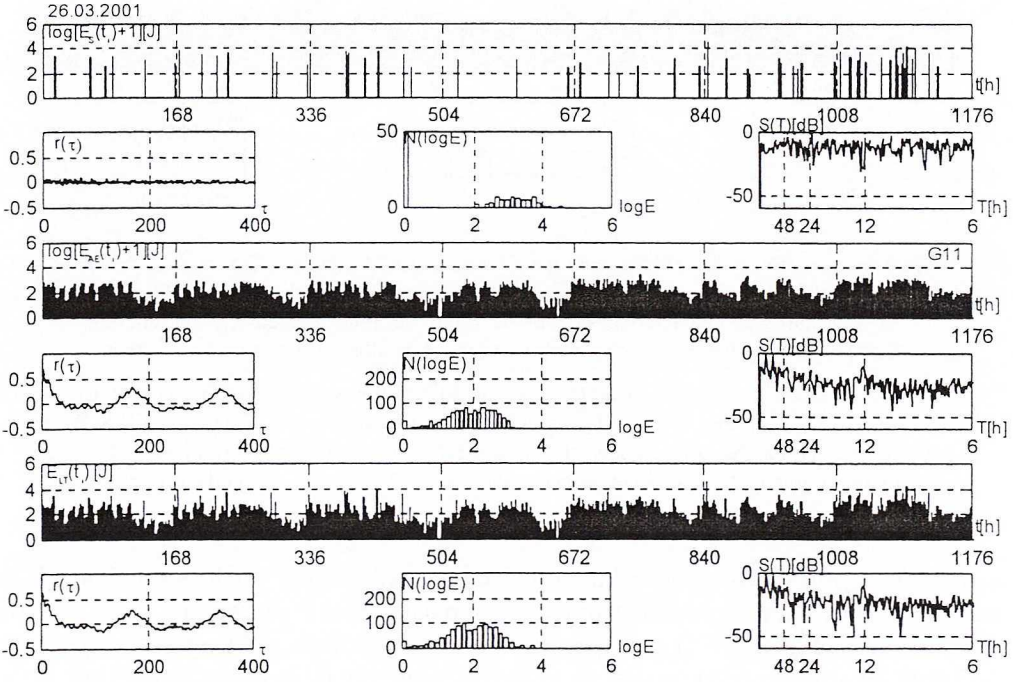

Fig. 1. Time series of hourly logarithmic energies of seismic and AE events {from the top to the bottom: $\log[E_s(t_i) + 1]$, $\log[E_{AE}(t_i) + 1]$, $E_{LT}(t_i)$}, observed at the longwall 331 of the ZGP and (below any time series) its statistical characteristics (from the left to the right): autocorrelation function, energy-frequency histogram, spectrum. For more detailed information see the text

Rys. 1. Szeregi czasowe godzinowej logarytmicznej energii zdarzeń sejsmicznych i sejsmoakustycznych {od góry: $\log[E_s(t_i) + 1]$, $\log[E_{AE}(t_i) + 1]$, $E_{LT}(t_i)$}, obserwowane w rejonie ściany 331 ZG Piekary oraz (pod każdym z szeregów czasowych) ich charakterystyki statystyczne (od lewej): funkcja autokorelacji, histogram rozkładu energii, widmo. Dokładniejszy opis podano w tekście

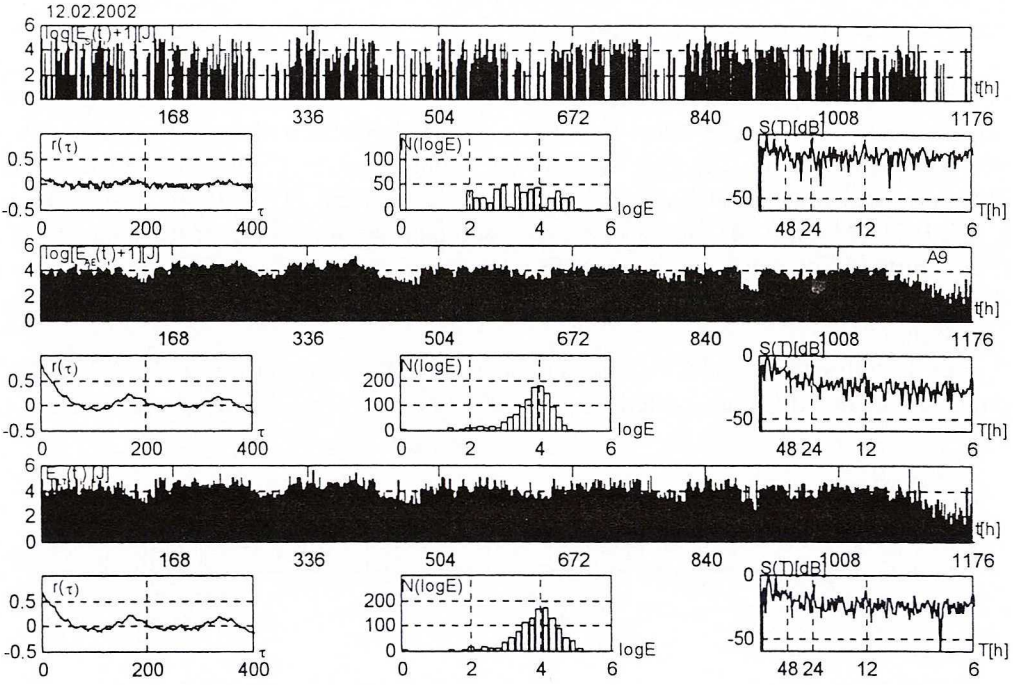

Fig. 2. Time series of hourly logarithmic energies of seismic and AE events {from the top to the bottom: $\log[E_s(t_i) + 1]$, $\log[E_{AE}(t_i) + 1]$, $E_{LT}(t_i)$}, observed at the longwall 37 of the WCM and (below any time series) its statistical characteristics (from the left to the right): autocorrelation function, energy-frequency histogram, spectrum. For more detailed information see the text

Rys. 2. Szeregi czasowe godzinowej logarytmicznej energii zdarzeń sejsmicznych i sejsmoakustycznych {od góry: $\log[E_s(t_i) + 1]$, $\log[E_{AE}(t_i) + 1]$, $E_{LT}(t_i)$}, obserwowane w rejonie ściany 37 KWK Wesoła oraz (pod każdym z szeregów czasowych) ich charakterystyki statystyczne (od lewej): funkcja autokorelacji, histogram rozkładu energii, widmo. Dokładniejszy opis podano w tekście

$\log[E_s(t_i) + 1]$ are shown with the horizontal axis scaled in hours. In the second line there are three frames showing statistics of the seismic data, from the left to the right they are: autocorrelation function $r(\tau)$, hourly energy-frequency histogram $N(\log E)$ and spectrum $S(T)$ with logarithmic (dB) vertical scale and horizontal scale in hours (not in Hz). In the third line only AE data, $\log [E_{AE}(t_i)]$ are shown and below them, in the fourth line there are–like in the second line–autocorrelation, hourly energy-frequency histogram and the spectrum of AE data. In the fifth line the values of $E_{LT}(t_i)$ are visible [note that they are not the sum of lines 1 and 3; energies E_s and E_{AE} are sumed up first-and then logarithmed] and these are the input data for our predicting algorithms. The last, sixth line, shows (as before) statistics of $E_{LT}(t_i)$.

Time series $E_{LT}^{G23}(t_i)$ and $E_{LT}^{A11}(t_i)$ bring no new information and need not to be shown. The input data are worthy of some comments:

a1. The time series of seismic-only data from the 331/ZGP (Fig. 1, first line) is almost unautocorrelated, with $r(\tau)$ generally below the 5% significance limit (equal to $1.96/\sqrt{1176} \approx 0.06$) and the spectrum is almost flat except of small peak at 24 h. This time series cannot be usefully predicted (i.e. $V_N = 1$) from its own past.

a2. The time series of seismic-only data from the hazardous 37/WCM (Fig. 2, first line) is slightly autocorrelated and spectral peaks at 12 hours (as a result of the 4-shift production regime), 24 hours and 168 hours are visible (V_N is — slightly — less than 1). **This means that seismic events, as defined in the chapter 1, are influenced to a degree by production process.** Very likely, that the smaller the lower tremor-defining threshold (here: $1 \cdot 10^2$ J), the stronger this influence is.

a3. Energy-frequency histogram of seismic-only data is important, showing the effects of doubly-censoring: seismic events are censored from below (at 10^2 J) due to the definition of seismic event and — as a result of some formal regulations — from above (at 10^5 J). The upper limit is not sharp. **The most important piece of data is plainly destroyed.**

a4. In the case of both coal mines, energy-frequency histograms of seismic-only data are far from normal distribution.

a5. AE hourly data (Fig. 1 and Fig. 2, third line) are evidently autocorrelated with periods of 12, 24 and 168 hours visible mainly in their spectra. **So, AE is influenced by the same production process as the seismic-only emission and we conclude that both AE nad seismic–only emission are parts of the same CPE process.**

a6. $E_{LT}(t_i)$ is autocorrelated and — except of high-energy far end — distributed like the relevant AE, as a result of equation (1) [but it should be repeated that $E_T^*(t) \neq E_T(t)$ and the acf of the genuine $E_T(t)$ is not a sum of AE and seismic acf's].

a7. Hourly energies of AE, rescaled and expressed in the same as E_s physical units (J) can (but need not) be of the same order of magnitude as the typical seismic events: for example (Fig. 2, third line), the hourly energies of AE from A9-37/WCM, oscillate near the level of 4 logarithmic energy units (or 10^4 J/hour).

The performance of 9 predictors [i.e. autoregressive models (4) of various, fixed or variable order] has been analysed and the resulting values of V_N, observation-normalized variance of prediction errors, are shown in tables: Table 1–4 for, respectively, time series of $E_{LT}^{G11}(t_i)$, $E_{LT}^{G23}(t_i)$, $E_{LT}^{A9}(t_i)$ and $E_{LT}^{A11}(t_i)$. Denoting by m_k models of the order K with parameters a_j, $j = 1, \ldots, K$ estimated from the relevant autocorrelation function and by M_k models with parameters $b_j, j = 1, \ldots, K$ estimated from observations $E_{LT}(t_i)$ (and by η_t a random error) they are:

$$m_1: \qquad x_t = a_1 x_{t-1} + \eta_t \tag{6a}$$

$$m_2: \qquad x_t = a_1 x_{t-1} + a_2 x_{t-2} + \eta_t \tag{6b}$$

$$m_3: \qquad x_t = a_1 x_{t-1} + a_2 x_{t-2} + a_3 x_{t-3} + \eta_t \tag{6c}$$

$$m'_{168} \qquad x_t = a_1 x_{t-1} + a_2 x_{t-2} + a_{168} x_{t-168} + \eta_t \tag{6d}$$

$$M_k: \qquad x_t = \sum_{j=1}^{K} b_j x_{t-1} + \eta_t, \qquad \mathrm{K} = 3, 4, 5, 12 \tag{6e}$$

$$M_{opt}: \qquad x_t = \sum_{j=1}^{K(t)} b_j x_{t-1} + \eta_t \tag{6f}$$

where $0 \leq K(t) \leq 12$ is the variable model order minimizing FSIC (Finite Sample Information Criterion) of Broersen and Wensink (1998) criterion, which is a modified form of the well-known AIC (Akaike Information Criterion, Akaike 1974) criterion selecting the best autoregressive model/linear predictor order. Models m_3 and M_3

TABLE 1

Quality (V_N) of prediction with 9 linear predictors (6a)–(6f), for various values of the window width (*NW*) and step length (*LS*). Estimated for the time series of 1176, predicted and then observed, hourly values of $E_{LT}(t_i)$ $i = 1, ..., 1176$, with AE observations (with: $\overline{E}_{LT} \approx 1.92$, $\sigma_E^2 \approx 0.70$) from the geophone G23 at the longwall 331/ZGP

TABLICA 1

Wartości wskaźnika V_N jakości prognozy, wyznaczone dla dziewięciu predyktorów liniowych (6a)–(6f), wykorzystując okna obserwacyjne o różnej szerokości *NW*, poruszające się krokiem o długości *LS*. Estymowano na podstawie 1176 kolejno prognozowanych i obserwowanych wartości szeregu czasowego $E_{LT}(t_i)$ $i = 1, ..., 1176$. Obserwacje AE ($\overline{E}_{LT} \approx 1{,}92$, $\sigma_E^2 \approx 0{,}70$) otrzymano za pomocą geofonu G23 z rejonu ściany 331/ZGP

LS [h]	*NW* [h]	m_1	m_2	m_3	m'_{168}	M_3	M_4	M_5	M_{12}	$M_{opt} \leq 12$
1	168	0.535	0.497	0.487	0.498	0.489	0.494	0.499	0.516	0.518
	336	0.537	0.497	0.487	0.482	0.489	0.490	0.492	0.491	0.486
	504	0.534	0.496	0.487	0.477	0.488	0.490	0.491	0.484	0.485
24	168	0.538	0.497	0.487	0.498	0.487	0.494	0.501	0.515	0.505
	336	0.538	0.497	0.486	0.482	0.487	0.489	0.491	0.492	0.487
	504	0.535	0.496	0.487	0.479	0.488	0.491	0.492	0.483	0.483
168	158	0.551	0.504	0.495	0.505	0.495	0.501	0.505	0.532	0.524
	336	0.545	0.503	0.495	0.488	0.494	0.497	0.498	0.497	0.496
	504	0.538	0.498	0.489	0.480	0.489	0.490	0.491	0.480	0.480
0	168	0.535	0.506	0.511	0.509	0.503	0.502	0.500	0.521	0.521
	336	0.539	0.496	0.487	0.478	0.486	0.485	0.488	0.483	0.483
	504	0.541	0.497	0.486	0.483	0.486	0.486	0.484	0.480	0.479

TABLE 2

Quality (V_N) of prediction with 9 linear predictors (6a)–(6f), for various values of the window width (NW) and step length (LS). Estimated for the time series of 1176, predicted and then observed, hourly values of $E_{LT}(t_i)$ $i = 1, ..., 1176$, with AE observations (with: $\overline{E}_{LT} \approx 1.97$, $\sigma_E^2 \approx 0.59$) from the geophone G11 at the longwall 331/ZGP

TABLICA 2

Wartości wskaźnika V_N jakości prognozy, wyznaczone dla dziewięciu predyktorów liniowych (6a)–(6f), wykorzystując okna obserwacyjne o różnej szerokości NW, poruszające się krokiem o długości LS. Estymowano na podstawie 1176 kolejno prognozowanych i obserwowanych wartości szeregu czasowego $E_{LT}(t_i)$, $i = 1, ..., 1176$. Obserwacje AE ($\overline{E}_{LT} \approx 1{,}97$, $\sigma_E^2 \approx 0{,}59$) otrzymano za pomocą geofonu G11 z rejonu ściany 331/ZGP

LS [h]	NW [h]	m_1	m_2	m_3	m'_{168}	M_3	M_4	M_5	M_{12}	$M_{opt} \leq 12$
1	168	0.605	0.580	0.567	0.581	0.570	0.565	0.563	0.582	0.572
	336	0.606	0.578	0.563	0.560	0.564	0.557	0.556	0.559	0.560
	504	0.603	0.574	0.558	0.550	0.559	0.554	0.554	0.553	0.552
24	168	0.609	0.579	0.565	0.581	0.566	0.558	0.555	0.576	0.582
	336	0.607	0.578	0.563	0.562	0.563	0.556	0.554	0.558	0.566
	504	0.605	0.574	0.557	0.551	0.557	0.551	0.553	0.553	0.552
168	168	0.614	0.583	0.563	0.585	0.564	0.559	0.554	0.568	0.557
	336	0.608	0.576	0.560	0.556	0.560	0.555	0.556	0.555	0.554
	504	0.604	0.572	0.556	0.548	0.556	0.550	0.552	0.548	0.549
0	168	0.610	0.578	0.559	0.579	0.558	0.557	0.563	0.589	0.577
	336	0.637	0.591	0.570	0.563	0.568	0.560	0.558	0.564	0.562
	504	0.631	0.589	0.566	0.565	0.566	0.559	0.556	0.563	0.560

have been included to enable comparison of correlation-based and observation-based algorithms: because we do not know the true autocorrelation values (only their estimates) it was believed that observation-based predictors are better. Model m'_{168} has been included to take advantage of the apriori "known" periodicity (of period 168 hours). Alike, M_{12} model was hoped to take advantage of the 12 and 24 hours periodicity visible in some spectra. Model M_{opt}, of order varying with time, has been estimated for any new data window (even with $LS = 1$) so as to minimize the FSIC.

Three window widths (1, 2 and 3 weeks) — denoted NW in the Table 1–4 — as well as four values of the window step-lengths [$LS = 0$, 1, 24, 168 hours, where $LS = 0$ means fixed, unmovable data window placed at $E_{LT}(t_i)$, $i = -NW + 1, -NW + 2, ..., -1, 0$] have been used.

TABLE 3

Quality (V_N) of prediction with 9 linear predictors (6a)–(6f), for various values of the window width (NW) and step length (LS). Estimated for the time series of 1176, predicted and then observed, hourly values of $E_{LT}(t_i)$ $i = 1, ..., 1176$, with AE observations (with: $\overline{E}_{LT} \approx 3.44$, $\sigma_E^2 \approx 0.64$) from the sensor A11 at the longwall 37/WCM

TABLICA 3

Wartości wskaźnika V_N jakości prognozy, wyznaczone dla dziewięciu predyktorów liniowych (6a)–(6f), wykorzystując okna obserwacyjne o różnej szerokości NW, poruszające się krokiem o długości LS. Estymowano na podstawie 1176 kolejno prognozowanych i obserwowanych wartości szeregu czasowego $E_{LT}(t_i)$, $i = 1, ..., 1176$. Obserwacje AE ($\overline{E}_{LT} \approx 3{,}44$, $\sigma_E^2 \approx 0{,}64$) otrzymano za pomocą czujnika A11 z rejonu ściany 37/WCM

LS [h]	NW [h]	m_1	m_2	m_3	m'_{168}	M_3	M_4	M_5	M_{12}	$M_{opt} \leq 12$
1	168	0.622	0.589	0.582	0.590	0.585	0.587	0.589	0.610	0.597
	336	0.647	0.600	0.588	0.587	0.589	0.582	0.585	0.577	0.579
	504	0.666	0.609	0.592	0.591	0.592	0.584	0.585	0.578	0.582
24	168	0.634	0.593	0.584	0.594	0.583	0.581	0.585	0.597	0.593
	336	0.659	0.605	0.594	0.594	0.593	0.584	0.587	0.574	0.588
	504	0.675	0.614	0.596	0.596	0.595	0.585	0.588	0.58	0.580
168	168	0.662	0.609	0.595	0.609	0.597	0.594	0.602	0.596	0.593
	336	0.679	0.617	0.597	0.605	0.598	0.589	0.594	0.585	0.596
	504	0.693	0.622	0.600	0.604	0.600	0.590	0.591	0.582	0.593
0	168	0.752	0.661	0.623	0.661	0.622	0.609	0.609	0.596	0.622
	336	0.669	0.608	0.588	0.609	0.588	0.578	0.582	0.572	0.582
	504	0.663	0.604	0.585	0.593	0.585	0.576	0.576	0.564	0.576

In all the cases the mean values of prediction errors, estimated after finishing the whole experiment, have been so small that the variances (of prediction errors) have not differed — from the respective mean squared errors — up to the fourth decimal place: we conclude than, that the predictors are unbiased (what is known in estimation theory) and the normalized variances may be used as a performance measure.

Analysing, contained in tables Table 1–4 prediction results (as measured with the V_N index), the following conclusions can be formulated:

b1. No single algorithm, model order, window width (NW) or step length (LS), has been the best in all circumstances.

b2.1. Achievable quality of prediction, as measured with V_N, has been, for the best cases in the tables, contained in the interval (0.45–0.57). It also means that the

TABLE 4

Quality (V_N) of prediction with 9 linear predictors (6a)–(6f), for various values of the window width (*NW*) and step length (*LS*). Estimated for the time series of 1176, predicted and then observed, hourly values of $E_{LT}(t_i)$ $i = 1, ..., 1176$, with AE observations (with: $\overline{E}_{LT} \approx 3.83$, $\sigma_E^2 \approx 0.51$) from the sensor A9 at the longwall 37/WCM

TABLICA 4

Wartości wskaźnika V_N jakości prognozy, wyznaczone dla dziewięciu predyktorów liniowych (6a)–(6f), wykorzystując okna obserwacyjne o różnej szerokości *NW*, poruszające się krokiem o długości *LS*. Estymowano na podstawie 1176 kolejno prognozowanych i obserwowanych wartości szeregu czasowego $E_{LT}(t_i)$, $i = 1, ..., 1176$. Obserwacje AE ($\overline{E}_{LT} \approx 3{,}83$, $\sigma_E^2 \approx 0{,}51$) otrzymano za pomocą czujnika A9 z rejonu ściany 37/WCM

LS [h]	*NW* [h]	m_1	m_2	m_3	m'_{168}	M_3	M_4	M_5	M_{12}	$M_{opt} \leq 12$
1	168	0.527	0.494	0.473	0.494	0.481	0.468	0.472	0.493	0.480
	336	0.538	0.499	0.472	0.492	0.476	0.458	0.457	0.466	0.457
	504	0.548	0.505	0.476	0.496	0.478	0.459	0.457	0.462	0.459
24	168	0.55	0.504	0.478	0.504	0.474	0.456	0.46	0.483	0.473
	336	0.555	0.507	0.477	0.500	0.474	0.456	0.454	0.461	0.453
	504	0.562	0.513	0.480	0.503	0.478	0.459	0.457	0.463	0.456
168	168	0.585	0.522	0.487	0.522	0.489	0.464	0.465	0.480	0.474
	336	0.584	0.525	0.486	0.518	0.487	0.467	0.466	0.473	0.473
	504	0.592	0.536	0.494	0.524	0.495	0.476	0.472	0.474	0.472
0	168	0.678	0.593	0.541	0.593	0.536	0.513	0.495	0.488	0.536
	336	0.675	0.644	0.599	0.637	0.598	0.589	0.561	0.528	0.561
	504	0.664	0.611	0.563	0.592	0.560	0.546	0.538	0.514	0.560

prediction uncertainty (equivalently: the confidence interval width for prediction) can be reduced to (67–75%) of its value given the distribution of observations $E_{LT}(t_i)$;

b2.2. Achievable quality of prediction, as measured by Q_N (with σ_s^2 the variance of seismic-only data as a benchmark), is very different for the two coal mines. For very sparse and evidently non-normally distributed seismic-only ZGP data (Fig. 1), $\sigma_s^2 \approx 0.504$. The variance (σ_s^2) of the $E_{LT}^{G23}(t_i)$ equals 0.70 and — in the Table 1 — the minimum value of $V_N \approx 0.477$. Alike, the variance of the $E_{LT}^{G11}(t_i)$ equals 0.59 and minimum of $V_N \approx 0.548$ (Tab. 2) so that, according to (5c), $0.64 < Q_N < 0.66$ for the 331/ZGP. The prediction uncertainty, as measured by the confidence interval (for prediction) width, can be reduced to ~80% of its value

resulting from the distribution of the (331/ZGP) seismic data only. Analogously, for the highly endangered 37/WCM, $\sigma_s^2 \approx 3.1$, for $E_{LT}^{A11}(t_i)$ $\sigma_s^2 \approx 0.64$ with minimum $V_N = 0.564$ (Tab. 3) and for the $E_{LT}^{A9}(t_i)$ $\sigma_s^2 \approx 0.51$ with minimum $V_N = 0.453$ (Tab. 4) so that, according to (5c), $0.075 < Q_N < 0.12$ (for the 37/WCM) and uncertainty (or confidence interval width for prediction) can be reduced to 27–35% of its value resulting from the seismic data only. **This has been obtained due to introducing the CPE concept and applying the linear prediction to the data of high seismic-only component [high value of $\sigma_b^2 \equiv \sigma_s^2$ in (5c)] or equivalently, to the highly hazardous and very active region.** The result illustrates a difference between the seismic-only and the CPE (total: seismic + AE) data prediction; the reverse of the medal is, obviously, the prediction of the CPE, not the tremors.

b3.1. Non-adaptive, constant-in-time predictors (those with $LS = 0$), apriori supposed to be the worst, performed with very mixed results: in two cases (G23-331/ZGP and A11-37/WCM) being the best or almost the best, in case of G11-331/ZGP performing on the average level and in case of A9-37/WCM being the worst. This is, presumably, the effect of stationarity or its lack.

b3.2. The special predictor m'_{168} [equation (6d)] introduced to take advantage of "known periodicity" in the data, proved to be the best or almost the best with the ZGP data (where production was stable), but performed rather poorly with the much less stable data from the WCM. We conclude (what is known) that apriori information can be very useful — but only when it is sure.

b4. Because predictor parameters have always been estimated inside the data window and then used to predict [the incoming value of $E_{LT}(t_{i+1})$] outside the data window, the optimum adaptive-in-time and adaptive-in-order predictor M_{opt} performed always well but has never garranteed to be (for finite, not strictly stationary data) the best.

b5. Inequalities $V_N(m_3) < V_N(m_2) < V_N(m_1)$ have always been observed, so that oversimplified m_1 and m_2 predictors should be avoided.

b6. In all cases, $V_N(m_3) \approx V_N(M_3)$.

b7. In no case predictor of order less than 5 has been the best. It is somewhat surprising and contrary to some previous findings (e.g. Rudajev et al. 1985, have found the best results for shorter predictors, probably due to using daily energies in their experiment, thus eliminating periodicities ≤ 24 hours). The differences between the $V_N(M_4)$ and $V_N(M_5)$ have been irregular and very small.

b8. For a given time series

- predictor m_3 with $LS = 24$ and $NW = 336$, has never been worse by more than 5% from the best one and more than 4,6% from M_{opt} with $LS = 24$ and $NW = 504$,
- predictor M_4 with LS = 24 and $NW = 336$, has never been worse by more than 2% from the best one,
- predictor M_{opt} with $LS = 24$ and $NW = 504$, has never been worse by more than 1% from the best.

Those are experimental results from, admittedly, a few examples. But the observed longwalls have been truly very different and the time series have been long, so the results can be — hopefully — useful at the other mines.

Graphical examples of prediction results are shown at Fig. 3–6, for the time series $E_{LS}^{G11}(t_i)$, $E_{LS}^{A9}(t_i)$, $E_{LS}^{A11}(t_i)$ and $E_{LS}^{G23}(t_i)$ respectively. The results have been obtained with the m_3 predictor with window NW = 336 moving with steps LS = 24 (hours).

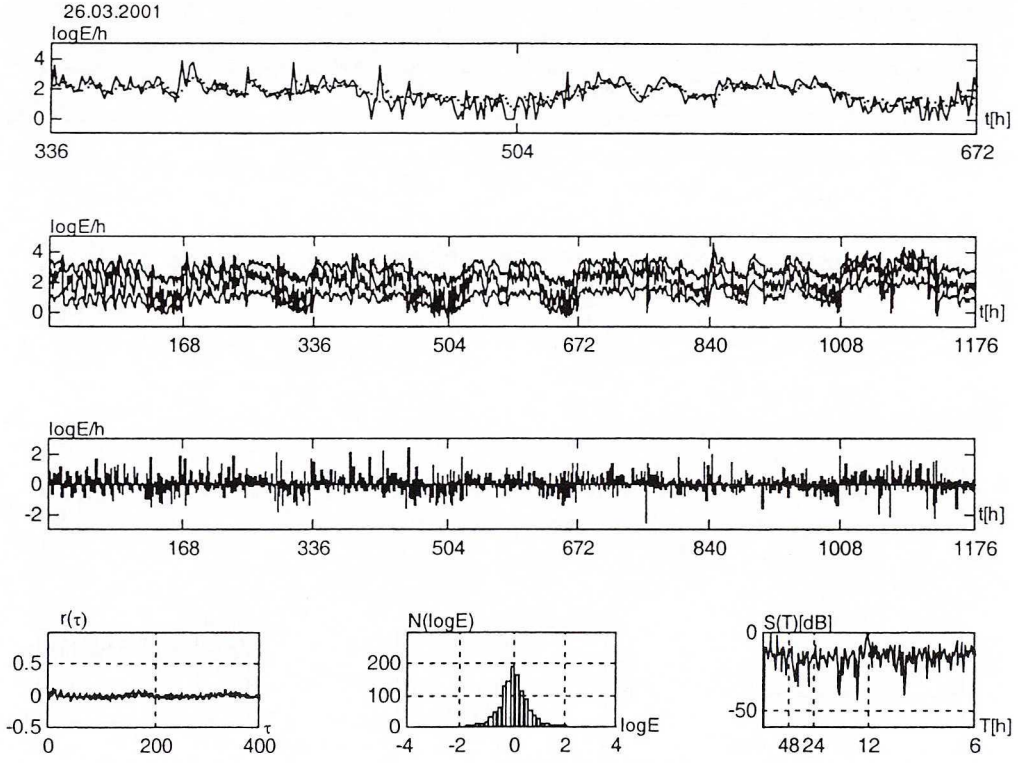

Fig. 3. Prediction results for the input data $E_{LT}^{G11}(t_i)$ [shown at Fig. 1]. From the top to the bottom:

- predicted (dots) and observed (continuous line) values of the $E_{LT}^{G11}(t_i)$, i = 336, ..., 672;
- observed values of the $E_{LT}^{G11}(t_i)$ i = 1, ..., 1176, between the 90% confidence limits for prediction;
- prediction errors;
- statistical characteristics of prediction errors (from the left to the right): autocorrelation function, energy-frequency histogram, spectrum

Rys. 3. Wyniki prognozy szeregu czasowego $E_{LT}^{G11}(t_i)$ [pokazanego na rys. 1]. Od góry:

- prognozowane (linia kropkowana) i obserwowane (linia ciągła) wartości $E_{LT}^{G11}(t_i)$ i = 336, ..., 672;
- wartości obserwowane $E_{LT}^{G11}(t_i)$, i = 1, ..., 1176 (w środku) oraz granice (90%) przedziału ufności dla prognozy;
- wartości błędów prognozy;
- charakterystyki statystyczne błędów prognozy (od lewej): funkcja autokorelacji, histogram rozkładu błędów, widmo szeregu czasowego błędów

The figures (Fig. 3–6), are built identically (and — as before — divided into "lines"). Except of the last line, horizontal (time) axis is scaled in hours, vertical (energy) axis — in logarithmic units. In the first line predicted (dots) and observed (continuous line) values of two– weeks lasting time series are visible; in the second line one can see **the observed values** [of the relevant $E_{LT}(t_i)$ time series] **between the 90% confidence limits for prediction** [note that the limits have always been estimated before the value

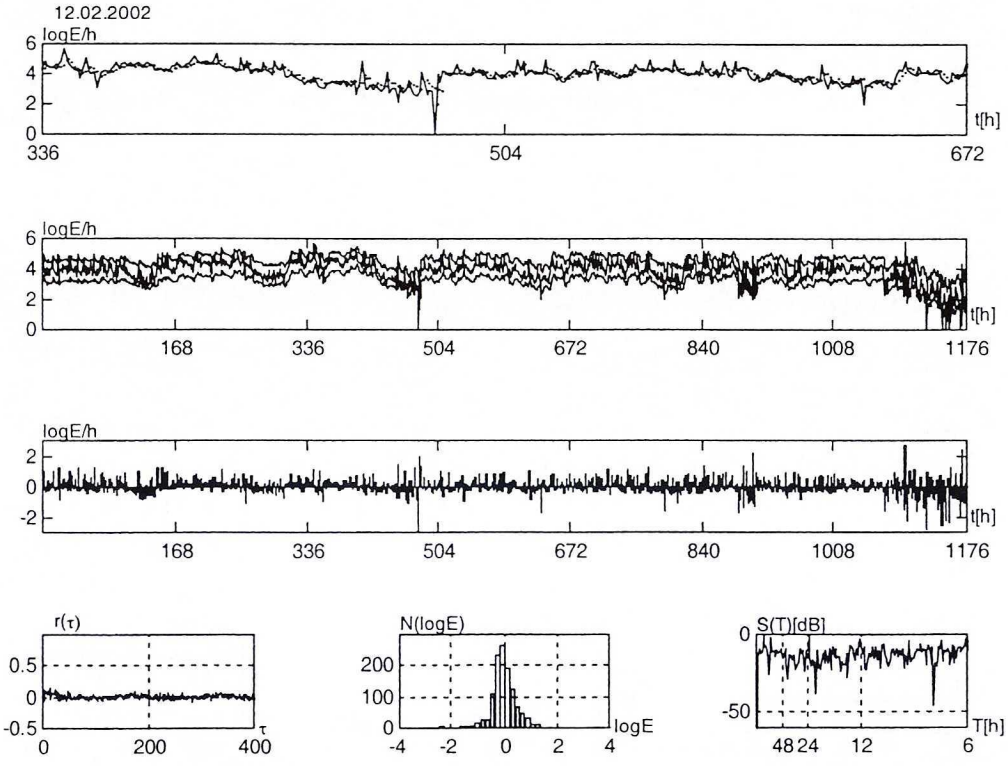

Fig. 4. Prediction results for the input data $E_{LT}^{A9}(t_i)$ [shown at Fig. 2]. From the top to the bottom:

- predicted (dots) and observed (continuous line) values of the $E_{LT}^{A9}(t_i)$, $i = 336, ..., 672$;
- observed values of the $E_{LT}^{A9}(t_i)$ $i = 1, ..., 1176$, between the 90% confidence limits for prediction;
- prediction errors;
- statistical characteristics of prediction errors (from the left to the right): autocorrelation function, energy-frequency histogram, spectrum

Rys. 4. Wyniki prognozy szeregu czasowego $E_{LT}^{A9}(t_i)$ [pokazanego na rys. 2]. Od góry:

- prognozowane (linia kropkowana) i obserwowane (linia ciągła) wartości $E_{LT}^{A9}(t_i)$ $i = 336, ..., 672$;
- wartości obserwowane $E_{LT}^{A9}(t_i)$, $i = 1, ..., 1176$ (w środku) oraz granice (90%) przedziału ufności dla prognozy;
- wartości błędów prognozy;
- charakterystyki statystyczne błędów prognozy (od lewej): funkcja autokorelacji, histogram rozkładu błędów, widmo szeregu czasowego błędów

of $E_{LT}(t_{i+1})$ has been observed] — for the full 1176-hours lasting data sequence. The third line exhibits (for the same 1176-hourly period) the prediction errors $[E_{LT}(t_i) - \overline{E}_{LT}(t_i)]$. In the last (the lowest) line, statistics of prediction errors are shown, from the left to the right: the autocorrelation function $r(\tau)$, histogram of prediction errors, spectrum of the errors sequence (with vertical scale in decibels and horizontal scale in hours).

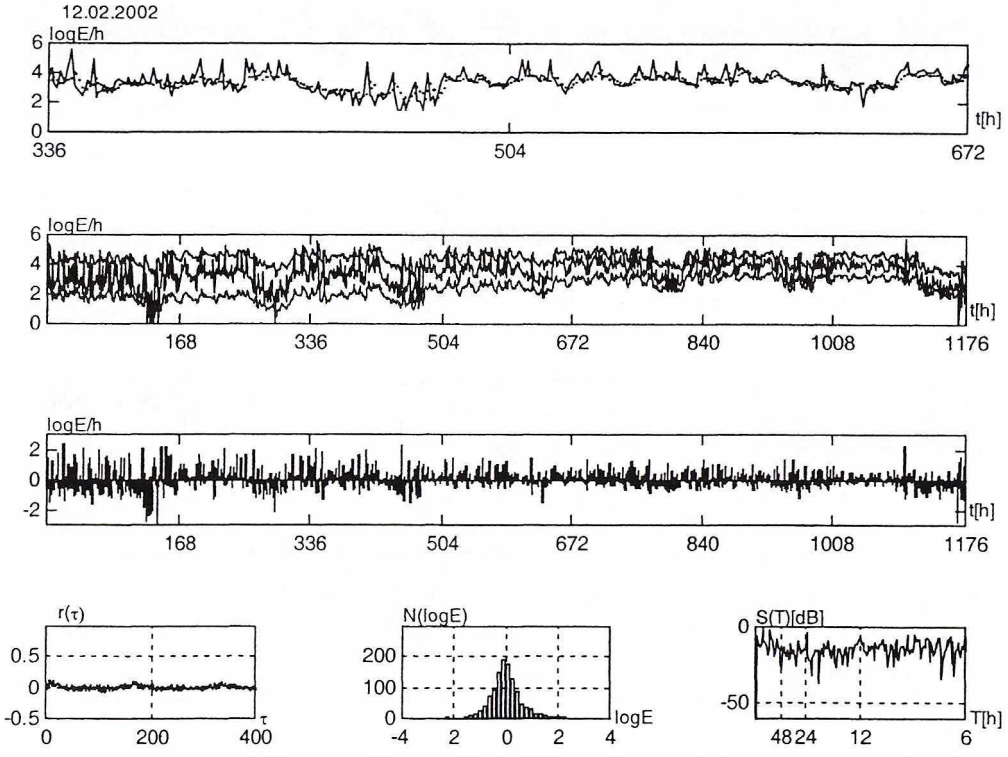

Fig. 5. Prediction results for the input data $E_{LT}^{A11}(t_i)$. From the top to the bottom:

- predicted (dots) and observed (continuous line) values of the $E_{LT}^{A11}(t_i)$, $i = 336, ..., 672$;
- observed values of the $E_{LT}^{A11}(t_i)$, $i = 1, ..., 1176$, between the 90% confidence limits for prediction;
- prediction errors;
- statistical characteristics of prediction errors (from the left to the right): autocorrelation function, energy-frequency histogram, spectrum

Rys. 5. Wyniki prognozy szeregu czasowego $E_{LT}^{A11}(t_i)$. Od góry:

- prognozowane (linia kropkowana) i obserwowane (linia ciągła) wartości $E_{LT}^{A11}(t_i)$, $i = 336, ..., 672$;
- wartości obserwowane $E_{LT}^{A11}(t_i)$, $i = 1, ..., 1176$ (w środku) oraz granice (90%) przedziału ufności dla prognozy;
- wartości błędów prognozy;
- charakterystyki statystyczne błędów prognozy (od lewej): funkcja autokorelacji, histogram rozkładu błędów, widmo szeregu czasowego błędów

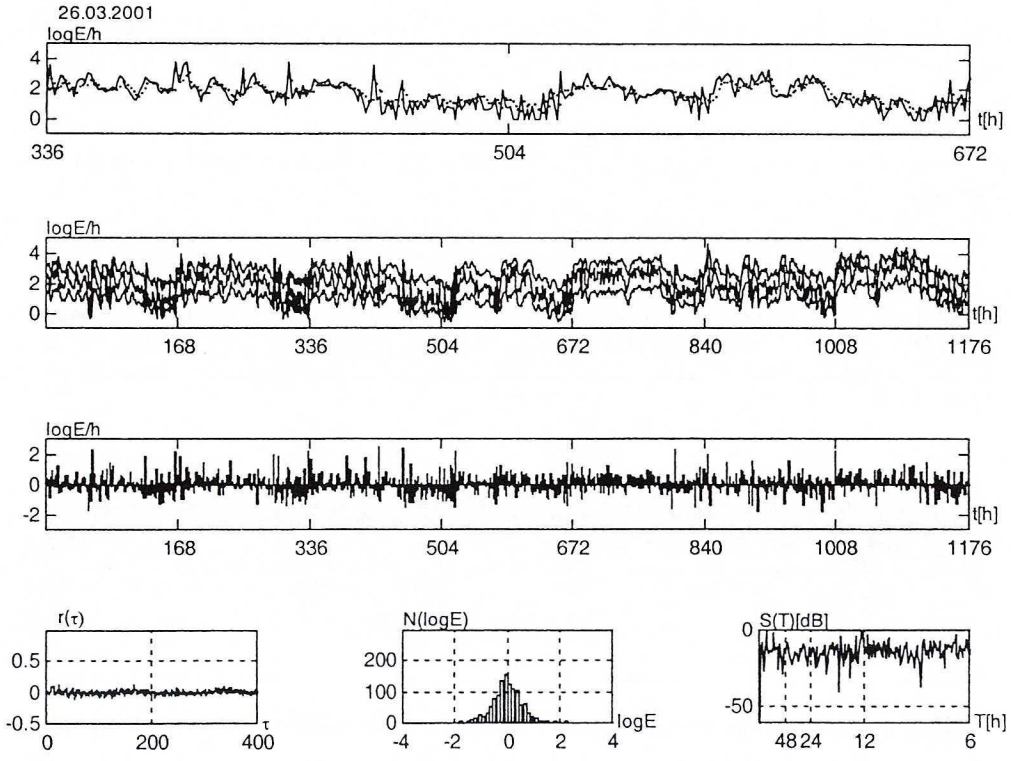

Fig. 6. Prediction results for the input data $E_{LT}^{G23}(t_i)$. From the top to the bottom:

- predicted (dots) and observed (continuous line) values of the $E_{LT}^{G23}(t_i)$, $i = 336, ..., 672$;
- observed values of the $E_{LT}^{G23}(t_i)$, $i = 1, ..., 1176$, between the 90% confidence limits for prediction;
- prediction errors;
- statistical characteristics of prediction errors (from the left to the right): autocorrelation function, energy-frequency histogram, spectrum

Rys. 6. Wyniki prognozy szeregu czasowego $E_{LT}^{G23}(t_i)$. Od góry:

- prognozowane (linia kropkowana) i obserwowane (linia ciągła) wartości $E_{LT}^{G23}(t_i)$, $i = 336, ..., 672$;
- wartości obserwowane $E_{LT}^{G23}(t_i)$, $i = 1, ..., 1176$ (w środku) oraz granice (90%) przedziału ufności dla prognozy;
- wartości błędów prognozy;
- charakterystyki statystyczne błędów prognozy (od lewej): funkcja autokorelacji, histogram rozkładu błędów, widmo szeregu czasowego błędów

A few conclusions can be drawn looking at these figures:

c1. Comparing observed values of $E_{LT}(t_i)$ and relevant confidence interval (for prediction) limits, one can qualitatively conclude that the predictors "generally well" adopt themselve to the changing conditions and that 90% confidence intervals are — at least approximately — trustworthy.

c2. Autocorrelation of prediction errors is generally small, below the 5% significance level ($1.96/\sqrt{1176} \approx 0.06$) but peaks in the spectra are still evident, confirming earlier fidings (e.g. b7) that the m_3 predictor results can be further improved (by some other linear predictors).

c3. The histograms of prediction errors are always "bell-shaped" but they are not strictly normal and frequently do not pass the standard normality test [the author has conjectured (Kornowski 2002a) that they can be better approximated by the so-called symmetric-α-stable densities with α slightly less than 2]. Anyway, to get quick and simple prediction intervals and hazard estimates, normality (of logarithmic energy errors) may be assumed.

c4. Prediction errors of the A9-37/WCM (from the highly hazardous region) — see Fig. 4 — are the smallest of the four analysed time series, with 90% — confidence interval of the width of (approximately) 1 logarithmic unit. It means that prediction accuracy is, in this case, similar to the observation accuracy ("to the order of magnitude").

It should be repeated, that the linear prediction — investigated here–informs us of the total (CPE) energy emitted during the prescribed time intervals from the observed region: this is not the same as "the time, place and magnitude of the nearest tremor", but the narrower are confidence limits and the shorter is T, the sampling interval of the CPE, the nearer are linear prediction results to the classical and much — desired forecast of "the tremor".

5. Comments and conclusions

The concept of the continuous process of (the total) emission, CPE, has been introduced and achievable results of the linear prediction have been checked for a few time series of hourly logarithmic energy of the CPE. The data has been collected at the two — of very different mining and geologic conditions — coal mines, so that the results, remaining examples only, are not tied to any single special conditions.

No data preprocessing, except of rescaling the AE energy with a constant coefficient and logarithmic transformation, has been applied — as this could complicate and expand the paper. Assessing the results, it should be remembered that the 1-D time series (e.g. the energy $E(t)$ of the CPE) is needed to obtain a well defined prediction and, as a result, not "the (badly defined) incoming tremor" but future values of the CPE are predicted. If the seismic hazard measure, associated with the CPE, or just predicted total energy, is considered useful, then:

- predictions $[\overline{E}_{LT}(t_{i+1})$ and $\sigma^2_E(t_{i+1})]$ are always estimable,
- the CPE (in all the many cases encountered by the author) is autocorrelated to a degree making prediction worthy of consideration.

The simple, low order, optimal linear predictor enables to achieve the performance index $0.45 < V_N < 0.57$ in the both coal mines and the benchmark-normalized index $0.56 < Q_N < 0.64$ in the not too hazardous ZGP and $0.075 < Q_N < 0.12$ in the hazardous

WCM — what is a very good result. Small improvements, on the order of a few percent, are still possible (in the context of the linear prediction) at a cost of more complicated algorithms.

The quality measures, V_N and Q_N, used here, are commonly accepted in prediction theory, but not best suited to seismological applications where — for small energies — prediction errors (large or not) can be unimportant at all, so that more complicated (weighted or conditioned) error measures could be considered. This is an interesting research field, but difficult as the current practice of data censoring deprives us the most important piece of information: as long as the highest energies are censored, they cannot be predicted.

As a final comment let me say that, in my opinion supported by the results of this paper, much better prediction results are still possible, if: wideband (in the energy and frequency domains) observational systems will be implemented in place of the two (AE and seismic) currently applied; and if energies of seismic and AE events are better estimated.

Theoretical studies of (models of) process leading to seismic event are needed to make a progress in the single-event prediction problem.

This research has been partly supported by the Polish KBN grant No 8T12B04521.

REFERENCES

Akaike H., 1974: A new look at the statistical model identyfication. IEEE Trans. Autom. Contr. AC-19, 716–723.

Broersen P.M.T., Wensink H.E., 1998: Autoregressive model order selection by a finite sample estimator for the Kullback-Leibler discrepancy. IEEE Trans. Signal Process. 46, 7, 2058–2061.

Dubiński J., Barański A., Gerlach Z., Mutke G., Nowak J., Syrek B., Szczerbiński J., 1996: Metoda sejsmologii górniczej oceny stanu zagrożenia tąpaniami. Instrukcja nr 1, Główny Instytut Górnictwa, Katowice, 15–35.

Geller R.J., 1997: Earthquake prediction: a critical review. Geophys. J. Int. 131, 3, 425–450.

Gibowicz S.J., Kijko A., 1994: An introduction to mining seismology. Academic Press, New York.

Jenkins G.M., Watts D.G., 1968: Spectral analysis and its applications. Holden-Day, San Francisco.

Kagan Y.Y., 1997: Are earthquakes predictable? Geophys. J. Int. 131, 3, 505–525.

Kagan Y.Y., 1999: Is earthquake seismology a hard, quantitative science? Pure Appl. Geophys. 155, 2–4, 233–258.

Kalouptsidis N., Carayannis G., Manolakis D., Koukoutsis E., 1985: Efficient recursive in order least squares FIR filtering and prediction. IEEE Trans. Acoust. Speech Sign. Process. ASSP-33, 1175–1187.

Kornowski J., 2002a: Podstawy sejsmoakustycznej oceny i prognozy zagrożenia sejsmicznego w górnictwie. Główny Instytut Górnictwa, Katowice.

Kornowski J., 2002b: Linear prediction of aggregated seismic and seismoacoustic energy emitted from a mining longwall. Acta Montana, in press.

Kornowski J., Kurzeja J., 2000: Korelacja energii wstrząsów górniczych z emisją sejsmoakustyczną i ocena możliwości jej wykorzystania w matematycznych modelach prognozy. XXIII Zimowa Szkoła Mech. Górotw.: „Geotechnika i budownictwo specjalne 2000", Wyd. Katedra Geomech. Górn. i Geotechniki AGH, Kraków, 181–195.

Kornowski J., Cianciara B., Sadlok R., Światłoch K., Trombik M., Zuberek W., 1996: Metoda sejsmoakustyczna oceny stanu zagrożenia tąpaniami. Instrukcja nr 1, Główny Instytut Górnictwa, Katowice, 36–54.

Lasocki S., 1995: Predykcja zagrożenia sejsmicznego. „Poradnik geofizyka górniczego" t. 2, Wyd. CPPGSMiE PAN, Kraków, 174–207.

Leadbetter M.R., Lindgren G., Rootzen H., 1983: Extremes and related properties of random sequences and processes. Springer Verlag, New York.

Lomnitz C., 1994: Fundamentals of earthquake prediction. J. Wiley, New York.

Marcak H., Zuberek W.M., 1994: Geofizyka górnicza. Śl. Wyd. Techn., Katowice.

Markel J.D., Gray A.H.Jr., 1976: Linear prediction of speech. Springer Verlag, New York.

Robinson E.A., 1967: Statistical communication and detection. Hafner Publ., New York.

Robinson E.A., Treitel S., 1980: Geophysical signal analysis. Prentice-Hall, Englewood Cliffs, N.J.

Rudajev V., Dragan V., Kašak K., 1985: Correlation of seismoacoustic and seismic data for the purpose of prediction of rockbursts. Publs. Inst. Geophys. Pol. Acad. Sc., M6(176), 249–262.

Silvia M.T., Robinson E.A., 1979: Deconvolution of geophysical time series in the exploration for oil and natural gas. Elsevier Sc. Publ., Amsterdam.

Strobach P., 1990: Linear prediction theory. A Mathematical Basis for Adaptive Systems. Springer Verlag, Berlin.

Vere-Jones D., 1970: Stochastic models for earthquake occurrence. J. Roy. Statist. Soc., Ser. B, 32, 1, 1–62.

Vere-Jones D., 1995: Forecasting earthquakes and earthquake risk. Int. J. Forecasting, 11, 503–538.

Wyss M., Dmowska R., 1997: Introduction. Pure Appl. Geophys., 149, 1–2.

REVIEW BY: PROF. DR HAB. INŻ. JÓZEF DUBIŃSKI, KATOWICE

Received: 11 December 2002