Transaction of Azerbaijan National Academy of Sciences, Series of Physical-Technical and Mathematical Sciences: Informatics and Control Problems, Vol. XXXVIII, No.3, 2018
www.icp.az/2018/3-07.pdf

U.E. SATTAROVA, M.E. ALIEV, T.A. BABAYEV, V.V. BAIRAMOV, P.F. PASHA

FORECASTING OF SEISMIC AND TECHNOLOGICAL PROCESSES USING AUTOREGRESSIVE MOVING AVERAGE METHOD

The autoregressive moving average (ARMA) method is used for forecasting of abnormal seismic and technological processes. The ARMA method is applied to seismoacoustic and technological signals to predict anomaly events. The analysis is verified through simulation studies.

Keywords: autoregressive moving average method, ARMA, signal behavior prediction, seismic signals, technological signals, MATLAB

1. Introduction. The importance of time series analysis and forecasting in science, engineering, and business has, in the past, increased steadily and it is still of actual interest for engineers and scientists. In process and production industry, of particular interest is time series forecasting where, based on some collected data, the future data values are predicted. This is important in process and production monitoring, in optimal processes control, etc. Mathematical models used for time series analysis are generally regression models, time-domain models and frequency-domain models. Regressive models are built using regression analysis, which is a collection of methods for the study of relationships between the variables and for estimation and prediction of values of one variable using the values of other variables incorporated in a joint time series [1, s.50-64; 2, s.83-110; 3-10]. For instance, to implement an efficient predictor for a variable of interest, the measurable variables representing the strong indicators for the same variable should first be identified.

2. Problem statement. The most popular regression models in engineering are the autoregression model (AR), moving-average model (MA), ARMA model, ARIMA model and CARIMA models. Autoregression models express the current value of a time series by a finite linear aggregate of previous values and by a shock μ_t [1, s.27-36].

$$Z_t = \alpha_1 Z_{t-1} + \alpha_2 Z_{t-2} + \cdots + \alpha_v Z_{t-v} + \mu_t \quad (2.1)$$

where a_1 to a_v are the autoregression parameters, μ_t is the white noise and v is the model order. The validity of an autoregressive model assumes that the time series to be modeled is stationary. Also, because of some possible internal cumulative effects, the autoregressive process will only be stable if the values of parameters a are within a certain range. It is common to write the autoregressive equation in terms of deviations $\hat{Z}_t = Z_t - \mu_t$ generally using the variable Z and its deviation $\hat{Z} = Z - \mu$. The individual terms of the time series now become $\hat{Z}_t, \hat{Z}_{t-1}, \hat{Z}_{t-2}, \hat{Z}_{t-3}, \ldots$, resulting in the autoregressive model [1, 2].

$$\hat{Z}_t = \phi_1 \hat{Z}_{t-1} + \phi_2 \hat{Z}_{t-2} + \phi_3 \hat{Z}_{t-3} + \cdots + \phi_p \hat{Z}_{t-p} + \alpha_t \quad (2.2)$$

where $\mu, \phi_1, \phi_2, \phi_3, \ldots, \phi_p, \alpha_t$ are unknown parameters to be estimated from the observation data. Introducing the autoregressive operator:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \ldots - \phi_p B^p \quad (2.3)$$

The autoregressive model can be written in the compact form

$$\phi(B)\hat{Z}_t = \alpha_t \quad (2.4)$$

A crucial problem in modeling of autoregressive time series is the selection of the order of the model to be built. A useful approach in this case is the analysis of the related partial autocorrelation function and the inverse autocorrelation function, because using the autocorrelation function itself is computationally complicated in the case of building of higher order models. Alternatively, fitting the time series shape by models of progressively higher order can be used, along with the analysis of the residual sum of squares for each order.

Another approach frequently used in modeling of univariate time series is based on the moving-average model [3]:

$$\hat{Z}_t = \alpha_t - \theta_1\alpha_{t-1} - \theta_2\alpha_{t-2} - \theta\alpha_{t-3} - \cdots - \theta_q\alpha_{t-q} \quad (2.5)$$

Which expresses $\hat{Z}_t$ in terms of an infinite weighted linear sum of $\alpha_t, \alpha_{t-1}, \alpha_{t-2}, \ldots, \alpha_{t-q}$. Introducing the moving-average operator of order q

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \quad (2.6)$$

The moving-average model can be written in the compact form as:

$$\hat{Z}_t = \theta(B)\alpha_t \quad (2.7)$$

The model contains (q+2) unknown parameters $\mu, \theta_1, \theta_2, \theta_3, \ldots \theta_q, \alpha_t$ to be estimated from the observation data.

The combination of the AR and MA models makes up the ARMA model [1, s.50-64]:

$$\hat{Z}_t = \phi_1\hat{Z}_{t-1} + \phi_2\hat{Z}_{t-2} + \phi_3\hat{Z}_{t-3} + \cdots + \phi_p\hat{Z}_{t-p} + \alpha_t - \theta_1\alpha_{t-1} - \theta_2\alpha_{t-2} - \theta\alpha_{t-3} - \cdots - \theta_q\alpha_{t-q} \quad (2.8)$$

Rewriting the model as:

$$\hat{Z}_t - \phi_1\hat{Z}_{t-1} - \phi_2\hat{Z}_{t-2} - \phi_3\hat{Z}_{t-3} - \cdots - \phi_p\hat{Z}_{t-p} = \alpha_t - \theta_1\alpha_{t-1} - \theta_2\alpha_{t-2} - \theta\alpha_{t-3} - \cdots - \theta_q\alpha_{t-q} \quad (2.9)$$

And rearranging it as:

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)\hat{Z}_t = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q)\,\alpha_t \quad (2.10)$$

The model can finally be written in compact form as:

$$\phi(B)\hat{Z}_t = \theta(B)\alpha_t \quad (2.11)$$

Where B is a delay operator. The derived compact model contains (p+q+2) unknown parameters $\mu, \phi_1, \phi_{,2}\ \phi_{,3} \ldots, \phi_p$ and $\theta_1, \theta_2, \theta_3, \ldots \theta_q, \alpha_t$ that are to be estimated from the given time series data. In practice, for the representation of actually occurring stationary time series, it is frequently adequate enough to take p and q not greater than 2. The presence of both autoregressive and moving-average terms in the ARMA model enables the representation of complex time series with fewer parameters than would be needed using a corresponding AR model.

When $\phi_1(B) = 1$, we have ARMA (p, q) = MA(q), and, when $\Theta(B) = 1$,we have ARMA (p, q) = AR(p). Such processes are often denoted as ARMA (0, q) and ARMA (p, 0) to stress the fact that the moving average model and the autoregressive model are members of the family of ARMA models [2, s.83-110].

Let

$$Z_t - \phi_1 Z_{t-1} - \ldots - \phi_p Z_{t-p} = \alpha_t \quad (2.12)$$

we assume that the roots of $\phi(z)$ are outside the unit circle. When $\tau > p$, the linear combination minimizing the mean square linear prediction error is

$$\text{f}(p) = \sum_{j=1}^{p} \phi_j Z_{\tau-j} \quad (2.13)$$

We will discuss this result later. Now we will use it to obtain the partial autocorrelation function (PACF) for $\tau > p$ namely:

$$\phi_{\tau\tau} = corr\{Z_\tau - \text{f}(p), Z_0 - \text{f}(p)\} = corr\{\alpha_\tau, Z_0 - \text{f}(p)\} = 0 \quad (2.14)$$

Since, by causality, $Z_{\tau-j}$ does not depend on the future noise value a_τ. When $\tau \le p, \phi_{pp} \neq 0$ and $\phi_{11}, \ldots, \phi_{p-1,p-1}$ are not necessarily zero.

3. Using Arma Method For Prediction Of Seismic and Technological Signals. Since the beginning of seismology 100 years ago it has been the hope of seismologists to be able to predict earthquakes in order to help populations across the globe avoid destruction and casualties. Nonetheless, earthquakes continue to occur without warning. Throughout the decades, many have claimed to have noticed something unusual before earthquakes occurred, something akin to precursors to them. These included light phenomena, sounds, the anomalous activity of small precursor earthquakes, strange animal behavior, a change in ground water level, and even dreams. There is no reason to reject all such claims as the ranting of people who recently suffered a major earthquake. In some cases, people had already pointed out such phenomena before the earthquake struck. However, such experiences did not help in making predictions. Hence there predominated a pessimistic view about earthquakes, one that viewed them as among the many inevitably unpredictable forces of nature. The Establishment of Seismic Stations Network at ANAS Institute of Control Systems was a step forward [11-13]. The ARMA model can be used to capture signals as well as time series for predicting future values both of the signal generated by these stations and any technological process. The major advantage of time series is possible to predict the future value based on the previous historical data. The study of the past sequence of historical data may be more valuable. The time series is helpful to predict the next sequence of future values. The utility of time series method is specifically for trend analysis of technological signals, trade market, finance, climatologic and earthquake prediction. ARMA modeling is based on a unique decomposition of a strictly proper order of the numerator is less than the order of the denominator, discrete transfer function. Once the time series model has been developed and tested it can be used for forecasting the future time series values at various time distances *d*. Of course, the forecasting does not deliver the exact future values of data that the given time series will really have, but rather their estimates. Using the autoregressive model:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \varepsilon_t \quad (3.1)$$

Based on a one-step movement along the time series:

$$Z_{t+1} = \phi_1 Z_t + \phi_2 Z_{t-1} + \varepsilon_{t+1} \quad (3.2)$$

We can formally write the predicted value to be:

$$\hat{Z}_{t+1} = \phi_1 Z_t + \phi_2 Z_{t-1} \quad (3.3)$$

For the two-steps ahead prediction, based on a two-steps movement along the time series, we can also formally write

$$Z_{t+2} = \phi_1 (\phi_1 Z_t + \phi_2 Z_{t-1} + \varepsilon_{t+1}) + \phi_2 Z_{t-1} + \varepsilon_{t+1} \quad (3.4)$$

And the predicted value to be

$$\hat{Z}_{t+2} = \phi_1 Z_{t+1} + \phi_2 Z_t \quad (3.5)$$

Or we have:

$$\hat{Z}_{t+2} = \phi_1 (\phi_1 Z_{t+1} + \phi_2 Z_t) + \phi_2 Z_t \quad (3.6)$$

The associated block diagram for the ARMA configuration is shown in Fig. 1. The Matlab programs for ARMA prediction of seismic and technological signals are given in Appendix. The technological signal and its prediction result using ARMA method is shown in Fig. 2. The 100 samples of recorded technological signal are applied to forecast the next 10 samples of technological signals using ARMA method.

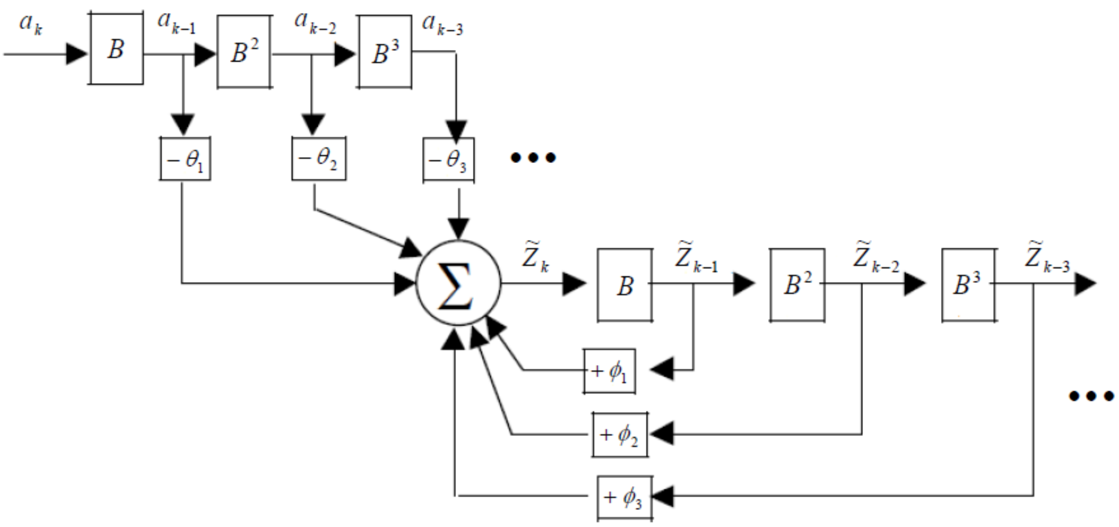

Fig. 1. The block diagram for the ARMA configuration

4. Conclusion. The ARMA prediction method was applied to forecast the seismic signal behavior. The analysis was verified through simulation studies using MATLAB software. It is presented in this work a constructive approach to build time series forecasting models, assuming no prior knowledge about the behavior of the seismic and technological signals. It is shown that the ARMA prediction method is a simpler and more efficient algorithm for forecasting univariate time series of seismic and technological signals. The Matlab programs are given in Appendix.

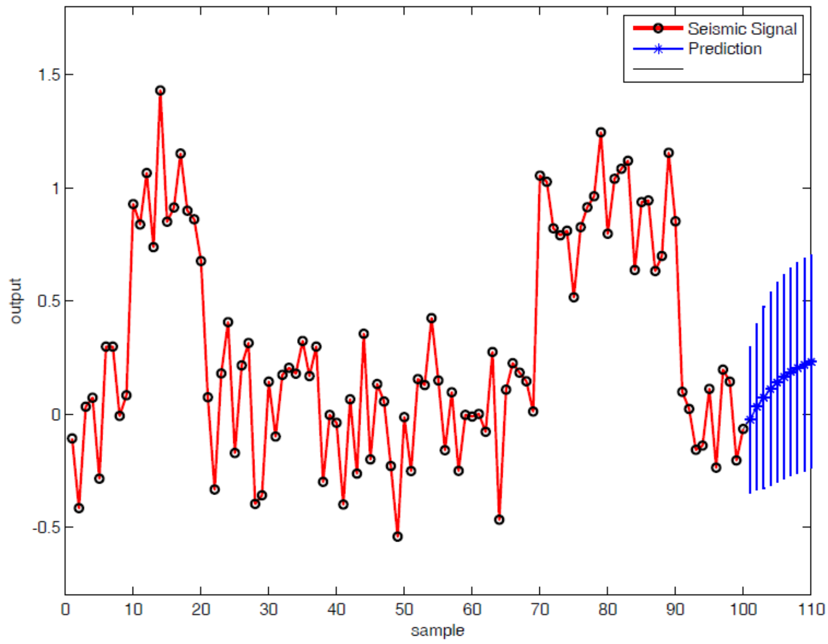

Fig. 2. The signal and its prediction result using ARMA method

Acknowledgements

The article was supported by the State Oil Fund of Azerbaijan Republic in 2017 in the frame of grant project.

References

1. Ajoy K. Palit and Dobrivoje Popovic. “Computational Intelligence in time series forecasting”, Springer, 2005, 272 p.
2. Brockwell PJ and Davis RA. (2010) Introduction to Time Series and Forecasting, 2nd edition, Springer-Verlag, New York., 435 p.
3. Paulo Cortez, Miguel Rocha. “Evolving Time Series Forecasting ARMA Models”, journal of Heuristics July 2004, Volume 10, Issue 4, pp 415–429.
4. V.N. Ovcharuk. Spectral analysis of signals acoustic emission. // http://ejournal.khstu.ru/media/2013/TGU_4_187.pdf.
5. Kim J.H. (2003). Forecasting autoregressive time series with bias-corrected parameter estimators. International Journal of Forecasting, 19, pp.493–502.
6. Poskitt D.S. (2003). On the specification of cointegrated autoregressive moving-average forecasting systems. International Journal of Forecasting, 19, pp.503–519.
7. Riise T. & Tjøstheim D. (1984). Theory and practice of multivariate ARMA forecasting. Journal of Forecasting, 3, pp.309– 317.
8. Dahl C.M. & Hylleberg S. (2004). Flexible regression models and relative forecast performance. International Journal of Forecasting, 20, pp.201–217.
9. Rashid G. Alakbarov, Fahrad H. Pashaev, Oqtay R. Alakbarov. "Forecasting Cloudlet Development on Mobile Computing Clouds", International Journal of Information Technology and Computer Science(IJITCS), Vol.9, No.11, pp.23-34, 2017. DOI: 10.5815/ijitcs.2017.11.03
10. Chunxiao Bao, Hong Hao, Zhong-Xian Li. Integrated ARMA model method for damage detection of subsea pipeline system.// Engineering Structures, Volume 48, March 2013, pp.176-192
11. Telman Aliyev, Akif Ali-zada, Gurban Etirmishli, Gambar Guluyev, Fahrad Pashayev, Abbas Rzayev. Intelligent Seismoacoustic System for Monitoring the beginning of Anomalous Seismic Process. // Seismic Instruments, 2011, Vol 47, No. 1, pp. 15-23.
12. Aliev T.A., Abbasov A.M., Guluyev G.A., Pashaev F.H., Sattarova U.E. System of robust noise monitoring of anomalous seismic processes // Soil Dynamics and Earthquake Engineering, 53 (2013), pp.11-26.
13. Aliev T., Quluyev Q., Pashayev F., Sattarova U. and Rzayeva N. (2016) Intelligent Seismic-Acoustic System for Identifying the Location of the Areas of an Expected Earthquake. Journal of Geoscience and Environment Protection, 4, pp.147-162.

APPENDIX

ARMApred.m Matlab file

```
%
% This script demonstrates the process of ARMA modeling and prediction
% for the seismic batch process data.
%
% First, clear memory and reset the RNG's.
%
clear
rand('state',0); randn('state',0);
%
% load in the data.
%
%load batch.txt
clc;
clear; t=[1:100]; x1=ones(1,100); x=zeros(1,100);
for i=10:20 x(1,i)=-1; end
for i=70:90 x(1,i)=-1; end
batch=0.25*randn(1,100)+(randn(1,1)+0.2*sin(20*t)*x1')*x;
%
% Take out the mean.
%
```

```
batchmean=mean(batch); batch=batch-batchmean;
%
% Estimate the autocorrelations.
%
myr=zeros(20,1);
for k=1:20,  myr(k)=r(batch,k); end
%
% Estimate the PACF.
%
phihat=pacf(myr,10)
%
% It appears that phihat(1,1) and phihat(2,2) are nonzero, but
% after that, the values of phihat are quite small and could just
% be due to random sampling variability.
%
% From the PACF and autocorrelations, it appears that an ARMA(2,0) model
% is appropriate for these data. Estimate phi1 and phi2.
%
phi1=(myr(1)*(myr(2)-1))/(myr(1)^2-1) phi2=(myr(1)^2-myr(2))/(myr(1)^2-1)
%
% We get phi1=0.5355, phi2=0.2534.
%
% So, our model is z(n)=a(n)+phi1*z(n-1)+phi2*z(n-2). The phi
% function is phi(B)=1-phi1*B-phi2*B^2. Since these zeros of phi(B) are outside of the unit
% circle, we have a stationary ARMA process.
%
%
% Next, we estimate the a's.
%
%
% We don't know anything before time 1, so we treat all earlier z's
% as if they were 0. The transient effect of this goes away as
% time goes on.
%
% Since z(1)=a(1)+phi1*z(0)+phi2*z(-1), we estimate a(1)=z(1).
% Since z(2)=a(2)+phi1*z(1)+phi2*z(0), we estimate a(2)=z(2)-phi1*z(1).
%
a=zeros(size(batch)); a(1)=batch(1); a(2)=batch(2)-phi1*batch(1);
%
% From here on in, we can use previous z's to predict the next z.
% The error in this prediction is our estimate of a.
%
% zpred(n)=phi1*z(n-1)+phi2*z(n-2)
% z(n)=a(n)+phi1*z(n-1)+phi2*z(n-2)
%
% So, we get
%
% a(n)=z(n)-zpred(n)
```

```
%
for k=3:length(batch),
batchnext=phi1*batch(k-1)+phi2*batch(k-2); a(k)=batch(k)-batchnext; end
%
% Estimate sigmaa.
%
sigmaa=std(a)
%
% Now, predict the future z's. The prediction formula is
%
% z(k+1)=phi(1)*z(k)+phi(2)*z(k-1)
%
n=length(batch);
for k=1:10, batch(n+k)=phi1*batch(n+k-1)+phi2*batch(n+k-2); end
%
% Note that these predictions are made with the demeaned series, so
% we'll still have to add the mean back in at the end.
%
% In order to put confidence intervals around our predictions, we
% need the Psi form of the model.
%
% Psi(B)=1/(1-phi1*B-phi2*B^2)=
% 1 + 0.5355 B + 0.5402 B^2 + 0.4250 B^3 + 0.3644 B^4 + 0.3028 B^5 + ...
%
psi=[0.5355; 0.5402; 0.4250; 0.3644; 0.3028; 0.2545; 0.2130; 0.1786; 0.1496]
v=zeros(10,1);
v(1)=sigmaa^2*1; v(2)=sigmaa^2*(1+psi(1)^2); v(3)=sigmaa^2*(1+psi(1)^2+psi(2)^2);
v(4)=sigmaa^2*(1+psi(1)^2+psi(2)^2+psi(3)^2);
v(5)=sigmaa^2*(1+psi(1)^2+psi(2)^2+psi(3)^2+psi(4)^2);
v(6)=sigmaa^2*(1+psi(1)^2+psi(2)^2+psi(3)^2+psi(4)^2+psi(5)^2);
v(7)=sigmaa^2*(1+psi(1)^2+psi(2)^2+psi(3)^2+psi(4)^2+psi(5)^2+psi(6)^2);
v(8)=sigmaa^2*(1+psi(1)^2+psi(2)^2+psi(3)^2+psi(4)^2+psi(5)^2+psi(6)^2+psi(7)^2);
v(9)=sigmaa^2*(1+psi(1)^2+psi(2)^2+psi(3)^2+psi(4)^2+psi(5)^2+psi(6)^2+psi(7)^2+psi(8)^2);
v(10)=sigmaa^2*(1+psi(1)^2+psi(2)^2+psi(3)^2+psi(4)^2+psi(5)^2+psi(6)^2+psi(7)^2+psi(8)^2+p
si(9)^2);
%
% We'll construct 1 standard deviation error bars for the predicted
% points.
%
ebar=sqrt(v);
%
% Finally, plot the data and the predicted data points, with error bars.
%
figure(1); plot(1:n,batch(1:n)+batchmean,'ko-');
hold on
%
% The new points. Use "*" markers and error bars.
%
```

```
errorbar((n+1):(n+10),batchmean+batch(n+1:n+10),ebar,'k*-');
%
% Connect the old and new points.
%
plot(n:n+1,batchmean+batch(n:n+1),'k-');
%
% label the axes.
%
xlabel('sample'); ylabel('output');
%
% reset the axes.
%
axis([0 110 -0.8 1.8]); print -deps batchpred.eps
```

c.m Matlab file

```
%
%  c=c(z,k)
%
%  Estimates the autocovariance c(k) for a time
%  series z.
%
function c=c(z,k)
n=length(z);  zbar=mean(z); c=0;
for i=1:n-k,  c=c+(z(i)-zbar)*(z(i+k)-zbar); end;
c=c/n;
```

pacf.m Matlab file

```
%
% phihat=pacf(r,n)
%
% Uses the Durbin recursive formulas to estimate the PACF coefficients.
%
function phihat=pacf(myr,n)
phihat=zeros(n,n); phihat(1,1)=myr(1);
for p=1:n-1,
  num=myr(p+1);
  for j=1:p,  num=num-phihat(p,j)*myr(p+1-j); end;
  denom=1;
  for j=1:p,
    denom=denom-phihat(p,j)*myr(j);
  end;
  phihat(p+1,p+1)=num/denom;
  for j=1:p,
    phihat(p+1,j)=phihat(p,j)-phihat(p+1,p+1)*phihat(p,p-j+1);
  end; end;
```

r.m Matlab file

```
%
%  r=r(z,k)
%
%  Estimates the autocorrelation r(k) from a data
```

```
% set z.
%
function r=r(z,k)
r=c(z,k)/c(z,0);
```

UOT 004.021

Ü.E. Səttarova, M.E. Əliyev, T.Ə. Babayev, V.V. Bayramov, P.F. Paşa

Avtoreqressiv gəzən orta qiymət metodlarının tətbiqi ilə seysmik və texnoloji proseslərin proqnozlaşdırılması

Anomal seysmik proseslərin inkişafının proqnozlaşdırılması üçün avtoreqressiv gəzən orta qiymət metodunun tətbiqi verilmişdir. ARMA metodu seysmik proseslərin və texnoloji proseslərin anomal istiqamətdə dəyişməsinin proqnozlaşdırılması üçün tətbiq edilmişdir. Tədqiqatların nəticəsi kompüter eksperimentləri vasitəsi ilə test edilmişdir.

Açar sözlər: avtoreqressiv gəzən orta qiymət, ARMA, siqnalın dəyişməsinin proqnozlaşdırılması, seysmik siqnal, texnoloji siqnal, MATLAB

УДК 004.021

У.Э. Саттарова, М.Е. Алиев, Т.А. Бабаев, В.В. Байрамов, П.Ф. Паша

Прогнозирование сейсмических и технологических процессов с применением авторегрессионного метода скользящего среднего

Описывается авторегрессионный метод скользящего среднего с целью прогнозирования развития аномальных сейсмических процессов. Метод авторегрессии скользящего среднего применялся также для прогнозирования изменения состояния технологических процессов. Результаты исследования проверялись при помощи вычислительных экспериментов.

Ключевые слова: метод авторегрессии скользящего среднего, ARMA, прогнозирование изменения сигналов, сейсмический сигнал, технологический сигнал, MATLAB

Institute of Control Systems of ANAS
Azerbaijan University of Architecture and Construction
Presented 05.03.2018