DOI 10.1007/s10559-022-00510-w
Cybernetics and Systems Analysis, Vol. 58, No. 5, September, 2022

MATHEMATICAL MODELS OF EXTREME MODES IN ECOLOGICAL SYSTEMS

I. Doroshenko,[1] O. Knopov,[2] and L. Vovk[3] UDC 517.9

Abstract. *The authors consider the basic stages of creating a computer system for decision 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 assessment 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 comprehension of the process under consideration.*

Keywords: *model, time series, trend, maximum likelihood method (MLM), least squares method (LSM), least absolute deviations method (LADM), minimax method (MMM).*

Under the condition of scientific and technical progress, the rapid development of the level of anthropogenic and technical influence on the environment substantially increased. Thus, many biosphere parameters reached critical values and became dangerous to human life and mankind's existence as a whole.

The problem arising in this connection makes it necessary to develop a system of decision support in ecological security. This is one of the most difficult multi-factorial problems. Its solution is only possible with the use of system analysis methods.

Let us consider the basic stages of creating a computer decision-support system. The first stage is the analysis of the key factors of pollution caused by ecological and technological catastrophes, typical failures of projects, exploitation of dangerous objects, etc. The second stage involves a quantitative assessment of the risk function and creation of an ecological and mathematical model that allows estimating the current state of nature and forecasting the future. The main goal of the model is to achieve comprehension of the process under consideration.

Let us give an example taken from [1] and deal with the analysis of soils' main properties. Following [1], it will suffice to consider (i) acids neutralization; (ii) adsorption of phosphorus and toxic materials; and (iii) oxygen donation.

Let us review each of them.

(i) **Acids Neutralization.** This capacity is substantial for maintaining the stability of freshwater habitats necessary for protecting fish and other aquatic life from acidity fluctuations. Soils' ability to neutralize acids substantially depends on their types.

Analyzing the situation over the last 100–150 years in the industrial regions of North America and Europe, we may conclude that soils can buffer large perturbations due to a great amount of sulfuric and nitric acidic rains caused by the burning of organic fuels. However, the effect of such impacts is not uniform for different regions because of the large variability in soil type. The paper [1] contains information about the acidification of lakes. The rate of lake response to acid inputs depends on two fundamental qualities of watershed soils: the ability to retain sulfate and the ability to supply base cations. It is necessary to know the history of acid emissions and time trends in pH and in the exhaustion of the watershed soil buffering capacity. Unfortunately, there is little information about it in the literature, and the data available

[1]Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine, *i.doroshenko@chnu.edu.ua*. [2]V. M. Glushkov Institute of Cybernetics, National Academy of Sciences of Ukraine, Kyiv, Ukraine, *a.knopov@gmail.com*, [3]National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine, *lilyvovk@gmail.com*.
Translated from Kibernetyka ta Systemnyi Analiz, No. 5, September–October, 2022, pp. 105–121. Original article submitted May 2, 2022.

 1060-0396/22/5805-0764

are pure. Moreover, experiments in this field are expensive and often impossible. We have to use the information available. That is why it is necessary to develop methods allowing to make inferences based on incomplete information such as statistical methods of small sample processing.

(ii) **Adsorption of Phosphorus and Toxic Materials.** Chemicals being added to soils are known to fall within four groups: (a) those getting into the ground and then in surface waters; (b) taken up by plants; (c) volatilized into the atmosphere, and (d) stored in the soil. In cases (a)-(c), when chemicals are input into 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 the nonlinear and time-delayed effects arising. A potential danger of chemicals is their ability to accumulate in soils thus the threshold for the soil's capacity to hold the chemical can be reached.

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

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

(a) What level of soil acidity is harmful to human health? When will it be accessed (or whether it will be accessed) for some regions?

(b) What is the law according to which the soil properties change?

(c) How to estimate the threshold moment for the soil's capacity to withstand the chemical effect?

MODELS

1. This kind of problems arises in the analysis of different natural and human activity-related phenomena. The models we design have to reflect rather efficiently the phenomenon under consideration and answer the questions we put in each specific case. With the possibility to describe the model by analytic expressions, we can considerably advance in it. On the other hand, we can tend to design the empirical model based on the summing up of the available observation set. We can define the characteristics of this process with the help of modeling or with estimation on the observation background. Note that when we investigate such a complicated process, we usually cannot get a satisfactory determinate model. Therefore, it is natural (see [2] for detals) to consider the investigated process as a 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 a function of an unknown kind containing the unknown parameters: $g(t) = g(t, \alpha)$ or $g(t)$ is a function of an 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 will be called trend extracting.

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

$$f(y(t)) = g(\xi(t-1)) + W(t), \tag{1}$$

where $\xi(t-1)$ is the observation vector to time moment $t-1$, $W(t)$ is a subsequence of independent random magnitudes, and f and g are some functions. Relation (1) is a stochastic difference equation. The theory of these equations was analyzed by Kalman [3, 4] and others. The main task is to develop an algorithm to forecast the process $y(t)$ based on some observation vector $\xi(t)$ at time $t-1$. This problem is investigated in full in the case where 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, is not 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 defined by the 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 analyzed in detail in [5, 6, etc.].

Likewise, we can consider a continuous analog of model (2), where operator D is given by

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

The equation of kind (2) includes a rather wide class of models, including the autoregression model and autoregression model with moving average. Such models describe rather well the processes related to risk assessment for ecology end economic systems. For example, in forecasting phosphate reserves in the soil, when data at previous time moments are known, the difference equation (2) is a quite natural model of the real process. Some difficulties may arise because the trend is generally time-dependent, so coefficients of $A(D)$ may be time-dependent too. The forecasting methods have been developed in such case as well [7]. But unlike the previous case, the Wiener–Kolmogorov method is used here, and the solution of the forecast problem reduces to some difference or differential equations.

Note that for some substances propagating in the soil it is sometimes convenient to treat the forecasted parameter t as a vector one. Then the model under study can be presented as (2), but the operator D has the form

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

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

Such systems are analyzed in [8].

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

Let us make some important notes on the models that describe soil pollution, accumulation of chemical substances and acids in soils, etc. Since our purpose is to obtain estimates related to the presence of chemicals, acids, and heavy metals in soil, it is natural to solve our problems using the mathematical methodology of random fields. We will analyze the so-called geochemical field, where the content of some chemical substance will be treated as a coordinate and time function. Thus, a random geochemical field will be a 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, $t \in T \subseteq R$, $x \in R^m$, $m \geq 1$. Chemical concentration is considered in a domain if $m = 2$ and in a space if $m = 3$. In experimental studies, a mean value of a point spatial variable within some domain or volume at time t is only available. For instance, we can describe the mean value of the function $\xi(\omega, t)$ within the domain S at the point $x \in S$ at time t as follows:

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

Similarly, we can determine the mean value of the function $\xi(\omega, t)$ within the volume V with the center of gravity at the point $x \in S$ at time t:

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

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

(i) **Uniformity of the Random Field $\xi(t, x)$ on x in a Restricted or Broad Sense** [9]. This property is manifested in the fact that at any point of the field, a random function has the same average characteristics. In other words, we can repeat the experiments over and over again in order to obtain as much as desired realizations of the

random function and thus to make the corresponding statistical inferences. This hypothesis will be inapplicable if the regular change of the values of $\xi(t, x)$ is observed, i.e., when the observation point moves away from the central part of the field. Sometimes, it is convenient to consider the field with stationary increments rather than a uniform field, i.e., joint distribution of the random values

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

under fixed t stays constant when the variables x_k are simultaneously moved with respect to h.

(ii) **Continuity.** Usually, soil properties change slowly and continuously. That is why this condition looks quite natural.

Apparently, properties (i) and (ii) are basic for the natural transition from the observations of function $\xi(t, x)$ to its average characteristics $y(t, x)$. Thus, if we fix some domain 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 the problem is being solved for some specific area or region.

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

ANALYSIS

Before solving the problems of finding the estimation, forecast, and distribution for some functionals, we will make a preliminary analysis of the time series. In our opinion, it consists of the following steps:

1a. Observation independence checking.

2a. Trend presence hypothesis checking.

3a. Observed value distribution hypothesis checking.

Then we perform a more detailed time series analysis. It consists of the following steps:

lb. Finding the class of models subject to investigation.

2b. Model parameter statistical estimation.

3b. Forecast of the time series values at the next time moments.

There is no need to dwell on the preliminary statistical analysis of time series. These problems are elucidated in many textbooks and monographs on mathematical statistics. We performed the statistical analysis for the data from [1]. The statistical information was processed by a software complex [10]. The main conclusion is the fact that, as a rule, each time series from [1] has a trend. The observation errors can be treated as independent, but normality does not always take place.

Problem of Trend Revealing. Let us consider the problem of trend revealing. First, note some specific properties of the above-mentioned time series.

1. As pointed out in [1, etc.], there is a lack or absence of information for many regions.

2. The noise distribution is not always known.

Taking into account these facts and performing a preliminary analysis of the time series, we can use different methods for trend revealing. The most known are: the maximum likelihood method (MLM), the least squares method (LSM), the least absolute deviations method (LADM), and the minimax method (MMM). We will describe each of them later and only mention that LADM is the most preferable in our case. The least absolute deviation estimators (LADE) are referred to as robust estimators, i.e., the estimators that have advanced stability with respect to the changes in the error distribution law, the presence of a rough error, etc. Moreover, the LADE are considerably more efficient in the case of a small sample. A series of studies of the LADM qualitative behavior has been done recently. First, the consistency of the estimators should be emphasized. Rather simple and easy-to-test consistency conditions are provided in [11, 12]. Note that rather general models are analyzed, where the parameter is generally contained in the regression function nonlinearly, and the noise may be dependent at different instants of time. Moreover, the statement of asymptotic normality can be proved under quite natural conditions.

One more circumstance should be mentioned. In considering small samples, the asymptotic properties such as consistency, asymptotic normality, etc. cannot be the criteria of estimators' quality. In this case, the estimators' stability (robustness) with respect to the observation distribution laws and the error variance are primary. In this respect, in the

case of unknown distribution we should prefer LADM and MMM. In the case of large samples, we certainly should try to get estimators' consistency and asymptotic normality.

1. **The Maximum Likelihood Method (MLM)** is preferable because of its properties in the cases where the noise distribution is known. There are classical results widely known for specialists in statistics: under some natural conditions, MLM estimators are consistent, asymptotically normal, and effective. However, their calculations imply great difficulties, that is why this method is not often used in practice. We will not describe these estimators in detail. Their properties are detailed in [2, 4].

2. **The Least Squares Method (LSM).**

The problem formulation is as follows. Consider the problem of estimating the integer nonlinear parameters contained in the regression function. The observation model is as follows:

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

where $\alpha \in A$ is a vector parameter belonging to some feasible set A, $g(t_j, \alpha)$ is a regression function, and ξ_j is random errors of the observations.

The observations may be independent or dependent but should be a stationary (in a narrow sense) sequence of random values with the unknown distribution function that is independent of α.

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

The random vector α_n that minimizes the sum of squared residuals

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

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

Due to the simplicity of the criterion and the fact that it is well adapted to the calculation procedures, the LSM is the most popular and well-studied method. It coincides with the MLM in the case of the Gaussian observation errors. It is most convenient for the calculations when the unknown parameter is contained in the regression function linearly. Here, we will not provide an example of the calculation algorithms and consistency, asymptotic normality, and asymptotic efficiency statements, we will only refer to the studies [2, 11]. Turning to the time series that describe the distribution of elements in the soil, we can make some conclusions regarding the trend. For instance, from Fig. 6 [1], we can see that the trend can be approximated by a straight line, for Fig. 11 [1] the approximating curve is a parabola, etc. But a linear relation cannot always be obtained. That is why it becomes important to investigate the LSE properties when a nonlinear dependence on the parameters takes place. Yet a deep analysis of estimators' properties is made in this case as well. However, finding the estimators involves significant calculation difficulties.

Note that even in the linear case, the calculation schemes have such a grave disadvantage as the necessity to save in the memory and to process large amounts of numerical information. The recursive computation schemes for the least square estimators described in [13, 14] do not have these disadvantages. Their essence is as follows. Let the following random sequence be observed:

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

where $\varphi(i)$ is a known function, α is an unknown parameter, ξ^i is a sequence of independent random variables. The iterative procedure for estimating the stochastic approximation parameter α when $x^0 = 0$ is as follows:

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

Thus, for the least squares estimators we have determined the iterative procedure that allows estimating the parameter by $(s+1)$ observations when the LSE based on s observations and the observation at the time $(s+1)$ are known. For the sake of completeness, we provide the statement about the convergence of the iterative procedure (4) with the probability one for the 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 known sequence dependent on the unknown parameter x^*. The problem is to estimate x^* using the observations y_k. We will 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 $\{c_s, s \geq 1\}$ and $\{\beta_s, s \geq 1\}$ are the sequences of positive numbers and $\inf c_s > 0$. The following statement is true [15].

THEOREM 1. Let the following conditions be satisfied:

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

(a) $\sum_{s=1}^{\infty} \beta_s c_s = \infty$,

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

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

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

(d) For any $\varepsilon > 0$ $\inf_{s \geq 1} \inf_{\varepsilon < \| x - x^* \| < \varepsilon^{-1}} \langle m_s(x) - m_s(x^*), x - x^* \rangle > 0$, 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 a strongly consistent estimator of x^*.

3. **The Least Absolute Deviations Method (LADM).** The least absolute deviations estimator (LADE) of the parameter α is a random vector α_n that minimizes the sum of the absolute values of the residuals

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

The LADEs coincide with the maximum likelihood estimators when the noise has the Laplace distribution. They are called robust estimators, i.e., estimators that have advanced stability with respect to the variations in the error distribution law, the presence of rough errors, etc. Moreover, the LADEs are much more efficient in the case of a small sample. A number of studies of the LADM qualitative behavior have been carried out recently. First, noteworthy is the consistency of the estimators. Simple and easy-to-test consistency conditions are provided in [11, 12]. Note that rather general models are being analyzed, where the parameter is usually contained in the regression function with a nonlinear maximum, and the noise may be dependent at different instants of time. Moreover, the statement of asymptotic normality is proved under rather natural conditions.

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

As an illustration, we will provide some statements about the asymptotic behavior of the LADE.

THEOREM 2. Let us 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 is 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:

(i) $\overline{\lim\limits_{n\to\infty}} \sup\limits_{\theta_1-\theta_2\in K} [\varphi_n(\theta_1,\theta_2) - \varphi(\theta_1,\theta_2)] \le 0$,

where $\varphi_n(\theta_1,\theta_2) = \frac{1}{n}\sum\limits_{i=1}^{n} |\varphi_i(\theta_1) - \varphi_i(\theta_2)|$, and $\varphi(\theta_1,\theta_2)$ is uniformly continuous at the diagonal $\theta_1 = \theta_2$.

(ii) $\overline{\lim\limits_{n\to\infty}} \varphi_n(\theta_1,\theta_2) \le \infty$.

(iii) $\lim\limits_{n\to\infty} \inf\limits_{\theta\in K\setminus\delta(\theta_0,\varepsilon^{n-1})} \left\{ \frac{1}{n}\sum\limits_{i=1}^{n} E_{\theta_0}|x_i - \varphi_i(\theta)| - \frac{1}{n}\sum\limits_{i=1}^{n} E_{\theta_0}|x_i - \varphi_i(\theta_0)| \right\} > 0$,

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

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

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

Theorem 2 is a simplified version of the statement about the strong consistency of the least absolute deviations, given in [11], where the rate of the convergence of the estimator to the real value is found.

THEOREM 3. Let the conditions of Theorem 2 and the following conditions be satisfied.

(i) $g_i = \frac{\partial\varphi_i(\theta)}{\partial\theta}$ exists.

(ii) The random values of ξ_i have a bounded probability density $h(x)$ and the inequality

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

holds. 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 $x \in R^1$ and $\Phi(x)$ is a Gaussian distribution function.

Let us formulate one more statement about the estimators of the least absolute deviations when the regression function depends on random values. This kind of model was analyzed in [16, 17, etc.].

Let us describe the observation model.

Let $\{(x_i, y_i), i \ge 1\}$ be a stationary (in a narrow sense) metrically transitive random process with a discrete parameter specified on the probability space (Ω, θ, P), $x_i \in R^k$, $y_i \in R^m$, $k \ge 1$, $m \ge 1$, and $\|\cdot\|_1$ be 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 following conditions are satisfied.

(i) $E\|x_i\|_1 < \infty$.

(ii) for any i, $E(x_i / F_i) = f(\theta, y_i)$ with probability one, 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 a fixed but unknown value from the closed set $I \subseteq R^l$, $l \ge 1$.

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

(iv) for any sequence $\{u_j, j\ge 1\}$ as $\|u_j\|_1\to\infty$, $j\to\infty$, $\|f(u_j,y_i)\|_1\to\infty$ with probability one as $j\to\infty$.

Let us introduce the notation:

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

(v) for any $j=\overline{1,k}$

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

(vi) for any $u\in I$, $u\ne\theta$, there exists $j: 1\le j\le k$ such that with probability one

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

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

Given the observations $\{(x_i,y_i)\}, 1\le i\le n\}$, we need to estimate the unknown parameter θ. Let us consider the LADE

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

The following theorem holds [11, 16, 17].

THEOREM 4. Let the conditions (i)–(vi) be satisfied. Then there exists at least one vector θ_n that satisfies condition (5) and

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

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

4. **The Minimax Method (MMM).** This method is often used to solve various problems, but these estimators have hardly been qualitatively analyzed. The difficulty is that it is impossible to apply the ergodic theorem and the central limit theorem to analyze the asymptotic properties. Nevertheless, due to the simplicity of the objective function, it is convenient to use these estimates at the preliminary stage for the calculation of the LSE and LADE. The practical expediency of this approach is noted in [10], where a software complex for the statistical processing of experimental data using the above-mentioned reasons is described. Note that wide use of the LADE and of the minimax estimator (MME) was suppressed by the calculation difficulties related to the minimization of nonlinear, nonsmooth, and nonconvex functions. In the case of the linear regression function, this difficulty can be avoided by reducing the criterion minimization problem to a linear programming one. The most recent achievements in nonsmooth optimization make it possible to solve the problem in the case of nonlinear regression function as well. In [10], the r-algorithm (the generalized gradient descent with the space dilation in the two consequent difference directions of subgradients ([18]) is described. The estimate is calculated under the constraints imposed on the set of parameters:

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

In the general case, finding the global optimum is a difficult problem because of the nonconvexity of the objective function. That is why it is expedient in practice to take into account some a priori information about the localization of the global minimum. A natural way is to divide the problem into two parts. The first stage is to find an approximate problem solution under some simplifying prepositions. The MMM can be used as a first approximation. In the absence of rough errors, the MME is known to approximate the LADE rather well since it is robust but has a simpler objective function. Then at the second stage, the minimum point of the original problem can be found by the nonsmooth optimization method.

Let us mention one more significant fact. As a rule, the type of the regression function cannot be described accurately enough based on the initial or visual data. That is why a large base of approximating functions that allows choosing the most suitable one is necessary. The more complete the set of approximating functions, the more possibilities for a more accurate estimation of the regression function.

Let us describe one more model of nonparametric estimation of regression function in some functional space. In our opinion, it rather adequately reflects the physical processes mentioned above. Suppose that the observations of the soil or another object take place at discrete instants of time with independent errors, and we need to determine the trend at any time on the entire observation interval. In other words, we need to recover the regression function in an optimal way with respect to some criterion, based on its distorted observations at discrete instants of time with additive noise. The MLM, LSM, and LADM can be taken as the criteria. A more exact formulation is as follows.

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

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

We will consider the LSE or LADE as the 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 the problems of minimization by criteria (7) and (7′) are more difficult than those considered above because the optimum is being found on some class of functions. Such estimation problems are called nonparametric. Their solution involves a lot of difficulties, both computing and qualitative, in the analysis of the statistical properties. However, numerical methods for finding the optimum are developed for these models as well. Among the studies in this field, noteworthy is the interesting approach developed in [19].

For a set of continuous functions in the space with a uniform metric, the LSE is analyzed in [12] and the LADE in [11]. Let us consider model (6) with criterion (7′) in more detail, assuming that the unknown function belongs to some set from the Hilbert space. The proofs of the basic statements will give an idea of the methods of analysis of such problems and the mathematical technique used there. We will prove the statement about the strong consistency and the functional limit theorem for the standardized estimators. In particular, we will prove the weak convergence of the standardized estimates to the standard statement of the Wiener process. Therefore, any linear functional of the estimator converges to the linear functional of the Wiener process. In particular, it is rather important in practice to know the distributions

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

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

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

$$= 1 - \frac{2}{\sqrt{2\pi T}} \sum_{n=0}^{\infty} \left[\int_{(2n+1)(b-a)-a}^{(2n+1)(b-a)+a} \exp\left\{-\frac{y^2}{2T}\right\} dy + \int_{(2n+1)(b-a)-b}^{(2n+1)(b-a)+b} \exp\left\{-\frac{y^2}{2T}\right\} dy\right]. \tag{8}$$

Thus, we can calculate the probabilities of locating inside a band of some characteristic. It may be the content of phosphates in the soil, heavy metals (for instance, cadmium), or the amount of different salts on the soil surface or in the depth. Moreover, given the distribution of the functional, we can find the first jump over the threshold, which can characterize the critical level of the inclusion of some substance.

Let us now 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 and satisfying the conditions:

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

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

2. For each $n \ge 1$, ξ_{jn}, $0 \le j \le n$, 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, $\{F_n, n \ge 1\}$ be a sequence of σ-algebras such that $F_n \subset F_{n+1}$, $n \ge 1$, and K be a compact subset of some Banach space with the 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:

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

(ii) for fixed n and ω, the function $Q_n(S, \omega)\colon K \to R$ 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

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

(iii) for some fixed element $S_0 \in K$ and for each $S \in K$, the relation

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

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

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

(iv) for any $\delta > 0$, there exists $\gamma_0 > 0$ and a function $c(\gamma)\colon R \to R$, $c(\gamma) \to 0$, $\gamma \to 0$, 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

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

The following assertion follows from Theorem 5 [12].

THEOREM 6. Assume that the conditions of Theorem 4 are satisfied and that

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

Then

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

Let us analyze the distribution of the functionals of the estimators α_n. Assume additionally that

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

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

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

The following auxiliary assertions take place [12].

LEMMA 1. Let the condition of Theorem 6 be satisfied 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

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

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

LEMMA 2. Let the conditions of Lemma 1 be satisfied and let

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

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

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

Lemma 1 and Lemma 2 imply the following theorem.

THEOREM 7. Assume that the conditions of Lemma 2 are satisfied. Then the sequence of random processes $\eta_n(t)$ weakly converges to a standard Wiener process as $n\to\infty$.

Problem of Attaining the Critical Threshold. Let us now turn to certain problems related to the behavior of some ecological systems, including those described above. As indicated in [20], real ecological systems are subject to various random influences. If the time of these random influences or perturbations is considerably less than the system operation time, it is possible to apply the rather developed tool of Markov processes to analyze the entire system (the dynamics of its development). Under these conditions, white noise is a natural model of random perturbations. For these perturbations, there is a well-developed mathematical apparatus that allows us to describe systems' dynamics as a stochastic differential equation:

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

where $a(t,x)$ is a drift coefficient corresponding to the determinate part of the system's development, and $\xi(t)$ is a white noise. Taking into account a random influence combined with the white noise (which is an idealized random process with a small time correlation) allows us to use the well-developed apparatus of the stochastic equation theory and to

obtain qualitative and quantitative results in modeling the behavior of process (9). Equation (9) determines the diffusion process with the densities of nonstationary probabilities satisfying the Kolmogorov equations (direct and inverse ones). For these processes, we can estimate some characteristics of the behavior of the ecological system mentioned above; e.g., the problems related to attaining the critical threshold by the system. Let us dwell on this problem.

Let (r_1, r_2) be the interval of feasible changes of $x(t)$ and T be the instant of time when $x(t)$ leaves the interval (r_1, r_2) for the first time. A significant assumption is that the probability of $x(t)$ abandoning the interval (r_1, r_2) for the first time is strongly positive, i.e.,

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

If $T < \infty$ then we say that the bound may be attained in a finite time. This bound is called attractive. It is evident that there are two interesting groups among attractive bounds: (i) absorption bounds and (ii) permitting bounds. As to the type of bounds, there are some results in the stochastic equation theory that provide a full and complete reply under some conditions, for instance, in [21, 22]. We will not dwell on the results related to the time of intersecting some level (threshold). Let us turn our attention to some of them.

For simplicity, we will consider a uniform in time stochastic equation, assuming that in (16)

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

The following statement 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 solution of the 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 of the theory of diffusion processes. It uses Markov moments and stochastic differential calculation apparatus [21].

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

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 the previous theorem.

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

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

be the Laplace transform 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 the Laplace transform, we can write the other characteristics of the variable $\tau_x[a,b]$.

One of the most important questions is calculating the probability of the first exit of the random process $x(t)$ from some interval. First of all, we are interested in finding the probability of the random process $x(t)$ being at point a earlier than at point b and vice versa. It is obvious that when $x \notin (a,b)$, the random process $x(t)$ appears at the nearest point earlier with probability 1. Therefore, we will suppose that $x \in (a,b)$. Let $P_a(x,b)$ be the probability that the random process $x(t)$ appears at point a earlier than at point b, when it is at the point x for the first time.

The following statement 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 solution of the equation

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

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

Proof. Let

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

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

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

According to the 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,

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

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

The theorem is proved.

COROLLARY 1. Let $\Phi(z)$ be a Laplace function. Then

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

To prove (11), it will suffice 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:

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

Note that the probability that the random process $x(t)$ stays within the interval $[a,b]$ is given by the formula (8) in the case of the Wiener process. By virtue of the Girsanov theorem [23], under the natural conditions, by replacing the measure and passing to another probability space, the space $x(t)$ will also be a Wiener-transformed measure, relatively. And then we can apply (8) to the obtained process.

So, using model (16), we can estimate the time of onset of the critical level of saturation of some element (for instance, the maximum permissible level of phosphorus in the soil) and take pre-planned precautions.

Note that in the previous section, we touched upon the problem of exceeding the level in the model described by Eq. (16). Rather extensive studies in finding the time of attaining some level and staying within the interval were carried out for stationary, Gaussian, and Poisson processes and these results will be used for sure in the case of an adequate model.

More detail can be found in [24], but we cannot dwell on them right now. We can only provide some facts that are useful for the solution of the problem given above.

Consider a stationary Gaussian process. The main idea is to compare this process with a normal process $x^*(t)$ whose maximum can be easily calculated. As an example, let us consider the process

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

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

It is obvious that the random process $x^*(t)$ is Gaussian. It can be easily rearranged as

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

For such processes, the following statement is true.

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 $0 < T < \pi/\omega$ and $u > 0$.

Proof. Let $N = N_u(t)$ be the number of times the process $x^*(t)$ abandons the level u on 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

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

Equality (13) is equivalent to (12); therefore, the lemma is proved.

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

LEMMA 4. Let $x_1(t)$ and $x_2(t)$ be Gaussian continuous processes with $Ex_1(t) = Ex_2(t) = 0$ and $Ex_1^2(t) = Ex_2^2 = 1$ and let $r_1(t,s)$ and $r_2(t,s)$ be 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 the maximums $M_1(t)$ and $M_2(t)$ of these processes 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 uses the piecewise-linear approximation of processes $x_i(t)$ by processes $x_n^i(t)$, taking the values $x_i(t)$ on the set $\{jq_n,\ j = 0, 1, \ldots\}$ and being linear inside the intervals $(jq_n,\ jq_{n+1})$, $q_n \to 0$, $n \to 0$. 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 can be easily seen 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 statement.

THEOREM 11 [25]. Let $u, T \to \infty$ so that $\frac{T}{2\pi}\lambda_2^{1/2}\exp\left\{-\frac{u^2}{2}\right\} \to \tau \ge 0$ and the correlation function $r(t)$ of the process $x(t)$ satisfy 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

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

We have analyzed the properties of the extremums of the Gaussian process without obtaining the information about the location of the extremums.

Let $L(t)$ be a point at which the process $x(t)$ attains its maximum for the first time on the interval $[0, T]$. We may suppose that $L(t)$ has a uniform distribution for any stationary process. But this is generally incorrect. This takes place if $x(t) = A(t - \varphi)$, where φ is a random variable uniformly distributed on $[0, 2\pi]$. For the stationary Gaussian process, $L(t)$ is always symmetric on $[0, T]$, and the possible jumps at points 0 and T have equal values. One of the methods to eliminate points like 0 and T is to move them away to infinity. The question arises whether $L = L(t)$ is asymptotically uniform as $T \to \infty$. For Gaussian stationary processes, this fact is a consequence of the asymptotic independence of the maximums on the intersecting intervals. More precisely, the following statement takes place.

THEOREM 12. Let $x(t)$ be 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. W. M. Stigliani, "Changes in valued "Capacities" of soils and sediments as indicators of nonlinear and time-delayed environmental effects," Environmental Monitoring and Assessment, Vol. 10, Iss. 3, 245–307 (1988).
2. R. L. Kashyap and A. R. Rao, Dynamic Stochastic Models from Empirical Data, Academic Press, New York–San Francisco–London (1976).
3. R. E. Kalman and R. S. Bucy, "New result in linear filtering and prediction theory," J. Basic Eng., Vol. 83 (1), 95–108 (1961).
4. R. Sh. Liptser and A. N. Shiryaev, The Statistics of Random Processes [in Russian], Nauka, Moscow (1974).
5. J. H. Laning and R. H. Battin, Random Processes in Automatic Control. McGraw-Hill Book Company Inc., New York–Toronto–London (1956).
6. S. Darlington, "Nonstationary smoothing and prediction using network theory concepts," IRE Trans. on Circuit Theory, Vol. 6, Iss. 5, 1–13 (1959).
7. A. Ya. Dorogovtsev, "Remarks on the processes generated by some differential equations," Visnyk KDU, Vol. 1, No. 3, 124–146 (1962).
8. P. S. Knopov, "On forecast of random fields," Teoriia Optimalnykh Reshenii, No. 3, 74–83 (1963).

9. I. I. Gikhman and A. V. Skorokhod, An Introduction to the Theory of Random Processes [in Russian], Nauka, Moscow (1965).
10. T. A. Bardadym, A. V. Borshchevskyi, L. B. Vovk, and A. V. Ivanov, The Complex of Programs for Statistical Processing of Experimental Data [in Russian], Inst. Cybern. AS USSR, Kyiv (1991). (Prepr. AN USSR, Institute of Cybernetics; 91–41).
11. N. N. Leonenko and A. V. Ivanov, Statistical Analysis of Random Fields [in Russian], Vyshcha Shkola, Kyiv (1989).
12. A. Ya. Dorogovtsev, Theory of Estimates of the Parameters of Random Processes [in Russian], Vyshcha Shkola, Kyiv (1982).
13. Yu. M. Ermoliev, Stochastic Programming Methods [in Russian], Nauka, Moscow (1976).
14. M. T. Wasan, Stochastic Approximation, Cambridge Univ. Press, Cambridge (1969).
15. Yu. M. Kaniovskii, P. S. Knopov, and Z. V. Necrylova, The Limit Theorem for The Stochastic Programming Method [in Russian], Naukova Dumka, Kyiv (1980).
16. J. Dupacovï and R. J.-B. Wets, "Asymptotic behavior of statistical estimators and optimal solutions for stochastic optimization problem," IIASA WP 86–41, Laxenburg, Austria (1986).
17. E. I. Kasitskaya and P. S. Knopov, "About one approach to the nonlinear estimating problem," in: Proc. 6th USSR–Japan Symp. Probability Theory and Mathematical Statistics (Kyiv, USSR, August 5–10, 1991), Kyiv (1991), pp. 151–157.
18. B. I. Mudrov and V. L. Kushko, Measurement Processing Methods [in Russian], Radio i Svyaz', Moscow (1983).
19. A. A. Gaivoronskii, A. N. Golodnikov, and D. Le Fung, "Solution of limiting extremal problems for minimizing functionals," Cybern. Syst. Analysis, Vol. 16, No. 1, 97–103 (1980). https://doi.org/10.1007/BF01099368.
20. Yu. M. Shiryaiev, Nonlinear Waives, Dissipative Structures and Catastrophes in Ecology [in Russian], Nauka, Moscow (1987).
21. I. I. Gihman and A. V. Skorokhod, Stochastic Differential Equations [in Russian], Naukova Dumka, Kyiv (1968).
22. K. Ito, Probability Processes, Vol. 2 [Russian translation], Inostr. Lit., Moscow (1963).
23. I. V. Girsanov, "On transforming a certain class of stochastic processes by absolutely continuous substitution of measures," Theory of Probability and Its Applications, Vol. 5, No. 3, 285–301 (1960).
24. M. R. Leadbetter, G. Lingren, and H. Rootzen, Extremes and Related Properties of Random Sequences and Processes. Springer-Verlag, New York–Heidelberg–Berlin (1986).
25. D. Slepian, "The one-sided barrier problem for Gaussian noise," Bell System Technical J., Vol. 41, Iss. 2, 463–501 (1962).