UDC 517.9

I. DOROSHENKO
Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine,
e-mail: *i.doroshenko@chnu.edu.ua*.

O. KNOPOV
V.M. Glushkov Institute of Cybernetics of NAS of Ukraine, Kyiv, Ukraine,
e-mail: *a.knopov@gmail.com*.

L. VOVK
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine,
e-mail: *lilyvovk@gmail.com*.

MATHEMATICAL MODELS OF EXTREME MODES IN ECOLOGICAL SYSTEMS

Abstract. The authors consider the basic stages of creating a computer system for decision-making support in ecological security. The key factors of environmental pollution caused by ecological and technological catastrophes and typical failures in the operation of hazardous facilities are analyzed. Methods for quantitative estimation of the risk function are proposed. Ecological and mathematical models are created that make it possible to assess the current state of the environment, to make predictions, and provide a comprehension of the process under consideration.

Keywords: model, time series, trend, maximal likelihood method (MLM), least squares method (LSM), least modules method (LMM), minimax method (MMM).

Under condition of scientific and technical progress rapid development of the level of antropogenical and technical influence on the environment essentially increased. Thus, many biosphere parameters reached critical indices and so are becoming dangerous to human life and to the mankind existence in whole.

The problem arising in this connection make it necessary working out the system of decision making support in ecological security. This is one of the most difficult multi-factorial problems. Its decision is possible only with the use of system analysis methods.

Let's consider the basic stages of the creation of computer system of decision making support. The first stage consists in the analysis of basic factors of pollution after the ecological and technological catastrophes, typical failures of project, exploitation of dangerous objects and so on. The second stage involves the risk function quantitative estimation and creation of the ecological and mathematical model allowing to estimate current condition of nature and to give forecast for future. The main goal of the model design is to achieve the comprehension of the process under consideration.

Give an example taken from [1] dealing with the analysis of soils' main properties. Following [1], it's enough to consider such ones: l) acids' neutralization; 2) adsorption of phosphorus and toxic materials; 3) oxygen donating.

Let us review each one.

1) **Acids' neutralization**. This capacity is substantial for maintaining the stability of fresh water habitats necessary for fish and other aquatic life protecting from acidity fluctuations. Soils' ability to neutralize acids essentially depends on their types.

Analyzing the situation over the last 100-150 years in the industrialized regions of North America and Europe the conclusion of the soils' buffering capacities large perturbations can be done. This is because of the great amount of sulfuric and nitric acidic rains resulting from burning of organic fuels. But the effect of such impacts is not uniform for different regions because of the large variability in soil type. [1] contains information about the acidification of lakes. The rate of lake response to acid inputs depends on two watershed soils fundamental qualities — the ability to retain sul-

fate and the ability to supply base cations. It is necessary to know the history of acid emissions, the time trends in pH and in the exhaustion of the watershed soils buffering capacity. Unfortunately, in literature there is little information about it and data available are pure. Besides, the experiments in this sphere are expensive and often impossible. We have to use the information available. That's why it's necessary to develop methods allowing to make inference using insufficient amount of information such as statistical methods of small samples processing.

2) **Adsorption of phosphorus and toxic materials**. It's known that chemicals being added to soils falls into 4 groups: i) those getting into ground and then in surface waters; ii) taken up by plants; iii) volatilized into the atmosphere; iv) stored in soil. In the cases i)- iii), when chemicals are input in the environment, their effects are easy to observe even within a short period after pollution. The fourth way is the most interesting for investigation because of nonlinear and time-delayed effects arising. The potential danger of chemicals consists in their ability to accumulate in soils thus the *threshold* for the soil's capacity to hold the chemical can be reached.

3) **Oxygen donating**. The most efficient mechanism of producing biochemical energy necessary for any organic life is oxidation of organic carbon to carbon dioxide by molecular oxygen. Thus, oxygen donating is extremely essential for the life in biosphere.

Thus, summarizing data concerning three main soils properties we meet the necessity to design a model (or a series of models) answering the following questions:

1. What level of soil acidity is harmful for human health? When will it be accessed (or whether it is accessed) for some region?

2. What is the law according to which the soil properties change?

3. How to estimate the moment of *threshold* for the soil's capacity to hold the chemical reaching?

MODELS

1. This kind of problems arises in study of different natural and human activity connected phenomena. The models we design have to reflect the phenomenon under consideration effectively enough and answer the questions we put on each concrete case. Having a possibility to describe model by analytic expressions we can considerably advance in it. On the other hand, we can aspire to design the empirical model based on the summing up of the accessible observation set. We can definite this process characteristics with the help of modeling or with estimation on the observation background. Note that when we investigate so difficult process, we can't get satisfactory determinate model as a rule. Therefore, it is natural (it is described in detail in [2]) to consider the investigated process as sum of two parts: determinate $g(t)$ and random $W(t)$. Thus, the observations model is:

$$Y(t) = g(t) + W(t),$$

where $g(t)$ is the function of unknown kind containing the unknown parameters: $g(t) = g(t, \alpha)$ or $g(t)$ is the function of unknown kind belonging to some class of functions. It is necessary to estimate either α or the whole function $g(t)$. This kind of problem we will name trend extracting.

2. Another class of problems will be given by the following model:

$$f(y(t)) = g(\xi(t-1)) + W(t), \qquad (1)$$

where $\xi(t-1)$ is observation vector to time moment $t-1$, $W(t)$ is a subsequence of independent random magnitudes, f and g are some functions. The relation (1) is a stochastic difference equation. These equations theory is investigated in Kalman's works [3, 4] and others. The main task is to acquire the algorithm to

find forecast of the process $y(t)$ based on some observation vector $\xi(t)$ within the time moment $t-1$. This problem is investigated so fully in case when the outcoming signal y satisfies the difference equation:

$$A(D)\,y(t) = B(D)W(t)\,,\ A(0) = B(0) = 1\,, \tag{2}$$

and $y(t)$, in general, isn't observed immediately, in fact we observe the process with white noise:

$$x(t) = y(t) + \eta(t)\,,$$

where $\eta(t)$ is some discrete white noise.

Operators D, $A(D)$, and $B(D)$ are determined by equalities

$$Dy(t) = y(t-1)\,,\ A(D) = I - \sum_{j=1}^{m_1} A_j D^j\,,\ B(D) = \sum_{j=1}^{m_2} B_j D^j\,,\ m_1 > m_2\,.$$

This kind of models is studied in detail in [5, 6] and others.

Likewise, we can consider the continuous analogy of the model (2), when operator D is given below:

$$Dy(t) = \frac{\partial}{\partial t}\, y(t).$$

The equation of kind (2) includes wide enough class of models, between them autoregression model and autoregression model with moving average. Such models well enough describe processes connected with risk estimation for ecology end economic systems. For example, in phosphate storing in soil forecasting, when data in previous time moments are known, the difference equation of kind (2) is a natural enough model of the real process. Some difficulties may arise because of the trend is, in general, time-dependent, so coefficients of $A(D)$ may be time-dependent too. The forecast finding methods are developed in such case as well [7]. But unlike the previous case, here the Wiener–Kolmogorov method is used and the forecast problem resolving is being reduced to some difference or differential equations resolving.

Note, that sometimes in some substances spreading in soil forecasting t parameter is convenient to be treated as vector one. Then the model under investigation can be written as (2), but D operator will look like

$$D = \frac{\partial^m}{\partial t_1 \ldots \partial t_m}\,,\ t = (t_1, \ldots, t_m)\,,$$

and $W(t)$ is a multidimensional white noise.

Such systems are studied in [8].

Certainly, t is not interpreted here as a time parameter but as averaged coordinate of some region. Harmful substances forecast is made for neighboring regions.

Let's do some more essential notes about the models describing soil pollution, chemical substances and acids storing in soils etc. processes. As our purpose consists in estimates connected with chemicals, acids, heavy metals presence in soil obtaining it's natural to resolve our problems using the mathematical random field methodology. We will study so called geochemical field, where the content of some chemical substance as coordinate and time function will be treated. Thus, the random geochemical field will be function of time t and space variable x: $\xi(t, x) = \xi(\omega, t, x)$, where $\omega \in \Omega$ is an element of the probability space Ω, $t \in T \subseteq R$, $x \in R^m$, $m \geq 1$. If $m = 2$, chemicals concentration on an area is considered, if $m = 3$ — in the space. In experimental investigations we often have at our disposal only average meaning of point spatial variable within some area or volume at time moment t. For instance, average

meaning of the function $\xi(\omega, t)$ within the area S at the point $x \in S$ at time moment t we describe as follows:

$$g(t, x) = \frac{1}{S} \int\limits_{x+h \in S} \xi(x + h)dh.$$

The same way we can determine average meaning of the function $\xi(\omega, t)$ within the volume V with center of gravity at the point $x \in S$ at time moment t:

$$g(t, x) = \frac{1}{V} \int\limits_{x+h \in V} \xi(x + h)dh.$$

Let us list natural assumptions necessary to be done for the correct problem solution if we use average meaning $g(t, x)$ of the function $\xi(t, x)$.

i) **Random field** $\xi(t, x)$ **uniformity** on x in the restricted or broad sense [9]. This property develops in the fact that at any point of the field, random function has the same average characteristics. In other words, we have possibility to repeat experiments endlessly, to obtain as much as desired of random function realizations and so to make valid corresponding statistical inference. This hypothesis will be unfit if regular change of $\xi(t, x)$ meaning when observation point recedes from the central part of the field is observed. Sometime it is convenient to consider instead of uniform field the field with stationary increments, i.e. joint distribution of random values

$$\eta(t, x_k) = \xi(t, x_{k-1}) - \xi(t, x_k)$$

under fixed t stays constant when variables x_k are simultaneously moved on value h.

ii) **Continuity.** Usually, the change of any soil properties takes place slowly and continuously. That's why this condition looks natural enough.

Apparently, properties 1 and 2 are basic for natural transition from $\xi(t, x)$ function observations to its average characteristics $y(t, x)$. So, if we fix some area D_x by any point $x \in D$ we can go to the forecast estimation problem for any random function $y(t) = y(t, x)$ omitting index x and implying that problem is being solved for some concrete area or region.

Thus, we can consider the case when the sequence of random value observations $y(t_1), \ldots, y(t_n)$, for which some statistical inference should be made. We shall call the $y(t_1), \ldots, y(t_n)$ sequence time series.

ANALYSIS

Before we solve estimation, forecast, distribution finding problems for some functionals we have to make preliminary analysis of time series. In our opinion it consists of the items below:

1a. Observation independence checking.

2a. Trend presence hypothesis checking.

3a. Observed value distribution hypothesis checking.

Then we do more detailed time series analysis. It consists of such items:

lb. Finding the class of models being subject to investigation.

2b. Model parameter statistical estimation.

3b. Time series values at following time moments forecast.

It's no need to stop in detail at the preliminary statistical analysis of time series. These problems are widely elucidated in many textbooks and monographs on mathematical statistics. We made statistical analysis for data from [1]. Statistical information was processed by program complex [10]. The main conclusion consists in the fact that as a rule, each time series from [1] has a trend. Observation errors can be treated as independent, but normality does not always take place.

Problem of trend extracting. Let us consider now the problem of trend extracting. First note some specific properties of the mentioned above time series.

1. As it is pointed in [1] and other for many regions the lack or absence of the information takes place.

2. Noise distribution is not always known.

Taking into account these facts and having made the preliminary analysis of time series, we can use different methods for trend extraction. The most known are: the maximal likelihood method (MLM), the least squares method (LSM), the least modules method (LMM), minimax method (MMM). We shall later describe each one, now we mark only that in our case the most preferrable is LMM. The least modules estimators (LME) are referred to as robust estimators, i.e. estimators having advanced stability with respect to errors distributions law changes, rough error presence etc. Besides the LME are considerably more effective in small sample case. At the last time the series of investigations of the LMM qualitative behavior has been done. First the consistency of estimators should be stressed. Simple and easy enough checkable consistency conditions are given in [11, 12]. It should be noted that general enough models where the parameter is contained in the regression function generally speaking nonlinearly and the noise may have a form of dependent at different time moments are studied. Besides the statement of asymptotic normality is proved under natural enough conditions.

One more circumstance should be noted. When small samples are under consideration, the asymptotic properties such as consistency, asymptotic normality etc. can't be the criteria of estimators' quality. In this case the estimators' stability (robustness) with respect to observations distribution laws and the errors variance are primary. In this respect here in the case of unknown distribution we should prefer LMM and MMM. In the case of large samples, we certainly should seek estimators' consistency and asymptotic normality.

1. **The maximal likelihood method** (**MLM**) is most preferable because of its properties in the cases when noise distribution is known. There are classical results widely known for specialists in statistics, consisting in the facts that MLM estimators are under some natural conditions consistent, asymptotically normal and effective. But their calculations imply great difficulties, that's why it is not often used in practice. We would not describe at length these estimators. In [2, 4] their properties are spelled out in detail.

2. **The least squares method** (**LSM).**

The problem formulation is as follows. Consider the estimation problem in whole non-linear parameters which is contained in regression function. The observation model is given below:

$$x_j = g(t_j, \alpha) + \xi_j, \ \ j = \overline{1, n}, \tag{3}$$

where $\alpha \in A$ is a vector parameter belonging to some admissible multitude A, is a regression function, ξ_j are random errors of observations.

The observations may be independent or dependent but should constitute stationary in narrow sense sequence of random values with unknown distribution function which is not dependent on α.

It is necessary on x_j, $j = \overline{1, n}$, observations (3) to estimate α.

The random vector α_n minimizing the sum of residuals squares

$$\alpha_n = \arg\min_{u \in A} L_n(u), \ \ L_n(u) = \sum_{j=1}^{n} [x_j - g(t_j, u)]^2$$

is named the least squares estimator (LSE) of vector α.

Due to the criterion simplicity and the fact that it is well adapted to calculating procedures, LSM is the most widespread and well-studied method. It coincides with MLM

in the case of Gaussian observations errors. The most convenient case for calculations is one when unknown parameter is contained in the regression function linearly. Here we won't give example of calculation algorithms and consistence, asymptotic normality and asymptotic efficiency statements, only refer to some works [2, 11]. Turning to the time series describing elements distribution in the soil we can make some conclusions about the trend character. For instance, from Fig. 6 [1] it is seen that the trend can be approximated by straight line, for Fig. 11 [1] approximating curve is parabola etc. But a linear relation can't always be obtained. That's why the LSE properties investigations when nonlinear dependence on parameters takes place become actual. Yet in this case deep investigations of estimators' properties are made as well. But estimators finding implies essential calculating difficulties.

Note that even in linear case the calculating schemes suffer from such grave shortcoming as the necessity to save in memory and to process the great amounts of numerical information. The recursion computing schemes for least square estimators described in [13, 14] are devoid of these shortcomings. Their essence consists in following. Let the following random sequence is observed

$$\eta^i = \alpha\varphi(i) + \xi^i,\ i = 1, 2, \ldots,$$

where $\varphi(i)$ is known function, α is unknown parameter, ξ^i is independent random values sequence. Iterative procedure for the stochastic approximation parameter α estimation when $x^0 = 0$ looks as follows:

$$x^{s+1} = x^s + \frac{1}{\sum\limits_{i=1}^{s+1} \varphi^2(i)} [\eta^{s+1}\varphi(s+1) - \varphi^2(s+1)x^s]. \tag{4}$$

Thus, for the least squares estimators the iterative procedure enabling us to calculate parameter estimator by $(s+1)$ observations when the LSE basing on s observations and the observation at the time moment $(s+1)$ are known is determined. To make the picture complete, we give the statement about the iterative procedure (4) convergence with probability one for rather general case of nonlinear regression.

Suppose that the random sequence $\{y_k, k \geq 1\}$ is observed:

$$y_k = m_k(x^*) + x^k,\ x^*,\ z^k,\ m_k(x^*) \in R^N,$$

where z^k is a sequence of, generally speaking, dependent random vectors with $E(z^k) = 0$, $m_k(x^*)$ is a determined sequence depending on an unknown parameter x^*. The problem consists in estimating x^* using observations y_k. We use the following iterative procedure for estimating x^*:

$$x^{s+1} = x^s - \beta_s \xi^s,\ s \geq 1,\ x^l \in R^N,$$

where $\{c_s, s \geq 1\}$, $\{\beta_s, s \geq 1\}$, are the sequences of positive numbers, $\inf c_s > 0$. The following statement takes place [15]:

Theorem 1. Let the conditions are satisfied:

1. $u(s, x) \| m_s(x) - m_s(x^*) \| \leq k(1 + \| x \|)$,
2. $u(s, x)(1 + \| x \|) \leq k$,
3. Sequences $\{c_s, s \geq 1\}$, $\{\beta_s, s \geq 1\}$, $\{z_s, s \geq 1\}$, satisfy conditions:

a) $\sum\limits_{s=1}^{\infty} \beta_s c_s = \infty$,

b) $\sum_{s=1}^{\infty} \beta_s^2 c_s^2 < \infty$,

c) $\sum_{s=1}^{\infty} \beta_s c_s E \left\| z^s / \sigma_{z^s} \right\| < \infty$,

where σ_{z^s} is a σ-algebra, generated by random values $z^1, \ldots, z^s$.

d) For any $\varepsilon > 0$ $\inf_{s \geq 1} \inf_{\varepsilon < \|x - x^*\| < \varepsilon^{-1}} \left\langle m_s(x) - m_s(x^*), x - x^* \right\rangle > 0$, where $\langle \cdot \rangle$ is a scalar product in R^N, $\|x\| = \langle x, x \rangle^{1/2}$.

Then $\|x^s - x^*\| \to 0$ with probability 1, i. e. x^s is the strongly consistent estimator of x^*.

3. **The least modules method (LMM).** The least modules estimator (LME) of parameter α is a random vector α_n, minimizing residuals modules sum

$$\alpha_n = \arg\min_{u \in A} \breve{L}_n(u), \ \breve{L}_n(u) = \sum_{j=1}^{n} |x_j - g(t_j, u)|.$$

The LME coincide with the maximal likelihood estimators in cases when noise has Laplace distribution. They are referred to as robust estimator, i.e. estimators having advanced stability with respect to errors distribution law changes, rough errors presence etc. Moreover, the LME are considerably more effective in small sample case. At the last time the series of investigations of the LMM qualitative behavior has been done. First the consistency of estimators should be stressed. Simple and easy enough checkable consistency conditions are given in [11, 12]. It should be noted that general enough models where the parameter is contained in the regression function generally peaking nonlinearly and the noise may have a form of dependent at different time moments are studied. Besides in natural enough conditions the statement of asymptotic normality is proved.

One more circumstance should be noted. When small samples are under consideration, the asymptotic properties such as consistency, asymptotic normality etc. can't be the criteria of estimator quality. In this case the estimator stability (robustness) with respect to observations distribution law and the error variance are primary. In this respect here in the case of unknown distribution we should prefer LLM and MMM. In the case of large samples, we certainly should seek estimates consistency and asymptotic normality.

As an illustration we give some statements about the asymptotic behavior of the LME.

Theorem 2. Let we have the following observation model:

$$x_i = \varphi_i(\theta_0) + \xi_i, \ i = \overline{1, n}, \ \theta_0 \in K,$$

where K is a compact subset from R^l, ξ_i are equally distributed symmetric independent random values, $M_{\theta_0} \xi_i = 0$, $M_{\theta_0} \xi_i^2 < \infty$, and the conditions below are satisfied:

1) $\overline{\lim_{n \to \infty}} \sup_{\theta_1 - \theta_2 \in K} [\varphi_n(\theta_1, \theta_2) - \varphi(\theta_1, \theta_2)] \leq 0$,

where $\varphi_n(\theta_1, \theta_2) = \frac{1}{n} \sum_{i=1}^{n} |\varphi_i(\theta_1) - \varphi_i(\theta_2)|$,

and $\varphi(\theta_1, \theta_2)$ is uniformly continuous at the diagonal $\theta_1 = \theta_2$.

2) $\overline{\lim_{n \to \infty}} \varphi_n(\theta_1, \theta_2) \leq \infty$.

3) $\lim_{n \to \infty} \inf_{\theta \in K \setminus \delta(\theta_0, \varepsilon^{n-1})} \left\{ \frac{1}{n} \sum_{i=1}^{n} E_{\theta_0} |x_i - \varphi_i(\theta)| - \frac{1}{n} \sum_{i=1}^{n} E_{\theta_0} |x_i - \varphi_i(\theta_0)| \right\} > 0$,

where $\delta(\theta_0, r) = \{\theta \in \mathrm{K}, \ |\theta - \theta_0| \leq r\}$.

Let $\theta_n = \underset{\theta \in K}{\operatorname{argmin}} \sum_{i=1}^{n} |x_i - \varphi_i(\theta)|$. Then $P_\theta\{\lim\limits_{n->>\infty} \theta_n = \theta_0\} = 1$.

Under the additional conditions of smoothness of the functions $\varphi_i(\theta)$ on the parameter θ the central limit theorem for values $\sqrt{n}(\theta_n - \theta_0)$ takes place too.

The Theorem 2 is the simplified variant of the least modules strong consistency statement given in [11]. In the same work the rate of an estimator convergence to the real value is found.

Theorem 3. Let the conditions of Theorem 2 are satisfied and the following conditions are fulfilled:

1) $g_i = \dfrac{\partial \varphi_i(\theta)}{\partial \theta}$ exists.

2) Random values ξ_i have bounded probability density $h(x)$ and the inequality

$$|g(x) - g(0)| \le H|x|, \ H > 0$$

takes place. Then

$$\sup_{\theta \in K} \left| P_\theta \left\{ 2h(0) \left(\sum_{i=1}^{n} g_i^2(\theta_0) \right)^{1/2} (\theta_n - \theta_0) < x \right\} - \Phi(x) \right| \to 0, \ n \to \infty,$$

where $x \in R^1$, $\Phi(x)$ is Gaussian distribution function.

We formulate now one more statement about the least modules' estimators in situation when regression function depends on random values. This kind of models was studied in [16, 17] and others.

Describe the observation model.

Let $\{(x_i, y_i), i \ge 1\}$ is a stationary in narrow sense metrically transitive random process with discrete parameter specified on the probability space (Ω, θ, P), $x_i \in R^k$, $y_i \in R^m$, $k \ge 1$, $m \ge 1$, $\|\cdot\|_1$ is a norm on R^k defined as

$$\|a\|_1 = \sum_{j=1}^{k} |a_j|, \ a = (a_1, \ldots, a_k)^{\mathrm{T}} \in R^k.$$

For (x_i, y_i) the conditions below are satisfied:

1) $E\|x_i\|_1 < \infty$.

2) For any i with probability one $E(x_i / F_i) = f(\theta, y_i)$, where $f(u, z): I \times R^m \to R^k$ is a function continuous on I when z is fixed and measurable on z when u is fixed; θ is fixed but unknown value from closed set $I \subseteq R^l$, $l \ge 1$.

3) For any $c > 0$ $E\left\{ \max\limits_{\{u \in I, \|u\|_1 \le c\}} \|f(u, y_i)\|_1 \right\} < \infty$.

4) For any sequence $\{u_j, j \ge 1\}$ with $\|u_j\|_1 \to \infty$, $j \to \infty$, $\|f(u_j, y_i)\|_1 \to \infty$ with probability one when $j \to \infty$.

Introduce the notations:

$$\xi_i = (\xi_{1i}, \ldots, \xi_{ki})^{\mathrm{T}} = x_i - f(\theta, y_i), \ f = (f_1, \ldots, f_k)^{\mathrm{T}}.$$

5) For any $j = \overline{1,k}$

$$P\{\xi_{ji} < 0 / F_i\} = \frac{1}{2}.$$

6) For any there exists j: $1 \le j \le k$ such that with probability one

$$f_j(u, y_i) \ne f_j(\theta, y_i).$$

7) For any $\lambda > 0$ $P\{\xi_{ji} \in [-\lambda, 0] / F_i\} \times P\{\xi_{ji} \in [0, \lambda] / F_i\} > 0$.

Let we have observations $\{(x_i, y_i)\}, 1 \le i \le n\}$. It's needed to estimate the unknown parameter θ. Let us consider the LME

$$\theta_n \in \arg\min_{u \in I} F_n(u), \ \|\theta_n\|_1 < \infty,$$
$$F_n(u) = \frac{1}{n} \sum \|x_i - f(u, y_i)\|_1. \tag{5}$$

The following theorem takes place [11, 16, 17]:

Theorem 4. Let the conditions 1)-6) are satisfied. Then there exists at least one vector θ_n, satisfying the condition (5) and

$$P\{\lim_{n\to\infty} \theta_n = \theta_0\} = 1, \quad P\{\lim_{n\to\infty} F_n(\theta_n) = F(\theta)\} = 1,$$

where $F(u) = E\{\|x_i - f(u, y_i)\|_1\}$, $u \in I$.

4. **The minimax method (MMM).** This method is often used in resolving different problems, but these estimators are almost qualitatively unexplored. The difficulty consists in impossibility to imply the ergodic theorem and central limit theorem for the asymptotic properties' investigation. Nevertheless because of the simplicity of the aim function these estimates are convenient to be used in a preliminary stage for the calculation of LSE and LME. The practical expediency of this approach is noted in [10], where the program complex for statistical processing of experimental data using the reasons above is described. Note that the broad use of the LME and minimax estimator (MME) was suppressed by calculating difficulties connected with the minimization of the nonlinear, non-smooth and nonconvex functions. In the case of linear regression function this difficulty can be avoided reducing the criterion minimization problem to the linear programming one. The latest achievements in the sphere of the non-smooth optimization give possibility to solve the problem in the nonlinear regression function case too. In [10] r-algorithm (the generalized gradient descent with the space stretching in two consequent subgradients difference direction ([18]). The estimate is calculated when the restrictions on the parametric set are given:

$$\alpha \in A \subseteq R^m, \ A = \{\alpha: \beta_i \le \alpha_i \le \gamma_i, \ i = \overline{1, m}\}.$$

In a general case the global optimum finding is a difficult problem because of the purpose function nonconvexity. That's why on practice it's expedient to take into account some a priori information about the global minimum localization. The natural way consists in dividing a problem on two parts. The first stage consists in finding of an approximate problem decision in some simplifying prepositions. MMM can be used as the first approximation. It's known that in absence of rough errors the MME well enough approximates the LME, as it is rough but has simpler purpose function. Then on the second stage the minimum point of the initial problem is being found by non-smooth optimization method.

Let us stress one more substantial fact. As a rule, the regression function type cannot be described precisely enough basing on initial or visual data. That's why the broad basis of approximating functions giving the possibility to choose the most suitable is essential. The more complete is the set of approximating functions the more possibilities are accessible for the more precise regression function estimation.

Let's describe one more regression function non-parametric estimation model in some functional space. In our opinion it gives adequate enough reflection of physical processes mentioned above. Suppose that soil or another object observations are made in discrete time moments with independent errors and we would like to determine the trend in any time moment on the whole observation interval. In other words, we have to

restore the regression function in the optimal respectively to some criterion manner, basing on its distorted observations in discrete moments with additive noise. The MLM, LSM, LMM can be taken as criteria. The more precise formulation is the following.

Let K is some real set compact with respect to some metric on [0, 1], $\{\xi_{jn}, 0 \le j \le n\}$ is a set of real independent random values with finite first two moments. The observation model is following:

$$x_{jn} = \alpha_0\left(\frac{j}{n}\right) + \xi_{jn}\ ,\ 0 \le j \le n\ ,\ \alpha_0 \in K\ . \tag{6}$$

We will consider the LSE or LME as estimates:

$$L_n(\alpha_n) = \min_{\alpha \in K} \sum_{j=0}^{n} \left[x_{jn} - \alpha\left(\frac{j}{n}\right)\right]^2, \tag{7}$$

$$\breve{L}_n(\breve{\alpha}_n) = \min_{\alpha \in K} \sum_{j=0}^{n} \left| x_{jn} - \alpha\left(\frac{j}{n}\right)\right|. \tag{7′}$$

Note, that minimization problems by the criteria (7) and (7′) are more difficult than studied above because the optimum is being found on some class of functions. The estimation problems of this kind are called nonparametric. Their decision is connected with great difficulties of both compute and qualitative connected with statistical properties investigation character. But for these models numerical methods of optimum finding are developed as well. Among the works in this sphere, we mark the interesting approach developed in [19].

For the set of continuous functions in the space with uniform metric the LSE is studied in [12], and the LME in [11]. We'd like to describe more detailed the model (6) with the criterion (7′) supposing that the unknown function belongs to some set from Hilbert space. The proofs of the basic statements will give an idea of research methods for such problems and the mathematical technique used in them. The statement about the strong consistency and the functional limit theorem for the standardized estimators will be proved. In particular, the standardized estimates weak convergence to the standard Wiener process statement will be proved. Therefore, any linear functional of the estimator converges to the linear functional of Wiener process. In particular, on practice the knowledge of the distributions below

$$\sup_{0 \le t \le T} \alpha_n(t)\ ,\quad \sup_{0 \le t \le T} \breve{\alpha}_n(t)$$

is important enough.

As the weak convergence statements are proved to calculate these functionals distributions it is enough to know the functional $\sup\limits_{0 \le t \le T} W(t)$, where $W(t)$ is the standard Wiener process. This distribution is known and given, for example, in [19]. It looks as follows:

$$P\left\{ a \le \sup_{0 \le t \le T} W(t) \le b \right\} = \tag{8}$$

$$= 1 - \frac{2}{\sqrt{2\pi T}} \sum_{n=0}^{\infty} \left[\int_{(2n+1)(b-a)-a}^{(2n+1)(b-a)+a} \exp\left\{ -\frac{y^2}{2T} \right\} dy + \int_{(2n+1)(b-a)-b}^{(2n+1)(b-a)+b} \exp\left\{ -\frac{y^2}{2T} \right\} dy \right].$$

Thus, we have possibility to calculate probabilities of the location inside a band of some characteristic. It may be phosphate content in soil, heavy metal (for instance, cadmium) or different salts quantity on the soil surface or in the depth. Moreover, knowing the distribution of functional we can find the first jump over the threshold which can characterize the critical level of some substance inclusion.

Now we formulate the problem more strictly.

1. Let $K = \{x(t), t \in [0,1]\}$ be a set of real functions, compact in the sense of convergence in L_2, satisfying the conditions:

a) $\|x(t)\| \le 1$, where $\|x(t)\|^2 = \int_0^1 x^2(t)dt$;

b) $x(t)$ can have only a finite number of discontinuities of the first kind and at intervals of continuity satisfies the uniform Lipschitz condition: $|x(t_1) - x(t_2)| \le c|t_1 - t_2|$, the constant c being independent of the function x and the points t_1 and t_2, at points of discontinuity $x(t) = x(t-0)$.

2. For each are independent random variables satisfying the conditions:

a) $E\xi_{0n} = 0$;

b) $E[\xi_{0n}]^2 = \sigma^2 < \infty$.

We assume that for a fixed function $\alpha_0 \in K$ the random variables

$$x_{kn} = \alpha_0\left(\frac{k}{n}\right) + \xi_{kn},\ 0 \le k \le n,$$

are observed. It is necessary to estimate the unknown function $\alpha_0 \in K$ from the observations x_{kn}. Let us choose the least square estimate for α_n as the estimate for α_0. α_n is an element from K defined by the relation

$$\sum_{j=0}^{n}\left[x_{jn} - \alpha_n\left(\frac{j}{n}\right)\right]^2 = \min_{\alpha \in K}\sum_{j=0}^{n}\left[x_{jn} - \alpha\left(\frac{j}{n}\right)\right]^2.$$

The minimum is reached because of the compactness of the set K. It can be shown that $\alpha_n(t)$ can be chosen to be a separable measurable stochastic process. In what follows we will need the following assertions, proved in [12].

Theorem 5. Let (Ω, F, P) be a probability space, let $\{F_n, n \ge 1\}$ be a sequence of σ-algebras such that $F_n \subset F_{n+1}$, $n \ge 1$, and let K be a compact subset of some Banach space with norm $\|\cdot\|$. We assume that

$$\{Q_n(S) = Q_n(S,\omega), (S,\omega) \in K \times \Omega,\ n \ge 1\}$$

is a sequence of real functions satisfying the following conditions:

1) for fixed n and each $S \in K$ the function $Q_n(S,\omega)$: $\Omega \to R$ is F_n-measurable;

2) for fixed n and ω the function $Q_n(S,\omega)$: $K \to R$ is continuous on K; for each $n \ge 1$ and $\omega \in \Omega$ the element $S_n = S_n(\omega) \in K$ is given by the relation:

$$Q_n(S_n) = \min_{S \in K} Q_n(S);$$

3) for some fixed element $S_0 \in K$ and for each $S \in K$ the relation

$$P\{\lim_{n\to\infty} Q_n(S,\omega) = \Phi(S, S_0)\} = 1$$

is true for some real functions $\Phi(S, S_0)$: $K \to R$, continuous on K and such that

$$\Phi(S, S_0) > \Phi(S_0, S_0),\ S \ne S_0;$$

4) for any $\delta > 0$ there exists $\gamma_0 > 0$ and a function $c(\gamma)$: $R \to R$, $c(\gamma) \to 0$, $\gamma \to 0$, such that for each $S \in K$ and each $0 < \gamma < \gamma_0$ the relation

$$P\left\{\lim_{n\to\infty} \sup_{\{S:\, \|S - S'\| < \gamma,\, \|S - S_0\| > \delta\}} |Q_n(S) - Q_n(S')| < c(\gamma)\right\} = 1$$

holds. Then

$$P\{\lim_{n\to\infty} \|S_n - S_0\| = 0\} = 1.$$

The following assertion follows from Theorem 5 [12].

Theorem 6. Assume that the conditions of Theorem 4 are fulfilled and that

$$E|\xi_{jn}|^4 = \gamma < \infty.$$

Then

$$P\{\lim_{n\to\infty} \|\alpha_n(t) - \alpha_0(t)\| = 0\} = 1.$$

Now we proceed to the study of the distribution of functionals of the estimators α_n. We assume additionally that

7) α_0 is an internal point of K in the following sense:

a) $\|\alpha_0(t)\| < 1$;

b) for $\alpha_0(t)$ the condition 1b is fulfilled for a constant $\widetilde{c} < c$.

The following auxiliary assertions take place [12].

Lemma 1. Let the condition of Theorem 6 be fulfilled and let α_0 be an internal point of K. Then for some function $\varphi(t) \in K$ such that

$$\|\varphi\| = 1, \quad \int_0^1 \varphi^4(t)dt < \infty$$

the distribution of the random functional

$$\sqrt{n}\int_0^1 \varphi(t)[\alpha_n(t) - \alpha_0(t)]\,dt$$

converges weakly to the distribution of normal random variable with mean 0 and variance σ^2 as $n \to \infty$.

Lemma 2. Let the conditions of Lemma 1 be fulfilled and let

$$E|\xi_{jn}|^6 = \nu < \infty.$$

Then for each $0 \le t_1 \le t_2 \le 1$ the following inequality holds:

$$E\left\{\sqrt{n}\int_{t_1}^{t_2}[\alpha_n(t) - \alpha_0(t)]\,dt\right\}^6 \le c(t_2 - t_1)^2.$$

Lemma 1 and Lemma 2 imply the following theorem.

Theorem 7. Assume that the conditions of lemma 2 are fulfilled. Then the sequence of random processes $\eta_n(t)$ converges weekly to the standard Wiener process as $n \to \infty$.

Problem of critical threshold attainment. Let us now turn to certain problems, connected with the behavior of some ecological systems, among them there are described above. As indicated in [20], the real ecological systems are exposed to various random influences. If the time of these random influences or perturbations is considerably less than a system work time, it is possible to apply sufficiently developed instrument of Markov processes to the whole system analysis (the dynamics of its development). Under these conditions the natural model for random perturbations is the white noise. For these perturbations the mathematical instrument is well developed, it allows to describe systems' dynamic as the stochastic differential equation:

$$\frac{dx(t)}{dt} = a(t, x(t)) + \sigma(t, x(t))\,\xi(t), \tag{9}$$

where $a(t,x)$ is a drift coefficient, according to the determinate part of system development, and $\xi(t)$ is a white noise. The consideration of random influence, combined with white noise (which is an idealized random process with short time cor-

relation), allows to use well-developed apparatus of stochastic equation theory and to obtain qualitative and quantitative results about the modeling of the process (9) behavior. The equation of type (9) determines a diffusion process with written densities of transient probabilities satisfying the Kolmogorov equations (strait and reverse). For these processes it's possible to estimate some characteristics of ecological system behavior which is noted above; for instance, the problems, connected with a critical threshold system attainment. Let us stop on this problem.

Let (r_1, r_2) is the interval of permissible changes $x(t)$ and T is the moment of $x(t)$ first exit from the interval (r_1, r_2). The essential assumption is that probability of $x(t)$ first exit from the interval (r_1, r_2) will be strongly positive, that is

$$P\{\lim_{t\to T} x(t) = r_1\} > 0.$$

If $T < \infty$ then we say that the bound may be attained for a finite time. This bound is called attractive. It is evident that among attractive bounds there are two interesting groups: i) the absorption bounds and ii) the permitting bounds. Apropos of bounds type there are some results in stochastic equation theory giving exhaustive answer under some conditions, for instance, in [21, 22]. We won't do the detailed review of the results connected with moments of some level (threshold) intersection. Let us stop at some of them.

We will consider, for simplicity, the uniform in time stochastic equation, supposing that in (16)

$$a(t,x) = a(x),\ \sigma(t,x) = \sigma(x).$$

The following assertion takes place.

Theorem 8. Let $\sigma(x) > 0$ when $x \in [a,b]$. Then $\tau_x[a,b] = \inf\{t: x(t) \notin [a,b]\}$, $x(0) = x$ is finite with probability 1 when $x \in (a,b)$ and $E\tau_x[a,b] = \nu(x)$, where $\nu(x)$ is the decision of differential equation

$$\frac{1}{2}\sigma^2(x)\nu''(x) + a(x)\nu'(x) = -1,$$

satisfying the condition:

$$\nu(a) = \nu(b) = 0.$$

The proof of Theorem 8 is typical in the diffusion processes theory, it uses Markov moments and stochastic differential calculation apparatus [21].

Using Ito formula and Theorem 8, we can get the equation for obtaining the second moment of $\tau_x[a,b]$. Let us formulate this assertion.

Theorem 9. Under the conditions of Theorem 8, $E\{\tau_x[a,b]\}^2 = \nu_1(x)$ satisfies the equation

$$\frac{1}{2}\sigma^2(x)\nu_1''(x) + a(x)\nu_1'(x) = -2\nu(x),$$

where $\nu_1(a) = \nu_1(b) = 0$ and the function $\nu(x)$ is defined as in previous theorem.

As demonstrated above, we have found two first moments of the diffusion process first exit time from some interval. However, we can obtain the fuller description of random value $\tau_x[a,b]$ behavior. It is obvious, that when $x \in (a,b)$ $\tau_x[a,b]$ equals to either a or b. Let

$$\tau_x^1[a,b] = \tau_x[a,b],\ \text{if}\ \tau_x[a,b] = a,$$

and

$$\tau_x^2[a,b] = \tau_x[a,b],\ \text{if}\ \tau_x[a,b] = b.$$

Let

$$\varphi_{i\lambda}(x) = Ee^{-\lambda\tau_x^i[a,b]}$$

is Laplace transformation of $\tau_x^i[a,b]$, $i = 1,2$. Then, as shown, for instance, in [22],

$\varphi_{1\lambda}(x)$ satisfies the following equation:

$$L\varphi_{1\lambda}(x) = \lambda\varphi_{1\lambda}(x),\ \ \varphi_{1\lambda}(x)(a+0) = 1,\ \varphi_{1\lambda}(b-0) = 0,$$

where L satisfies the equation

$$L = a(x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2}.$$

Similarly, $\varphi_{2\lambda}(x)$ satisfies the following equation:

$$L\varphi_{2\lambda}(x) = \lambda\varphi_{2\lambda}(x),\ \ \varphi_{2\lambda}(x)(a+0) = 0,\ \varphi_{1\lambda}(b-0) = 1.$$

If we know Laplace transformation, we can write other characteristics of value $\tau_x[a,b]$.

One of the most important questions is the question about calculation of the random process $x(t)$ first exit probability from some interval, and as a first priority we are interested in the searching of hit probability of random process $x(t)$ in the point a earlier then in the point b, and on the contrary. It is evident, that when $x \notin (a,b)$, the random process $x(t)$ falls in a nearest point earlier with probability 1. Therefore, we will suppose that $x \in (a,b)$. Let $P_a(x,b)$ is the hit probability of random process $x(t)$ in the point a earlier then in the point b, when it is in the point x the first time.

The following assertion given in [21, 22] is true.

Theorem 10. Let $x \in (a,b)$ and $\sigma(x) > 0$. Then $P_a(x,b)$ satisfies the equation

$$P_a(x,b) = \frac{u(x) - u(b)}{u(a) - u(b)},$$

where $u(x)$ is the decision of the equation

$$\frac{1}{2}\sigma^2(x)u''(x) + a(x)u'(x) = 0, \tag{10}$$

which is not equal to 0 identically on the interval $[a,b]$.

Proof. Let

$$\psi(x) = \frac{u(x) - u(b)}{u(a) - u(b)}.$$

It is evident, that $\psi(x)$ satisfies (10). Using Ito formula, it is easy to show that

$$E x\tau_x[a,b] = \psi(x).$$

According to assumption, $x\tau_x[a,b] = a$ with probability $P_a(x,b)$ and $x\tau_x[a,b] = b$ with probability $P_b(x,a)$. Hence,

$$\psi(x) = E x\tau_x[a,b] = \psi(a)P_a(x,b) + \psi(b)P_b(x,a) = P_a(x,b).$$

Similarly, we can get the formula for $P_b(x,a)$.

The theorem is proved.

Corollary 1. Let $\Phi(z)$ is Laplace function. Then

$$P_a(x,b) = \frac{\int_x^b \Phi(z)dz}{\int_a^b \Phi(z)dz},\ \ P_a(x,\ a) = \frac{\int_a^x \Phi(z)dz}{\int_a^b \Phi(z)dz}. \tag{11}$$

For proof (11) it is enough to mention that

$$u(x) = \int_a^x \Phi(z)dz$$

satisfies the conditions of Theorem 10.

Corollary 2. If $a(x) = 0$, then $P_a(x,b)$ and $P_b(x,a)$ have simpler forms:

$$P_a(x,\ b) = \frac{b-x}{b-a},\quad P_b(x,\ a) = \frac{x-a}{b-a}.$$

Notice, that the probability of a random process $x(t)$ stay in the band $[a, b]$ is given by the formula (8) in the case of Wiener process. By virtue of Girsanov theorem [23], under the natural conditions by means of a measure replacement and a transition to other probability space the space $x(t)$ also will be Wiener transformed measure relatively. And then we can apply (8) to obtained process.

So, using the model (16), we can estimate the moment of approach of the critical level of some element saturation (for instance, a limit permissible level of phosphorus in soil) and take previous planed precautions.

Note, that in the above part we touched the problem of level exceeding in the model, described by equation (16). The sufficiently wide research concerning the finding of the moment of an attainment some level and a stay in the bond, is made for stationary, Gaussian and Poisson processes and, without doubt, these results will be used under an adequate model existing.

The fuller exposition of these questions is given in [24], but we have no possibility to top on it now. We only give some facts useful for the problem given above decision.

We will consider stationary Gaussian process. The main idea of the investigation is to compare this process with simply normal process $x^*(t)$, maximum of which is easily calculated. Let us consider as an example the process

$$x^*(t) = \eta \cos \omega t + \varsigma \sin \omega t,$$

where $\omega > 0$, η and ς are standard normal values.

It is evident that random process $x^*(t)$ is Gaussian. It is easy transformed to the following form:

$$x^*(t) = A\cos(\omega t - \varphi).$$

For such processes the following assertion takes place.

Lemma 3 [24]. Let $M^*(t) = \sup\limits_{t \in [0,T]} x^*(t)$. Then

$$P\{M^*(t) \le u\} = \Phi(u) - \frac{\omega T}{2\pi} \exp\left\{-\frac{u^2}{2}\right\} \tag{12}$$

for $0 < T < \dfrac{\pi}{\omega}$, $u > 0$.

Proof. Let $N = N_u(t)$ is a number of process' $x^*(t)$ exits out of level u in the interval $[0, T]$. Then

$$E(N) = \frac{\omega T}{2\pi} \exp\left\{-\frac{u^2}{2}\right\}$$

and

$$P\{M^*(t) > u\} = P\{x^*(0) > u\} + P\{x^*(0) \le u,\ N \ge 1\} = 1 - \Phi(u) + E(N) =$$

$$= 1 - \Phi(u) + \frac{\omega T}{2\pi} \exp\left\{-\frac{u^2}{2}\right\}, \tag{13}$$

where

$$\Phi(u) = \frac{1}{2\pi} \int_0^u e^{-\frac{t^2}{2}}\, dt.$$

The equality (13) is equivalent to (12), that is why the lemma is proved.

The comparison of an arbitrary Gaussian process with the process $x^*(t)$ is based on the following assertion [25].

Lemma 4. Let $x_1(t)$ and $x_2(t)$ are Gaussian continuous processes with $Ex_1(t) = Ex_2(t) = 0$, $Ex_1^2(t) = Ex_2^2 = 1$, and let $r_1(t, s)$, $r_2(t, s)$ are their covariation functions. Let for some $\delta > 0$ $r_1(t, s) \ge r_2(t, s)$ for each $0 \le t, s \le \delta$. Then these

processes' maximums $M_1(t)$ and $M_2(t)$ satisfy the relation:

$$P\{M_1(t) \le u\} \ge P\{M_2(t) \le u\}$$

for any $0 \le T \le \delta$.

The proof invokes a piecewise-linear approximation of processes $x_i(t)$ by processes $x_n^i(t)$, taking values $x_i(t)$ on the set $\{jq_n,\ j = 0,1,\ldots\}$ and being linear inside the intervals $(jq_n,\ jq_{n+1})$, $q_n \to 0$, $n \to 0$. Let us denote $M_i^n(t) = \max\{x_i(jq_n), 0 \le jq_n \le T\}$. It is obvious that $M_i^n(t) \to M_i(t)$ as $q_n \to \infty$. Put $q_n = 2^{-n}$. Then $\{M_i^n \le u\} \to \{M_i \le u\}$, that is why $P\{M_i^n(T) \le u\} \to P\{M_i(T) \le u\}$ as $n \to \infty$. It is easy to see that $P\{M_2^n(T) \le u\} \le P\{M_1^n(T) \le u\}$ [25]. These facts prove the lemma.

Using Lemmas 1 and 2 it is easy to obtain the following important assertion.

Theorem 11 [25]. Let $u, T \to \infty$ in such a way that $\frac{T}{2\pi}\lambda_2^{1/2}\exp\left\{-\frac{u^2}{2}\right\} \to \tau \ge 0$ and correlation function $r(t)$ of process $x(t)$ satisfies the following condition: $r(t)\ln t \to 0$ as $t \to \infty$. Then

$$P\{E(T) \le u\} \to e^{-\tau},$$

as $T \to \infty$, and

$$P\{a_t(E(T) - b_T) \le x\} \to \exp\{e^{-x}\},$$

as $T \to \infty$, where

$$a_T = (2\ln T)^{1/2}, \quad b_T = (2\ln T)^{1/2} + \frac{\ln\frac{\lambda_2^{1/2}}{2\pi}}{(2\ln T)^{1/2}}.$$

Above we investigated Gaussian process extremums properties without obtaining the information about the location of this extremum.

Let $L(t)$ is a point in which process $x(t)$ reaches its maximum for the first time in the segment $[0, T]$. The supposition that for any stationary process $L(t)$ has uniform distribution may arise. But generally speaking, it's incorrect. This will take place in the case when $x(t) = A(t - \varphi)$, where φ is a random variable uniformly distributed on $[0,2\pi]$. For stationary Gaussian process $L(t)$ is always symmetric on $[0, T]$ and the possible leaps at points 0 and T have equal values. One of the methods for the point like 0 and T elimination consists in their moving away to the infinity. The question arises whether $L = L(t)$ is asymptotically uniform when $T \to \infty$. For Gaussian stationary processes this fact is the consequence of the asymptotic independence of maximums on the intersecting intervals. More precisely, the following statement takes place.

Theorem 12. Let $x(t)$ is a stationary Gaussian process, $Ex(t) = 0$, $Ex^2(t) = 1$, $\lambda_2 = -r''(0) < \infty$, $r(t)\ln t \to 0$ as $t \to \infty$. Then

$$P\{L(T) \le lT\} \to l$$

as $T \to \infty$, $0 \le l \le 1$.

REFERENCES

1. Stigliani W.M. Changes in valued "Capacities" of soils and sediments as indicators of nonlinear and time-delayed environmental effects. *Environmental Monitoring and Assessment*. 1988. Vol. 10, Iss. 3. P. 245–307.
2. Kashyap R.L., Rao A.R. Dynamic Stochastic Models from Empirical Data. New York; San Francisco; London: Academic Press, 1976. 334 p.
3. Kalman R.E., Bucy R.S. New result in linear filtering and prediction theory. *J. Basic Eng*. 1961. Vol. 83 (1). P. 95–108.

4. Liptser R.Sh., Shiryaev A.N. The Statistics of Random Processes. Moscow: Nauka, 1974. 696 p.
5. Laning J.H., Battin R.H. Random Processes in Automatic Control. New York; Toronto; London: McGraw-Hill Book Company Inc., 1956. 450 p.
6. Darlington S. Nonstationary smoothing and prediction using network theory concepts. *IRE Transactions on Circuit Theory*. 1959. Vol. 6, Iss. 5. P. 1–13.
7. Dorogovtsev A.Ya. Remarks on the processes generated by some differential equations. *Visnyk KDU*. 1962. Vol. 1, N 3. P. 124–146. (In Ukrainian).
8. Knopov P.S. To question about a forecast to random fields. *Teoriia Optimalnykh Reshenii*. 1963. N3. P. 74–83. (In Russian).
9. Gikhman I.I., Skorokhod A.V. The Introduction to the Random Processes Theory. Moscow: Nauka, 1965. 655 p. (In Russian).
10. Bardadym T.A., Borshchevskyi A.V., Vovk L.B., Ivanov A.V. The complex of programs for statistical processing of experimental data. Kiev: Institute of Cybernetics AN USSR, 1991. 20 p. (Prepr. AN USSR, Institute of Cybernetics; 91–41). (In Russian).
11. Leonenko N.N., Ivanov A.V. Statistical Analysis of Random Fields. Kiev: Vyshcha Shkola, 1989. 216 p. (In Russian).
12. Dorogovtsev A.Ya. The Estimation Theory of Random Processes Parameters. Kiev: Vyshcha Shkola, 1982. 192 p. (In Russian).
13. Ermoliev Yu.M. Stochastic Programming Methods. Moscow: Nauka, 1976. 340 p. (In Russian).
14. Wasan M.T. Stochastic Approximation. Cambridge: Cambridge University Press, 1969. 212 p.
15. Kaniovskii Yu.M., Knopov P.S., Necrylova Z.V. Limit Theorem for Stochastic Programming Method. Kiev: Naukova dumka, 1980. 206 p. (In Russian).
16. Dupacová J., Wets R.J.-B. Asymptotic behavior of statistical estimators and optimal solutions for stochastic optimization problem. IIASA WP 86–41. 1986. Laxenburg, Austria.
17. Kasitskaya E.I., Knopov P.S. About one approach to the nonlinear estimating problem. *Proc. 6-th USSR–Japan Symposium Probability Theory and Mathematical Statistics* (5 – 10 August 1991, Kiev, USSR). Kiev, 1991. P. 151–157.
18. Mudrov B.I., Kushko V.L. Methods of Measuring Processing. Moscow: Radio i sviaz, 1983. 304 p.
19. Gaivoronskyi A.A., Golodnikov A.N., Le Din Fun. Decision method of limit extreme problems for functional's minimizing. *Cybernetics*. 1980. N 1. P. 88–92. (In Russian).
20. Shiryaiev Yu.M. Nonlinear Waives, Dissipative Structures and Catastrophes in Ecology. Moscow: Nauka, 1987. 368 p. (In Russian).
21. Gihman I.I., Skorokhod A.V. Stochastic Differential Equations. Kiev: Naukova dumka, 1968. 354 p. (In Russian).
22. Ito K. Probability Processes. Vol. 2. Moscow: Inostrannaya literatura, 1963. 136 p. (In Russian).
23. Girsanov I.V. On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability and Its Applications*. 1960. Vol. 5, N 3. P. 285–301. (In Russian).
24. Leadbetter M.R., Lingren G., Rootzen H. Extremes and Related Properties of Random Sequences and Processes. New York; Heidelberg; Berlin: Springer-Verlag, 1986. 336 p.
25. Slepian D. The one-sided barrier problem for Gaussian noise. *Bell System Technica1 Journal*. 1962. Vol. 41, Iss. 2. P. 463–501.

І.В. Дорошенко, О.П. Кнопов, Л.Б. Вовк

МАТЕМАТИЧНІ МОДЕЛІ ЕКСТРЕМАЛЬНИХ РЕЖИМІВ В ЕКОЛОГІЧНИХ СИСТЕМАХ

Анотація. У статті розглянуто основні етапи створення комп'ютерної системи підтримки прийняття рішень у галузі екологічної безпеки. Проаналізовано основні фактори забруднення довкілля після екологічних і техногенних катастроф, типові збої під час експлуатації небезпечних об'єктів. Запропоновано методи кількісного оцінювання функції ризику, створено еколого-математичні моделі, що дають змогу оцінити поточний стан довкілля й дати прогноз на майбутнє та зрозуміти суть процесу, що розглядається.

Ключові слова: модель, часовий ряд, тренд, метод максимальної правдоподібності (ММП), метод найменших квадратів (МНК), метод найменших модулів (МНМ), метод мінімаксу (МММ).

Надійшла до редакції 02.05.2022