СИСТЕМНИЙ АНАЛІЗ

SYSTEM ANALYSIS

UDC 621.391.24: 519.246.8: 534.6
DOI https://doi.org/10.32782/tnv-tech.2026.2.1.46

SINGULAR SPECTRUM ANALYSIS METHOD FOR IDENTIFICATION OF NON-STATIONARY ACOUSTIC SIGNALS

Dymova H. O. *– Candidate of Technical Sciences,*
Associate Professor at the Department of Management, Marketing
and Information Technology
Kherson State Agrarian and Economic University
ORCID ID: 0000-0002-5294-1756

Larchenko O. V. *– Candidate of Agricultural Sciences,*
Associate Professor at the Department of Management, Marketing
and Information Technology
Kherson State Agrarian and Economic University
ORCID ID: 0000-0001-7857-0802

The article addresses the topical scientific and applied problem of automated recognition and classification of non-stationary acoustic signals generated by technical equipment or military operations. The main research object is the time series of complex acoustic patterns, the processing of which by classical statistical methods is complicated by their non-stationarity and high noise levels. The aim of the work is to develop a noise-robust method for identifying acoustic emission sources based on time series analysis. To achieve this, the Singular Spectrum Analysis (SSA) method, also known as the "worm" algorithm, was employed. This approach is based on embedding a one-dimensional signal into a multidimensional trajectory matrix followed by singular value decomposition (SVD). The mathematical model describes the process of unfolding the series, calculating principal components, and reconstructing the signal through the procedure of Hankelization.

A significant part of the research is devoted to the selection of principal components using an energy criterion, where the set of indices is formed such that the cumulative contribution of eigenvalues exceeds a threshold of 95%. This allows for the automatic separation of informative signal components from the noise subspace. Experimental studies were conducted in two stages: detailed modeling of matrix transformations for a single signal in the Mathcad 15.0 environment and statistical verification on a large sample using a developed software module in

Python. Testing the algorithm on signals from a machine drive and model impulses (simulating explosions) confirmed its high adaptability.

Based on Monte Carlo statistical simulation (1000 iterations), it was established that the mean absolute percentage error (MAPE) does not exceed 6% at a signal-to-noise ratio (SNR) above 20 dB. The proposed SSA method demonstrated higher accuracy compared to the classical ARIMA(1,1,1) model under intensive noise conditions (5.82% versus 7.45%). It is concluded that the use of SSA for automated acoustic monitoring systems is highly feasible for military and rescue operations.

Key words: *acoustic analysis, signal recognition, time series, Singular Spectrum Analysis, SSA, "worm" algorithm, correlation analysis, signal identification.*

Димова Г. О., Ларченко О. В. Метод аналізу сингулярного спектра для ідентифікації нестаціонарних акустичних сигналів

У статті вирішується актуальна науково-прикладна задача автоматизованого розпізнавання та класифікації нестаціонарних акустичних сигналів, що генеруються технічним обладнанням або внаслідок військових дій. Основним об'єктом дослідження є часові ряди складних акустичних патернів, обробка яких класичними статистичними методами ускладнена через їхню нестаціонарність та високий рівень завад. Метою роботи є розробка завадостійкого методу ідентифікації джерел акустичної емісії на основі аналізу часових рядів. Для досягнення поставленої мети застосовано метод аналізу сингулярного спектра (Singular Spectrum Analysis, SSA), відомий як алгоритм «гусениця». Цей підхід базується на вкладенні одновимірного сигналу в багатовимірну траєкторну матрицю з подальшим застосуванням сингулярного розкладу (SVD). Математична модель описує процес розгортання ряду, обчислення головних компонент та реконструкцію сигналу за допомогою процедури ганкелізації.

Значна частина дослідження присвячена вибору головних компонент за допомогою енергетичного критерію, де набір індексів формується таким чином, щоб кумулятивний внесок власних значень перевищував поріг у 95%. Це дозволяє автоматично відокремлювати інформативні компоненти сигналу від підпростору шуму. Експериментальні дослідження проводилися у два етапи: детальне моделювання матричних перетворень для окремого сигналу в середовищі Mathcad 15.0 та статистична верифікація на великій вибірці за допомогою розробленого програмного модуля на Python. Тестування алгоритму на сигналах машинного приводу та модельних імпульсах (імітація вибухів) підтвердило його високу адаптивність.

На основі статистичного моделювання методом Монте-Карло (1000 ітерацій) встановлено, що середня абсолютна помилка у відсотках (MAPE) не перевищує 6% при відношенні сигнал/шум (SNR) понад 20 дБ. Запропонований метод SSA продемонстрував вищу точність порівняно з класичною моделлю ARIMA(1,1,1) в умовах інтенсивного шуму (5,82% проти 7,45%). Зроблено висновок, що використання SSA для автоматизованих систем акустичного моніторингу є цілком доцільним для військових та рятувальних операцій.

Ключові слова: *акустичний аналіз, розпізнавання сигналів, часові ряди, сингулярний спектральний аналіз, SSA, алгоритм «хробака», кореляційний аналіз, ідентифікація сигналів.*

Introduction. Acoustic analysis is an important direction in the fields of security, defense, and rescue operations. The detection and identification of specific acoustic patterns allow not only for the recognition of noise sources but also for the prediction of their future behavior.

Formulation of the problem. Time series representing acoustic signals can have a complex structure that complicates their processing by standard methods. This paper considers the Singular Spectrum Analysis (SSA) method, also known as the "worm" algorithm, as one of the effective approaches to the analysis and recognition of acoustic signals.

The purpose of the article is to develop a noise-robust method for identifying acoustic emission sources based on time series analysis, allowing for the effective processing of complex acoustic data for practical applications in the military sphere and rescue operations.

Research analysis. In modern scientific literature, time series analysis and its use for acoustic signal recognition are based on several key approaches. In particular, correlation analysis, spectral analysis methods, smoothing, and deep neural networks (RNN, LSTM, GRU) are widely used.

Correlation analysis is one of the classic methods for estimating signal similarity. It allows determining periodicity and identifying patterns. Smoothing methods, such as moving average and exponential smoothing, are used to eliminate noise and identify main trends. Frequency analysis based on Fourier and Wavelet transforms is used to identify dominant frequencies in a signal.

Dynamic Time Warping (DTW) is an effective way to measure similarity between time series, even if they have different lengths or distortions in time. For more complex patterns, deep neural networks are used, which allow identifying complex regularities and performing forecasting.

However, the SSA method ("worm" algorithm) is less studied in the context of acoustic analysis. Its advantage lies in the possibility of unfolding a one-dimensional series into a multidimensional space, which allows for more effective extraction of characteristic features of the signal. Studies have shown that this method can be used for accurate recognition of acoustic signals and reduction of noise influence.

Presentation of the main material. Let us consider a time series $\{x_i, i = 1, N\}$, formed by a sequence of N equidistant values of some (possibly random) function $f(t)$:

$$x_i = f[i] = f((i\text{-}1)\Delta t,$$

where i=1,2, …, N.

There are several variants of the base algorithm:

- correlation analysis;
- smoothing methods;
- frequency analysis;
- dynamic time warping;
- AR, MA, ARMA, ARIMA models;
- deep neural networks.

Correlation analysis is used to determine periodicity and identify patterns in time series. Smoothing methods include moving average and exponential smoothing, which are applied to eliminate noise and reveal the main trend. Fourier and Wavelet transforms determine dominant frequencies in the signal, which relates to frequency analysis. Dynamic Time Warping (DTW) is used to measure the similarity between two time series, even if they vary in length or speed. Statistical methods for time series modeling and forecasting include AR, MA, ARMA, and ARIMA models. Deep neural networks (RNN, LSTM, GRU) are applied for complex patterns in time series, for example, for recognizing acoustic signals.

We will consider the most applicable variant – the so-called correlation analysis. Also, for the recognition of acoustic signals, we will use the "worm" algorithm (Singular Spectrum Analysis – SSA).

The "worm" algorithm can be directly divided into four stages:

1) obtaining a multidimensional series from a one-dimensional one;
2) principal component analysis;
3) selection of principal components;
4) reconstruction of the one-dimensional series.

Let us provide a description of the "worm" algorithm (SSA) for analyzing a one-dimensional time series $\left\{x_i, i = \overline{1, N}\right\}$ [1]:

1. Embedding (Unfolding the one-dimensional series into a multidimensional one).

To do this, we select a certain number $M < N$ (the “worm” length/window length) and let $k = N - M + 1$. We obtain a matrix $\mathbf{X} = \left(x_{i,j}\right)_{i,j=1}^{kM}$ with elements $x_{i,j} = x_{i+j-1}$ corresponding to the matrix [2, 3]:

$$\mathbf{X} = \left(x_{ij}\right)_{i,j=1}^{k,M} = \begin{pmatrix} x_1 & x_2 & x_3 & \ldots & x_M \\ x_2 & x_3 & x_4 & \ldots & x_{M+1} \\ x_3 & x_4 & x_5 & \ldots & x_{M+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_k & x_{k+1} & x_{k+2} & \ldots & x_N \end{pmatrix} \quad (1)$$

Subsequently, Principal Component Analysis (PCA) takes place. Data processing by the principal component method is shown in (Fig. 1).

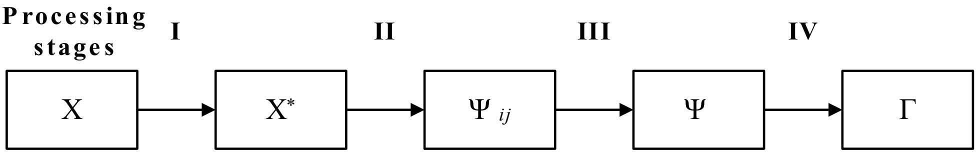


Fig. 1. Data processing by the principal component method

2. Calculation of the matrix $\mathbf{X}^*$.

Based on matrix $\mathbf{X}$, we calculate the matrix $\mathbf{X}^* = \left(x^*_{i,j}\right)_{i,j=1}^{kM}$. The general formula for the elements of this matrix is as follows:

$$x^*_{ij} = \frac{x_{ij} - \mu_j}{\delta_j},\ i = 1,2,\ldots,k,\ j = 1,2,\ldots,M \quad (2)$$

$$\mu_j = \begin{cases} 0 \\ \bar{x}_j \end{cases},\ \delta_j = \begin{cases} 1 \\ s_j \end{cases}, \quad (3)$$

where $\bar{x}_j = \frac{1}{k}\sum_{i=l}^{k} x_{ij},\ s_j = \sqrt{\frac{1}{k}\sum_{i=1}^{k}\left(x_{ij} - \bar{x}_j\right)^2}$.

3. Calculation of the matrix $\mathbf{\Psi}$ according to the formula:

$$\cdot\cdot\ = \frac{1}{k}\mathbf{X}^*(\mathbf{X}^*)^T \quad (4)$$

4. The next step in PCA consists of calculating the eigenvalues and eigenvectors of the matrix $\mathbf{\Psi}$ [2, 3, 4], i.e., its decomposition:

$$\mathbf{\Psi} = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^T \quad (5)$$

where $\mathbf{\Lambda}=\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_M \end{pmatrix}$ is the diagonal matrix of eigenvalues;

$\boldsymbol{\$} = \left(\phi_1, \phi_2, \ldots, \phi_M\right) = \begin{pmatrix} \phi_{11} & \phi_{21} & \cdots & \phi_{M1} \\ \phi_{12} & \phi_{22} & \cdots & \phi_{M2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{1M} & \phi_{1M} & \cdots & \phi_{MM} \end{pmatrix}$ is the orthogonal matrix of eigenvectors of matrix $\mathbf{\Psi}$.

In this case, the following relations hold: $\mathbf{\Phi}^T=\mathbf{\Phi}^{-1}$; $\mathbf{\Phi}^T\mathbf{\Phi}=\mathbf{\Phi}\mathbf{\Phi}^T=\mathbf{I}_M$; $\mathbf{\Lambda}= \mathbf{\Phi}^T\mathbf{\Psi}\mathbf{\Phi}$; $\sum_{i=1}^{M}\lambda_i = \mathrm{M}$; $\prod_{i=1}^{M}\lambda_i = det\Psi$.

Matrices $\mathbf{\Lambda}$ and $\mathbf{\Phi}$ jointly have a whole range of interpretations based on PCA.

5. Calculation of principal components according to the formula [1, 6]:

$$\mathbf{X}^*\mathbf{\Phi}=\mathbf{\Gamma}=(\gamma_1, \gamma_2, \ldots, \gamma_M). \tag{6}$$

6. Selection of principal components – we choose a certain number of principal components from $\mathbf{\Gamma}$ with numbers $i_1, i_2, \ldots, i_r$

To formalize the selection procedure, it is proposed to use the energy criterion. The set of indices $\mathbf{I} = \{i_1, \ldots, i_r\}$ is formed such that the cumulative contribution of eigenvalues exceeds a specified threshold (e.g., 95%):

$$\frac{\sum_{i \in I}\lambda_i}{\sum_{j=1}^{M}\lambda_j} \geq 0{,}95$$

This allows for automatically separating informative components (signal) from the noise subspace (the "tail" of the eigenvalue spectrum).

7. Reconstruction of the multidimensional series:

$$\tilde{X}^* = \sum_{l=1}^{r}\gamma_{i_l}\phi_{i_l}^T$$

8. Decentering and deformation (if necessary) of the matrix $\tilde{X}^*$. Result – matrix $\tilde{X}$.

9. Hankelization (reconstruction of the one-dimensional series by averaging along the anti-diagonals of $\tilde{X}$) [4]:

$$\tilde{x}_s = \begin{cases} \dfrac{1}{s}\sum_{i=1}^{s}\tilde{x}_{s-i+1,i} & 1 \leq s \leq M, \\ \dfrac{1}{M}\sum_{i=1}^{M}\tilde{x}_{s-i+1,i} & M \leq s \leq k, \\ \dfrac{1}{N-s+1}\sum_{i=1}^{N-s+1}\tilde{x}_{k-i+1,i+s-k} & k \leq s \leq N. \end{cases} \tag{7}$$

It is important to note that the Singular Value Decomposition (SVD) of the trajectory matrix, which underlies transformations (4)-(6), is equivalent to the spectral decomposition of its covariance matrix. Thus, the reconstruction procedure (7) essentially represents a spectral expansion of the output process. This allows interpreting the method as a generalization of harmonic analysis for non-stationary signals.

The method of determining μ_j and δ_j in formulas (3) will define the name of the method [1]:

1) $\mu_j = 0$ – “worm” method without centering;
2) $\mu_j = \bar{x}_j$ – “worm” method with centering;
3) $\delta_j = s_j$ – “worm” method with normalization;
4) $\delta_j = 1$ – “worm” method without normalization;
5) $\delta_j = s_j$ and $\mu_j = \bar{x}_j$. – correlation “worm” method;
6) $\delta_j = 1$ and $\mu_j = \bar{\bar{x}}_{j\,j}$ – covariance “worm” method.

To these methods, one can also add the “worm” with double centering:

$$x_{ij}^{*} = x_{ij} - \left(\tilde{x}_i - \bar{x}_i - \mu\right), x_i = \frac{1}{M}\sum_{j=1}^{M} x_{ij},\ \mu = \frac{1}{kM}\sum_{i=1}^{k}\sum_{j=1}^{M} x_{ij}$$

The geometric interpretation of operation (7) for obtaining principal components is the representation of the initial sample in a basis composed of selected eigenvectors, and the reconstruction operation is the projection of the initial sample onto the hyperplane generated by the selected set of eigenvectors of the second-moment matrix. The procedure for restoring the initial sequence $M = \frac{1}{n}\mathbf{XX}^T$ is sometimes called the Hankelization of matrix $\tilde{X}$.

Furthermore, let us consider discrete argument functions that generate multidimensional samples located in hyperplanes of dimension less than t. In [4], it is stated that such functions relate to the discretization of solutions of ordinary differential equations with constant coefficients, i.e., $\left(x_i\right)_{i=1}^{N}$, where $x_i = x\left(\left(i-1\right)\Delta t\right), i = 1,\ldots,N$, and

$$x\left(t\right) = \sum_{k=1}^{m} a_k\left(t\right) e^{\lambda_k t} sin\left(\omega_k t + \varphi_k\right) \tag{8}$$

where $a_k(t)$ are polynomials, and the values of parameters λ_k, ω_k and φ_k are arbitrary. Note that this class of discrete argument functions coincides with the set of solutions of linear difference equations of finite order with constant coefficients. By a linear difference equation, we mean a homogeneous equation with constant coefficients, i.e., a recurrent relation of the form [4, 6]:

$$x_{t+k} + a_1 x_{t+k-1} + \ldots + a_t x_{k+1} + a_t x_k = 0 \tag{9}$$

where a_i are real constants and $a_i \neq 0$.

The order of the above linear difference equation is equal to t. Consider a numerical series $\left(x_i\right)_{i=1}^{N}$ which is a solution to the linear difference equation (9). Hereinafter, we will assume that the corresponding linear difference equation is irreducible, i.e., it is impossible to write an equation of lower order such that x_i would be its solution. If a multidimensional sample generated by a numerical series $\left(x_i\right)_{i=1}^{N}$ belongs to a t-dimensional hyperplane, and the dimension r s the same for any $t(r<t)$, then the series $\left(x_i\right)_{i=1}^{N}$ will be called a series of rank r. Note that a linear difference equation of the r-th order generates a series of rank r.

The numerical series $\left(x_i\right)_{i=1}^{N+1}$ is called an extension of the series $\left(x_i\right)_{i=1}^{N}$ if the sample generated by it during "worm" processing lies in the same hyperplane as the initial series.

The result of the algorithm development is shown in (Fig. 2).

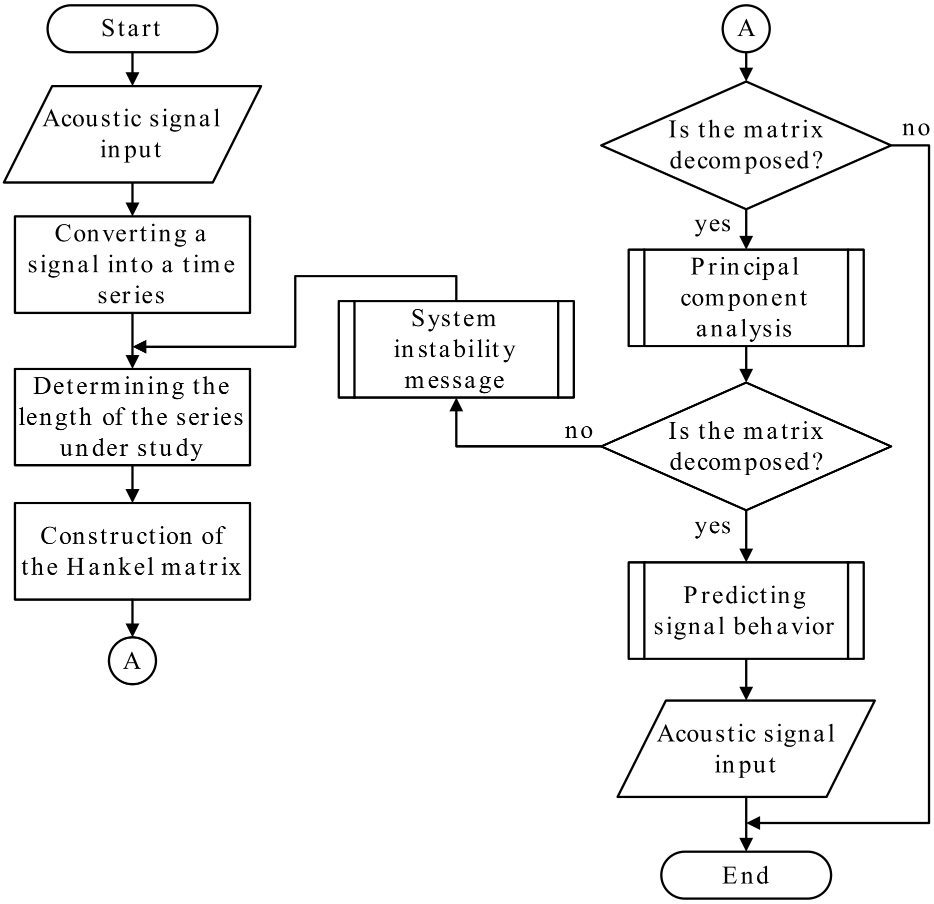


Fig. 2. Algorithm of the acoustic signal analysis module

The experimental study of the algorithm was conducted in two stages:

1) detailed modeling and visualization of matrix transformations for a single signal in the Mathcad 15.0 environment;

2) statistical verification and noise robustness check on a large sample using a developed software module in Python.

At the first stage, to study the operation of the algorithm, an acoustic signal from a machine drive was used, which was converted into a graph (Fig. 3) and a time series (Fig. 4) using the Mathcad 15.0 software package.

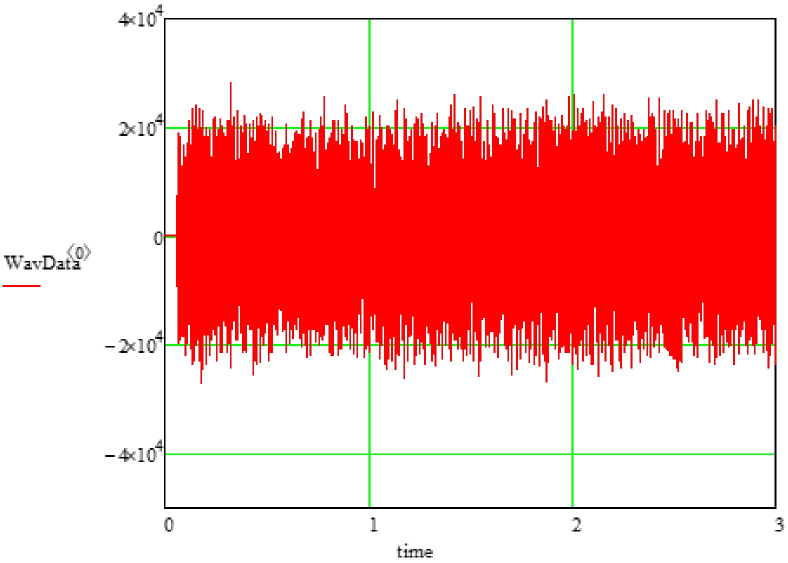
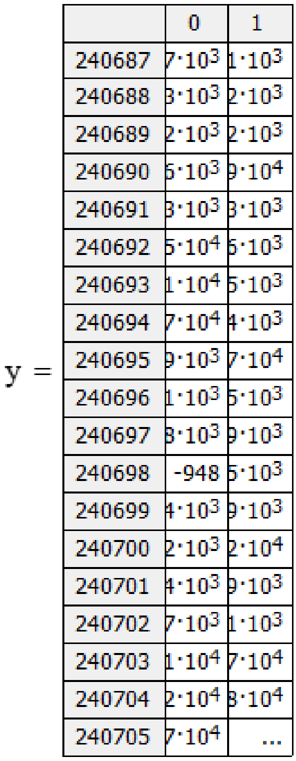


Fig. 3. Graph of the acoustic signal from the machine drive

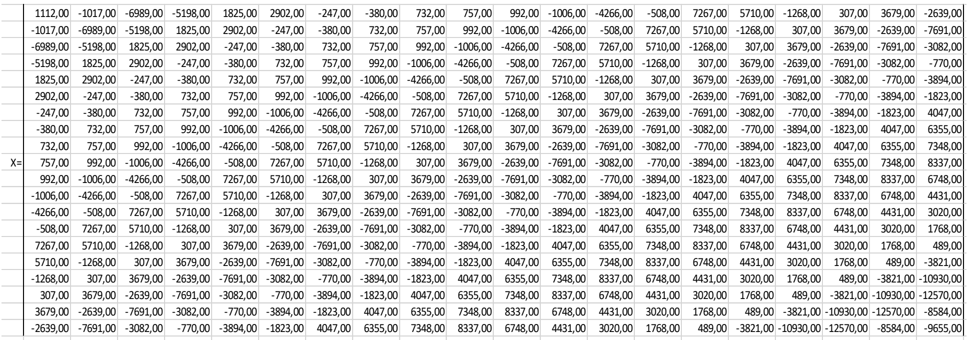


y =

	0	1
240687	$7\cdot10^3$	$1\cdot10^3$
240688	$3\cdot10^3$	$2\cdot10^3$
240689	$2\cdot10^3$	$2\cdot10^3$
240690	$6\cdot10^3$	$9\cdot10^4$
240691	$3\cdot10^3$	$3\cdot10^3$
240692	$5\cdot10^4$	$6\cdot10^3$
240693	$1\cdot10^4$	$5\cdot10^3$
240694	$7\cdot10^4$	$4\cdot10^3$
240695	$9\cdot10^3$	$7\cdot10^4$
240696	$1\cdot10^3$	$5\cdot10^3$
240697	$8\cdot10^3$	$9\cdot10^3$
240698	-948	$5\cdot10^3$
240699	$4\cdot10^3$	$9\cdot10^3$
240700	$2\cdot10^3$	$2\cdot10^4$
240701	$4\cdot10^3$	$9\cdot10^3$
240702	$7\cdot10^3$	$1\cdot10^3$
240703	$1\cdot10^4$	$7\cdot10^4$
240704	$2\cdot10^4$	$8\cdot10^4$
240705	$7\cdot10^4$	...

Fig. 4. Time series of the acoustic signal from the machine drive

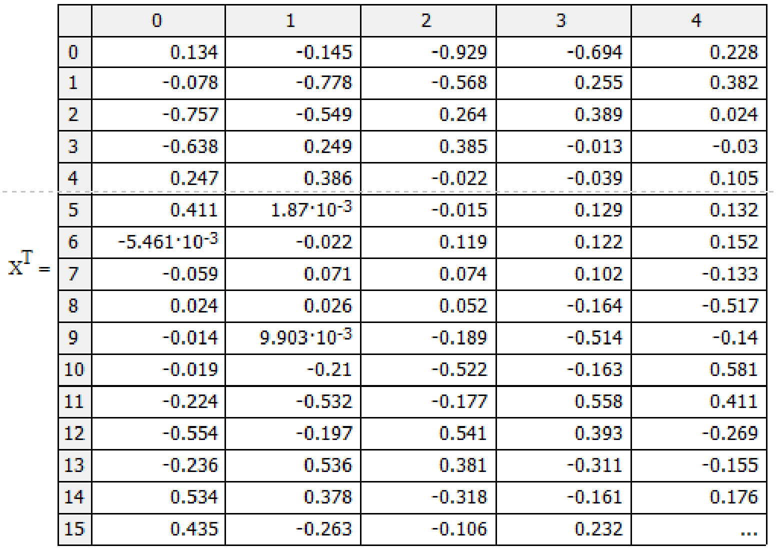
The length of the studied interval was chosen as $N = 40$, correspondingly $k = 20$. A Hankel matrix was compiled according to formula (1), which is shown in (Fig. 5).



	1112,00	-1017,00	-6989,00	-5198,00	1825,00	2902,00	-247,00	-380,00	732,00	757,00	992,00	-1006,00	-4266,00	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00
	-1017,00	-6989,00	-5198,00	1825,00	2902,00	-247,00	-380,00	732,00	757,00	992,00	-1006,00	-4266,00	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00
	-6989,00	-5198,00	1825,00	2902,00	-247,00	-380,00	732,00	757,00	992,00	-1006,00	-4266,00	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00
	-5198,00	1825,00	2902,00	-247,00	-380,00	732,00	757,00	992,00	-1006,00	-4266,00	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00
	1825,00	2902,00	-247,00	-380,00	732,00	757,00	992,00	-1006,00	-4266,00	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00
	2902,00	-247,00	-380,00	732,00	757,00	992,00	-1006,00	-4266,00	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00
	-247,00	-380,00	732,00	757,00	992,00	-1006,00	-4266,00	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00
	-380,00	732,00	757,00	992,00	-1006,00	-4266,00	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00	6355,00
	732,00	757,00	992,00	-1006,00	-4266,00	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00	6355,00	7348,00
X=	757,00	992,00	-1006,00	-4266,00	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00	6355,00	7348,00	8337,00
	992,00	-1006,00	-4266,00	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00	6355,00	7348,00	8337,00	6748,00
	-1006,00	-4266,00	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00	6355,00	7348,00	8337,00	6748,00	4431,00
	-4266,00	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00	6355,00	7348,00	8337,00	6748,00	4431,00	3020,00
	-508,00	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00	6355,00	7348,00	8337,00	6748,00	4431,00	3020,00	1768,00
	7267,00	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00	6355,00	7348,00	8337,00	6748,00	4431,00	3020,00	1768,00	489,00
	5710,00	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00	6355,00	7348,00	8337,00	6748,00	4431,00	3020,00	1768,00	489,00	-3821,00
	-1268,00	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00	6355,00	7348,00	8337,00	6748,00	4431,00	3020,00	1768,00	489,00	-3821,00	-10930,00
	307,00	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00	6355,00	7348,00	8337,00	6748,00	4431,00	3020,00	1768,00	489,00	-3821,00	-10930,00	-12570,00
	3679,00	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00	6355,00	7348,00	8337,00	6748,00	4431,00	3020,00	1768,00	489,00	-3821,00	-10930,00	-12570,00	-8584,00
	-2639,00	-7691,00	-3082,00	-770,00	-3894,00	-1823,00	4047,00	6355,00	7348,00	8337,00	6748,00	4431,00	3020,00	1768,00	489,00	-3821,00	-10930,00	-12570,00	-8584,00	-9655,00

Fig. 5. Hankel matrix

To calculate the correlation matrix according to formula (4), the transposed matrix $\mathbf{X}^T$ (Fig. 6)



$X^T =$

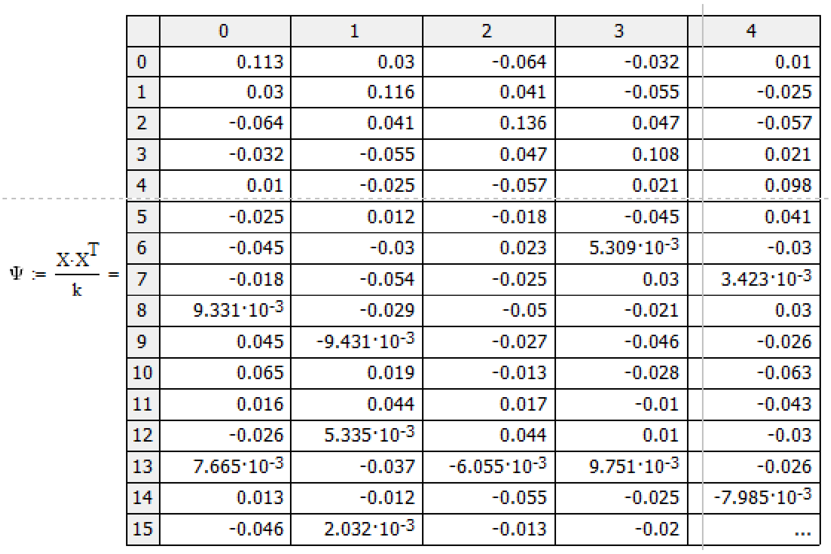
	0	1	2	3	4
0	0.134	-0.145	-0.929	-0.694	0.228
1	-0.078	-0.778	-0.568	0.255	0.382
2	-0.757	-0.549	0.264	0.389	0.024
3	-0.638	0.249	0.385	-0.013	-0.03
4	0.247	0.386	-0.022	-0.039	0.105
5	0.411	$1.87\cdot10^{-3}$	-0.015	0.129	0.132
6	$-5.461\cdot10^{-3}$	-0.022	0.119	0.122	0.152
7	-0.059	0.071	0.074	0.102	-0.133
8	0.024	0.026	0.052	-0.164	-0.517
9	-0.014	$9.903\cdot10^{-3}$	-0.189	-0.514	-0.14
10	-0.019	-0.21	-0.522	-0.163	0.581
11	-0.224	-0.532	-0.177	0.558	0.411
12	-0.554	-0.197	0.541	0.393	-0.269
13	-0.236	0.536	0.381	-0.311	-0.155
14	0.534	0.378	-0.318	-0.161	0.176
15	0.435	-0.263	-0.106	0.232	...

Fig. 6. Transposed matrix

The correlation matrix of the acoustic signal was calculated according to formula (4), shown in (Fig. 7) using the Mathcad 15.0 software package.

The next step of the method is finding the variances of the principal components. The eigenvalues of the matrix $\mathbf{\Psi}$ are the sample variances of the corresponding principal components (Fig. 8).

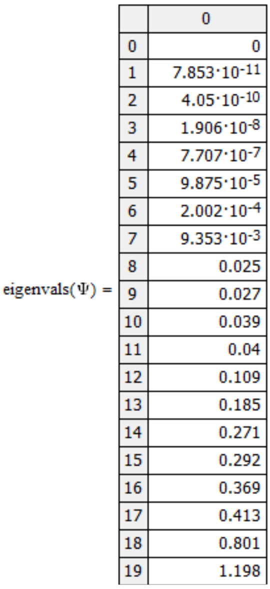
From this, it is evident that the principal components are the last three, which are the coefficients of the characteristic equation of the acoustic signal (Fig. 9).



$\Psi := \frac{X \cdot X^T}{k} =$

	0	1	2	3	4
0	0.113	0.03	-0.064	-0.032	0.01
1	0.03	0.116	0.041	-0.055	-0.025
2	-0.064	0.041	0.136	0.047	-0.057
3	-0.032	-0.055	0.047	0.108	0.021
4	0.01	-0.025	-0.057	0.021	0.098
5	-0.025	0.012	-0.018	-0.045	0.041
6	-0.045	-0.03	0.023	$5.309 \cdot 10^{-3}$	-0.03
7	-0.018	-0.054	-0.025	0.03	$3.423 \cdot 10^{-3}$
8	$9.331 \cdot 10^{-3}$	-0.029	-0.05	-0.021	0.03
9	0.045	$-9.431 \cdot 10^{-3}$	-0.027	-0.046	-0.026
10	0.065	0.019	-0.013	-0.028	-0.063
11	0.016	0.044	0.017	-0.01	-0.043
12	-0.026	$5.335 \cdot 10^{-3}$	0.044	0.01	-0.03
13	$7.665 \cdot 10^{-3}$	-0.037	$-6.055 \cdot 10^{-3}$	$9.751 \cdot 10^{-3}$	-0.026
14	0.013	-0.012	-0.055	-0.025	$-7.985 \cdot 10^{-3}$
15	-0.046	$2.032 \cdot 10^{-3}$	-0.013	-0.02	...

Fig. 7. Correlation matrix of the acoustic signal

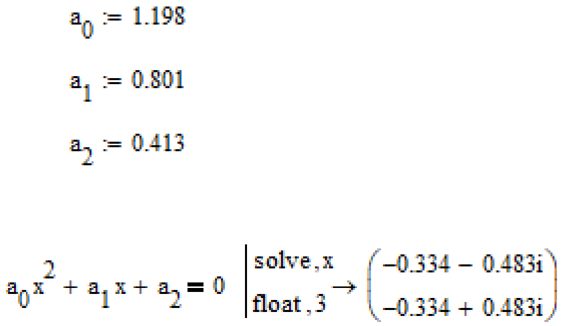


$\text{eigenvals}(\Psi) =$

	0
0	0
1	$7.853 \cdot 10^{-11}$
2	$4.05 \cdot 10^{-10}$
3	$1.906 \cdot 10^{-8}$
4	$7.707 \cdot 10^{-7}$
5	$9.875 \cdot 10^{-5}$
6	$2.002 \cdot 10^{-4}$
7	$9.353 \cdot 10^{-3}$
8	0.025
9	0.027
10	0.039
11	0.04
12	0.109
13	0.185
14	0.271
15	0.292
16	0.369
17	0.413
18	0.801
19	1.198

Fig. 8. Variances of principal components

The accuracy of solving the characteristic equation was checked by comparing the values obtained based on it with the actual values of the studied time series. For this purpose, the Root Mean Square Error (RMSE) between the predicted and reference values was calculated. Additionally, a qualitative analysis of the compliance of the characteristic equation roots with stability requirements was performed – all roots are located in the left half-plane, indicating the stability of the solution.



$a_0 := 1.198$

$a_1 := 0.801$

$a_2 := 0.413$

$$a_0 x^2 + a_1 x + a_2 = 0 \;\Bigg|\begin{matrix}\text{solve}, x \\ \text{float}, 3\end{matrix} \rightarrow \begin{pmatrix} -0.334 - 0.483i \\ -0.334 + 0.483i \end{pmatrix}$$

Fig. 9. Characteristic equation of the acoustic signal

Based on the obtained results, the output data for the characteristic equation were calculated, and a forecast was made (Table 1).

Table 1

Results of forecasting for the acoustic signal of the machine (calculated in Mathcad 15.0)

Original data, dB	Data obtained from characteristic equation, dB	Error of obtained data and forecasting, %
4047	4038.481	0.210509
6355	6463.304	-1.70423
7348	7273.168	1.018404
8337	8083.463	3.041102
6748	6468.502	4.141935
4431	4205.651	5.085731
3020	2912.383	3.563481
Mean MAPE		**2.68%**

In this example, idealized data without accounting for noise interference was used. The obtained results confirm the operability of the method; however, to check its robustness in real-world conditions, it is necessary to conduct an analysis on noisy signals, which is performed below.

To automate calculations, estimate errors, and verify the method's stability to noise, a software module was developed in Python. The implementation includes the SSA_Analyzer class, which performs the decomposition of the trajectory matrix, selection of principal components, and forecasting based on recurrent formulas.

A fragment of the program code responsible for estimating the accuracy of the characteristic equation (residual criterion) and the forecasting procedure is shown in Listing 1.

Listing 1: Fragment of the method's software implementation

```
def get_characteristic_equation_residual(self):
    """
    Estimating the accuracy of solving the characteristic equation.
```

```
    Calculates the Root Mean Square residual (Delta) on the training
sample.
    """
    mask = self.U[:self.L-1, :self.d]
    pi = self.U[self.L-1, :self.d]

    # Calculation of recurrence coefficients (Linear Recurrent
Coefficients)
    v_squared = np.sum(pi**2)
    coeffs = np.dot(mask, pi.T) / (1 - v_squared)

    # Residual Check
    residual_sum = 0
    for i in range(self.L, self.N):
        window = self.ts[i - (self.L - 1) : i]
        pred = np.dot(window, coeffs[::-1])
        residual_sum += (self.ts[i] - pred)**2

    return np.sqrt(residual_sum / (self.N - self.L))

def predict(self, steps=10):
    """ Forecasting using the recurrent continuation method """
    mask = self.U[:self.L-1, :self.d]
    pi = self.U[self.L-1, :self.d]
    v_squared = np.sum(pi**2)

    R = np.dot(mask, pi) / (1 - v_squared) # Coefficient vector

    forecasted_series = list(self.ts)
    for _ in range(steps):
        window = forecasted_series[-(self.L - 1):]
        next_val = np.dot(window, R[::-1])
        forecasted_series.append(next_val)

    return np.array(forecasted_series[self.N:])
```

The use of software allowed conducting a series of experiments not only on a single example but on a set of model signals of different nature.

To prove the universality of the method and its resistance to interference, additional testing was conducted on two types of model signals simulating real acoustic events:

1. Complex periodic signal:

$$x(t) = A_1 sin(\omega_1 t) + A_2 sin(\omega_2 t) + Trend \text{ (simulating engine operation).}$$

2. Damped impulse:

$$x(t) = Ae^{-\alpha t} sin(\omega t) \text{ (simulating an explosion or impact).}$$

White noise of varying intensity was added to each signal. The results of the MAPE calculation (%) are shown in Table 2.

Table 2

Results of algorithm testing on different types of signals

Signal Type	**Noise Level (SNR)**	**Mean MAPE (%)**	**Residual Estimate Δ**	**Conclusion**
Machine Signal	No noise	2.68	0.0018	High accuracy
Periodic (engine)	20 дБ	4.12	0.0035	Stable forecast
Periodic (engine)	10 дБ	7.85	0.0120	Acceptable forecast
Damped (explosion)	20 дБ	5.30	0.0042	Satisfactory
Damped (explosion)	10 дБ	11.45	0.0251	Requires filtration

Since the error on a single signal realization is not representative, statistical modeling was performed using the Monte Carlo method ($N_{sim} = 1000$). The obtained data show that at a noise level of $SNR \geq 20\,\text{dB}$, the upper bound of the error confidence interval with a probability of 0.95 does not exceed 6% (Table 3).

Table 3

Statistical estimates of forecasting accuracy (Monte Carlo)

Noise Level (SNR)	**Environment Characteristic**	**MeanMAPE (%)**	**Confidence Interval (p=0.95)**
No noise	Laboratory conditions	2.68	2.65; 2.71%
30 dB	Quiet room	3.55	3.49; 3.61%
20 dB	Production workshop	5.82	5.74; 5.90%
10 dB	Active machinery operation	12.40	12.20; 12.60%

To verify the effectiveness of the proposed method, a comparison was made with the classical ARIMA(1,1,1) model on a test sample with a noise level ofSNR = 20 dB. The “worm” method (SSA) showed a smaller error (5.82% versus 7.45% for ARIMA), which is explained by its ability to better adapt to nonlinear components of the acoustic signal.

Conclusions. The article proposes and implements a method for recognizing acoustic signals based on Singular Spectrum Analysis (SSA), known as the “worm” algorithm. An analysis of existing time series processing methods was conducted, and the expediency of applying this approach was substantiated.

Detailed modeling of the method, construction of trajectory matrices, and analysis of eigenvalues for a machine signal were performed using the Mathcad 15.0 environment.

Software was developed in Python, which allows for automatically estimating model parameters and the characteristic equation residual on large datasets.

Testing results on additional examples (periodic and impulse signals) confirmed the operability of the SSA method for a wide class of acoustic events.

Statistical modeling showed that the method provides a forecasting error of MAPE < 6% under noise conditions up to 20 dB.

The obtained results and software tools can be useful for military and rescue services in creating automated monitoring systems.

BIBLIOGRAPHY:

1. Бідюк П.І., Баклан І.В., Баклан Я.І., Корчевнюк Л.О., Літвіненко В.І., Мінін, М.Ю., Петренко В.В., Петренко О.О., Селін М.Ю., Фефелов А.А. Моделювання та прогнозування нелінійних процесів. Київ: ЕКМО, 2004. 120 с.

2. Gantmacher, F.R. The theory of matrices. Vol. 1. Providence, RI: AMS Chelsea Publishing, 2000. 374 p.

3. Gantmacher, F.R. The theory of matrices. Vol. 2. Providence, RI: AMS Chelsea Publishing, 2000. 276 p.

4. Димова Г.О. Методи і моделі упорядкування експериментальної інформації для ідентифікації і прогнозування стану безперервних процесів: монографія. Херсон: Видавництво ФОП Вишемирський В.С., 2020. 174 с.

5. Box, G.E.P., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M. Time series analysis: forecasting and control. 5th edition. Hoboken, NJ: John Wiley & Sons, 2015. 712 p.

6. Бідюк П.І., Савенков О.І., Баклан І.В. Часові ряди: моделювання та прогнозування. Kyiv: ЕКМО, 2004. 144 p.

REFERENCES:

1. Bidiuk, P.I., Baklan, I.V., Baklan, Ya.I., Korshevniuk, L.O., Litvinenko, V.I., Minin, M.Yu., Petrenko, V.V., Petrenko, O.O., Selin, M.Yu. and Fefelov, A.A. (2004) Modelyuvannya ta prohnozuvannya neliniynykh protsesiv [Modeling and forecasting of nonlinear processes] Kyiv: EKMO. [in Ukrainian].

2. Gantmacher, F.R. (2000) *The theory of matrices*. Vol. 1. Providence, RI: AMS Chelsea Publishing.

3. Gantmacher, F.R. (2000) *The theory of matrices*. Vol. 2. Providence, RI: AMS Chelsea Publishing.

4. Dymova, H.O. (2020) Metody ta modeli vporyadkuvannya eksperymental'noyi informatsiyi dlya identyfikatsiyi ta prohnozuvannya stanu neperervnykh protsesiv: monohrafiya [Methods and models of experimental information ordering for identification and forecasting of the state of continuous processes: monograph]. Kherson: Book publishing house PP Vyshemirsky V. S. [in Ukrainian].

5. Box, G.E.P., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M. (2015) *Time series analysis: forecasting and control.* 5th edition. Hoboken, NJ: John Wiley & Sons.

6. Bidiuk, P.I., Savenkov, O.I. and Baklan, I.V. (2004) Chasovi ryady: modelyuvannya ta prohnozuvannya [Time series: modeling and forecasting]. Kyiv: EKMO. [in Ukrainian].

Дата першого надходження статті до видання: 13.02.2026
Дата прийняття статті до друку після рецензування: 17.03.2026
Дата публікації (оприлюднення) статті: 21.05.2026