Advanced Statistical Approach for the Mathematical Modeling of Transfer Processes in a Layer Based on Experimental Data at the Boundary

Olha Chernukha [1,2], Petro Pukach [2], Halyna Bilushchak [2], Yurii Bilushchak [1,2] and Myroslava Vovk [3,*]

1 Pidstryhach Institute for Applied Problems of Mechanics and Mathematics, National Academy of Sciences of Ukraine, 3-b, Naukova Str., 79060 Lviv, Ukraine; cher@cmm.lviv.ua (O.C.); bil@cmm.lviv.ua (Y.B.)
2 Department of Computational Mathematics and Programming, Institute of Applied Mathematics and Fundamental Sciences, Lviv Polytechnic National University, 12 Bandera Str., 79013 Lviv, Ukraine; petro.y.pukach@lpnu.ua (P.P.); halyna.i.bilushchak@lpnu.ua (H.B.)
3 Department of Advanced Mathematics, Institute of Applied Mathematics and Fundamental Sciences, Lviv Polytechnic National University, 12 Bandera Str., 79013 Lviv, Ukraine
* Correspondence: myroslava.i.vovk@lpnu.ua

Abstract: In this work, a mathematical model of the transfer process in a layer under the condition of given experimental data on a part of the layer boundary is presented and investigated. Such research is important for the mathematical description of the objects and systems for which, based on physical considerations, it is impossible to correctly impose boundary or initial conditions, even in a sufficiently general form, but there are experimental data on the desired function or its derivative at the boundary of the body or at the initial time. The values of the desired function at the boundary are known at certain moments in time. The boundary condition is constructed by the experimental data and the initial-boundary value problem, with such a boundary condition, is formulated and solved. The influence of the statistical characteristics of the sample of experimental data on the solution to the initial-boundary value problem is analyzed, and a two-sided statistical estimation of the solution is determined. The confidence intervals for the coefficients of the regression equation and the corresponding confidence intervals for the sought function are established. The influence of the statistical characteristics of the sample on the sought function at the lower boundary of the layer is investigated. Numerical analysis of the solution to the initial-boundary value problem is carried out depending on the statistical characteristics of the sample. Various cases of samples by size and variance are considered. Numerical solutions are studied under the conditions of large and small time intervals of the considered process.

1. Introduction

The further development of approaches and methods of mathematical description of non-equilibrium processes of various physical nature in natural or artificial objects [1,2] is caused by the need to construct effective methodologies and estimates for modeling a number of socio-economic problems of the functioning of society, the economy, the defense capability of the state, forecasting the redistribution of pollution of anthropogenic origin in objects in natural environment, assessing the quality of drinking water and improving its purification on an industrial scale, establishing the influence of the diffusion of aggressive substances when assessing the reliability and durability of the operation of macro-structural elements and components to prevent the destruction of relevant materials (in particular, in dual-use technologies), etc.

Such research makes it possible to obtain a reliable forecast of the processes occurring in environmental objects and other systems, to take the necessary measures in time to

prevent their negative development. Without pretending to be complete, we will briefly dwell on the current practical and theoretical results obtained using statistical modeling methods [3–5] in recent years.

In [6], a mechanistic–statistical approach based on a combined reaction–diffusion model, which represents the dynamics of an organism in the growth area, is applied. The initial conditions and model parameters related to diffusion, reproduction, and mortality are estimated simultaneously in a Bayesian system using the adaptive sampling algorithm. In [7], an estimation of the parameter for population models based on partial differential equations is presented using a mechanistic–statistical model that combines a sub-model describing the studied dynamics and a stochastic sub-model describing the observation process. The problem of modeling the movement of the fish population is examined in [8]. The model has both deterministic and stochastic components. Modeling is carried out by calculating spatial statistics.

The authors of the work in [9] investigated statistical models that use partial differential equations (PDE) to describe dynamically developing natural systems. In [10], a data-driven methodology for identifying sources of pollutants in the atmosphere is presented. This approach combines a diffusion model with a quantitative probabilistic assessment of data uncertainty. The work in [11] is devoted to the use of ecological models based on partial differential equations, with a special emphasis on reaction–diffusion models. The choice of parameters is based on the definition of a probabilistic observation model and the appropriate data representation. A mathematical–statistical analysis of partial differential equations is carried out in [12] to analyze the impact of organism movements on the spatial dynamics of the population. Statistical techniques have been developed, that allow for a more accurate correlation of the conclusions of mathematical ecology and observations of empirical ecology. Because ecological processes evolve over time, they require statistical models that can adapt to the dynamics of change as new data are collected. Work [13] was devoted to the development of a model that combines the equations of ecological diffusion and logistic growth to characterize the processes of population colonization and to establish long-term equilibrium using a heterogeneous medium. Hierarchical Bayesian modeling is used in [14] for statistical inferences and probabilistic forecasts using the mechanical ecological diffusion model. The statistical model led to important ecological conclusions and proved to be an accurate forecasting method.

Anisotropic diffusion is the theoretical basis for damage removal in the problems of sensing and image processing applications. However, diffusion approaches require the choice of a special function at the boundary of the image region, the definition of which is usually problematic. On the basis of statistical image data in [15], the values of the image at the edge of the region were obtained, which made it possible to formulate problems of anisotropic diffusion in which the parameters of the image at the edge were studied from the training data. Statistical modeling was used in [16] to study reaction–diffusion equations for the purpose of image processing. In the article [17], in order to study the law of diffusion of cementation of cracks, the form of diffusion of cementation and the relationship between the radius of diffusion and time depending on the width of the crack were investigated using laboratory experiments and statistical analysis. To investigate the migration of heavy metals, nonlinear diffusion methods and an empirical method using a polynomial equation were applied in [18]. Experimental data representing heavy metal release were used for nonlinear models for calibration purposes. In work [19], the processes of natural carbonation of concrete samples of Portland cement over a long period were studied. A complex approach, consisting of the thermodynamic model and the statistical methodology for modeling long-term carbonation, was developed. Statistical modeling was used in [20] to investigate the process of chloride ions entering concrete, which is the main cause of concrete corrosion. The proposed method can be used to model the profile of chloride ions in concrete using only a few data samples for a given depth. Then, the stochastic diffusion of chloride ions can be modeled by the diffusion equation.

The statistical method for constructing cellular automata based on observation data was used in [21] to apply the diffusion equation and the Burgers equation. The study cited in [22] discussed the use of structure-of-research methods to identify statistical dependencies in high-dimensional physical processes. Large data sets for many processes in space and time require statistical models and methods that can adequately take into account such data. It was shown in [23] that the solution to the stochastic differential equation of advection and diffusion provides a flexible class of models for spatial–temporal processes, which is also computationally feasible for large data sets. Unlike the raw samples, the forecasts after statistical processing are calibrated, quantify the forecast uncertainty, and have a smaller mean absolute error.

The important practical problems described above led to the emergence and development of a number of new theoretical methods and methodologies of statistical modeling in recent years. In particular, in the monograph [24], a systematic description of various mass transfer models is carried out based on a statistical description where the input parameters and solutions are expressed by random processes and fields. A mathematical formulation of the main physical models of transport, diffusion, and spreading is carried out, and some analytical tools for statistical modeling are developed—in particular, in randomly layered media [25]. The algorithm for estimating the coefficients of the parabolic equation based on statistical observational data is presented. The solution to the diffusion problem is based on a probabilistic representation in the form of a functional of solution to a system of certain stochastic differential equations. The paper cited in [26] was devoted to the investigation of a linear parabolic stochastic partial differential equation of the second order with a small variance parameter using high-frequency data. Statistical conclusions for diffusion processes based on modeling results were obtained. In the work [27], a statistical scheme for observing the diffusion processes was proposed, which involved the convolution of diffusion processes and some nuclear functions in time. In [28], various procedures for statistical choice of input data that demonstrated adequate modeling of the evolution equation with a finite sample were presented. The authors of [29] considered the problem of uncertainty estimation in linear statistical inverse problems with high-dimensional parameters. In [30], a fractional Laplace operator was used to study anomalous diffusion processes, which leads to the problems described by fractional equations of nonlinear diffusion. The approach based on statistical linearization was proposed, which permits one to calculate, approximately and iteratively, the statistical characteristics of the diffusion process. The work in [31] was devoted to the consideration of various methods of visualization of data, which lie in layered structures glued together along the lower dimensional boundaries. Diffusion processes were used to represent noise in statistical models in spaces, for which there are no standard parametric probability distributions.

As the above analysis of sources on statistical modeling shows, it is not always possible to correctly impose boundary conditions on the region boundaries which proceed from physical considering, even in a sufficiently general way. This is due to the complexity and insufficiency (or impossibility) of carrying out relevant studies, and therefore, the analysis and necessary generalizations are lacking.

This work considers the initialboundary value problem of a parabolic type, which describes the processes of transfer of heat, mass, charge, etc., in a layer when experimental data on the desired function are available at one of the boundaries.

The purpose of the work is to obtain and investigate the mathematical model of the transfer process in a layer, under the condition that, on one of the two parts of the boundary of the layer, only experimental data are known regarding the value of the sought function at certain moments of time. The initial-boundary value problem is formulated and solved for a second-order parabolic equation with a boundary condition based on such experimental data. Moreover, the actual problem investigated in the work is the analysis of the influence of the statistical characteristics of the sample of experimental data on the solution to the initial-boundary value problem, the determination of a two-sided statistical estimation of the solution, and the establishment of confidence intervals. The aim of the investigation is

also to numerically analyze the solution to the initial-boundary value problem, depending on the statistical characteristics of the sample.

2. Statement of the Initial-Boundary Value Problem of Transfer Processes under Experimental Data at the Layer Boundary

We consider the physical process of transfer (of heat, mass, charge, etc.), which is described by a two-order partial differential equation, on the interval $[0, x_0]$, $x_0 \in R_+$. We assume that the coefficients of the problem are constant in the body region.

Conditions of the first kind are imposed on the sought function, but the system is under conditions of uncertainty. At the lower boundary of the layer, only experimental data on the desired function at certain moments of time are known.

In the case of a one-dimensional spatial coordinate, the transfer process is described by the following partial differential equation [32]:

$$\rho \frac{\partial f(t, x)}{\partial t} = d \frac{\partial^2 f(t, x)}{\partial x^2}, \tag{1}$$

where $f(t, x)$ is the desired function, ρ and d are the constant coefficients, t is time, and x is the spatial coordinate.

We assume that the initial condition is zero:

$$f(t, x)|_{t=0} = 0. \tag{2}$$

For $t > 0$ at the upper boundary of the layer, the constant source f_* acts:

$$f(t, x)|_{x=0} = f_* \equiv const. \tag{3}$$

At the lower boundary of the layer, the experimental data are known at N moments of time, as presented in Table 1.

Table 1. Template for experimental data.

t	t_1	t_2	...	t_i	...	t_N
$f(t)|_{x_0}$	$f_{x_0}(t_1)$	$f_{x_0}(t_2)$	...	$f_{x_0}(t_i)$	...	$f_{x_0}(t_N)$

Based on these experimental data, we construct a linear regression model [33,34]:

$$F(t) = at + b, \tag{4}$$

where the coefficients a and b are found by the method of least squares [35]

$$a = \frac{N \sum\limits_{i=1}^{N} t_i f_{x_0 i} - \left(\sum\limits_{i=1}^{N} t_i\right)\left(\sum\limits_{i=1}^{N} f_{x_0 i}\right)}{N \left(\sum\limits_{i=1}^{N} t_i^2\right) - \left(\sum\limits_{i=1}^{N} t_i\right)^2}, \tag{5}$$

$$b = \frac{\left(\sum\limits_{i=1}^{N} t_i^2\right)\left(\sum\limits_{i=1}^{N} f_{x_0 i}\right) - \left(\sum\limits_{i=1}^{N} t_i\right)\left(\sum\limits_{i=1}^{N} t_i f_{x_0 i}\right)}{N \left(\sum\limits_{i=1}^{N} t_i^2\right) - \left(\sum\limits_{i=1}^{N} t_i\right)^2}, \tag{6}$$

where $f_{x_0 i} = f_{x_0}(t_i)$, and the boundary condition at $x = x_0$ takes the form

$$f(t, x)|_{x=x_0} = F(t). \tag{7}$$

3. Finding the Solution to the Transfer Initial-Boundary Value Problem under Linear Regression at the Boundary

We reduce the initial-boundary value problems (1)–(3) and (7) to the problem with zero boundary conditions by substitution:

$$w(t, x) = f(t, x) - f_* \left(1 - \frac{x}{x_0}\right) - F(t)\frac{x}{x_0},$$

(8)

where $w(t, x)$ is a new desired function.

Then, we obtain the following problem:

$$\frac{\partial w(t, x)}{\partial t} + \frac{\partial F(t)}{\partial t}\frac{x}{x_0} = D\frac{\partial^2 w(t, x)}{\partial x^2}$$

(9)

with the initial condition

$$w(t, x)|_{t=0} = -f_* \left(1 - \frac{x}{x_0}\right) - F(t)|_{t=0}\frac{x}{x_0}$$

(10)

and zero boundary conditions

$$w(t, x)|_{x=0} = w(t, x)|_{x=x_0} = 0,$$

(11)

wherein $D = d/\rho$.

Let us apply the finite integral Fourier sin-transform [36] ($x \to y_n = n\pi/x_0$, $w(t, x) \to \widetilde{w}(t, y_n)$) to the initial-boundary value problem (9)–(11). Then, in the images, we obtain:

$$\frac{d\widetilde{w}(t, y_n)}{dt} + Dy_n^2\widetilde{w}(t, y_n) = \frac{dF(t)}{dt}\frac{(-1)^n}{y_n},$$

(12)

$$\widetilde{w}(t, x)|_{t=0} = \frac{-1}{y_n}\left(f_* + (-1)^{n+1}F(t)\right).$$

(13)

The solution to the problem in (12) and (13) is the following expression

$$\widetilde{w}(t, y_n) = -\frac{(-1)^n}{y_n}F(t) - e^{-Dy_n^2 t}\left[\frac{f_*}{y_n} - 2\frac{(-1)^n}{y_n}F(t)|_{t=0}+\right.$$

$$\left. + (-1)^n Dy_n \int F(t)e^{Dy_n^2 t}dt\Big|_{t=0} - \frac{(-1)^n}{y_n}Dy_n^2 \int F(t)e^{Dy_n^2 t}dt\right].$$

(14)

After applying the inverse Fourier transform, using (8), to relation (14), we obtain [37].

$$f(t, x) = f_*\left(1 - \frac{x}{x_0}\right) + F(t)\frac{x}{x_0} - \frac{2}{x_0}\sum_{n=1}^{\infty}\left[\frac{(-1)^n}{y_n}F(t)+\right.$$

$$+ e^{-Dy_n^2 t}\left[\frac{f_*}{y_n} - 2\frac{(-1)^n}{y_n}F(t)|_{t=0} + (-1)^n Dy_n \int F(t)e^{Dy_n^2 t}dt\Big|_{t=0} -\right.$$

$$\left.\left. - \frac{(-1)^n}{y_n}Dy_n^2 \int F(t)e^{Dy_n^2 t}dt\right]\right]\sin(y_n x).$$

(15)

If we take into account that $F(t)$ is a linear regression, i.e., Formula (4) holds, then (15) can be specified as follows:

$$(-1)^n y_n e^{-Dy_n^2 t}\int F(t)\,e^{Dy_n^2 t}\,dt = \frac{(-1)^n}{Dy_n}\left[at + b - \frac{a}{Dy_n^2}\right];$$

$$(-1)^n y_n e^{-Dy_n^2 t} \int F(t) \, e^{Dy_n^2 t} \, dt \Big|_{t=0} = \frac{(-1)^n}{Dy_n} \left[b - \frac{a}{Dy_n^2} \right];$$

$$f(t,x) = f_* \left(1 - \frac{x}{x_0} \right) + (at + b) \frac{x}{x_0} - \frac{2}{x_0} \sum_{n=1}^{\infty} \left[\frac{(-1)^n}{y_n} (at + b) + \right.$$

$$+ e^{-Dy_n^2 t} \left[\frac{f_*}{y_n} - \frac{(-1)^n}{y_n} (2b + at) \right] \sin(y_n x). \tag{16}$$

The resulting Formula (16) is valid for $t \in [0, t_N]$, since this is the time interval in which the sample of experimental data for the desired function is given.

Let us notice that the coefficients of linear regression a and b can acquire positive and negative values, and can be equal to zero.

4. Analysis of Statistical Characteristics of a Sample in the Case of Linear Regression

Let us analyze the effect of the statistical characteristics of the sample of experimental data set at the lower boundary of the layer on the coefficients of the linear regression (4), constructed using these data by the least squares method.

In the expressions (5) and (6), for the regression coefficients a and b, we consider the case where the sample size $N \to \infty$. Then, we obtain the following asymptotic relations:

$$\lim_{N \to \infty} a = \frac{\sum\limits_{i=1}^{\infty} t_i f_{x_0 i}}{\sum\limits_{i=1}^{\infty} t_i^2}; \quad \lim_{N \to \infty} b = 0. \tag{17}$$

We note that, if the number of experimental measurements changes over the same time interval, then the values of the linear regression coefficients are a and b. That is, comparing (5), (6), and (17), we can conclude that the larger the sample size N, the closer the angle of inclination of the regression line a becomes to the expression (17) and the smaller the absolute values of the free term b grow. Accordingly, for the samples, which the size is large enough, it is necessary to construct a linear regression by means of the method of least squares for the deviations of experimental values from their first value $f_{x_0}(t_i) \mp f_{x_0}(t_1)$, depending on $\mathrm{sgn} f_{x_0}(t_1)$.

We also note that the coefficient a can be presented in the following form [37]:

$$a = \frac{\mathrm{cov}(t, F_e(t))}{\sigma_t^2}, \tag{18}$$

where $\mathrm{cov}(t, F_e(t))$ is the covariance between t and $F_e(t)$ [38] and σ_t^2 is the variance of variable t.

It follows from Formula (18) that the coefficient a is directly proportional to the covariance of the variables t and $F_e(t)$, and it is inversely proportional to the variance of the variable t. And the smaller the variance σ_t^2, the higher the value of the parameter a.

Since $b = \frac{1}{N} \left(\sum\limits_{i=1}^{N} F_{ei} - a \sum\limits_{i=1}^{N} t_i \right)$, this coefficient can be presented in terms of covariance and variance as follows:

$$b = \frac{1}{\sum\limits_{i=1}^{N} t_i} \left(\sum\limits_{i=1}^{N} F_{ei} - N \frac{\mathrm{cov}(t, F_e(t))}{\sigma_t^2} \right). \tag{19}$$

It follows from Formula (19) that a decrease in variance under the same number of experiments can lead to both an increase and a decrease in the value of b, depending on the sign of the covariance and the magnitude of the values of F_{ei}. It should be noted that an increase in positive covariance reduces the value of the free term of the regression equation b, while a negative covariance increases it.

Also, linear regression coefficients can be presented using the correlation coefficient R:

$$a = R\frac{\sigma_F}{\sigma_t}, \quad b = \frac{1}{N}\left(\sum_{i=1}^{N} F_{ei} - R\frac{\sigma_F}{\sigma_t}\sum_{i=1}^{N} t_i\right), \tag{20}$$

where $R = \dfrac{N\sum\limits_{i=1}^{N} t_i F_{ei} - \left(\sum\limits_{i=1}^{N} t_i\right)\left(\sum\limits_{i=1}^{N} F_{ei}\right)}{\sqrt{\left(N\sum\limits_{i=1}^{N} t_i^2 - \left(\sum\limits_{i=1}^{N} t_i\right)^2\right)\left(N\sum\limits_{i=1}^{N} F_{ei}^2 - \left(\sum\limits_{i=1}^{N} F_{ei}\right)^2\right)}}$, $\sigma_t = \dfrac{1}{N}\sqrt{N\left(\sum\limits_{i=1}^{N} t_i^2\right) - \left(\sum\limits_{i=1}^{N} t_i\right)^2}$, and

$\sigma_F = \dfrac{1}{N}\sqrt{N\left(\sum\limits_{i=1}^{N} F_{ei}^2\right) - \left(\sum\limits_{i=1}^{N} F_{ei}\right)^2}$ are the mean square deviations of values t and $F_e(t)$,

correspondingly.

It follows from (20) that the coefficient of linear regression a is directly proportional to the correlation coefficient R, and its sign determines the sign of the parameter a (since the mean square deviation is always a non-negative quantity). At the same time, absolute values of a increase with increasing σ_F and decreasing σ_t. As for the coefficient b, when R and $\bar{t}$ are of the same sign, an increase in the correlation coefficient leads to a decrease in this parameter. In the case of different signs of the quantities R and $\bar{t}$, an increase in the correlation coefficient leads to an increase in the coefficient b.

If the coefficients a and b in the form of (18) and (19) are substituted into the expression for linear regression, then we obtain the following:

$$F_e(t) = \frac{1}{N\left(\sum\limits_{i=1}^{N} t_i^2\right) - \left(\sum\limits_{i=1}^{N} t_i\right)^2}\left\{\left[N\sum_{i=1}^{N} t_i F_{ei} - \left(\sum_{i=1}^{N} t_i\right)\left(\sum_{i=1}^{N} F_{ei}\right)\right]t+\right.$$

$$\left.+\left(\sum_{i=1}^{N} t_i^2\right)\left(\sum_{i=1}^{N} F_{ei}\right) - \left(\sum_{i=1}^{N} t_i\right)\left(\sum_{i=1}^{N} t_i F_{ei}\right)\right\}. \tag{21}$$

The asymptote (at $N \to \infty$) of the obtained expression (21) for the time interval $t \in [0; t_N]$ has the form:

$$\lim_{N\to\infty} F_e(t) = \frac{t\sum\limits_{i=1}^{\infty} t_i F_{ei}}{\sum\limits_{i=1}^{\infty} t_i^2}, \tag{22}$$

which is proportional to the running value of the time variable t, and, if the experimental values for F_{ei} are of the same sign, then it determines the sign of the asymptote.

We also note that there is no steady-state regime for the function $F_e(t)$, since the inequality $0 \leq t \leq t_N$ needs to be satisfied.

5. Two-Sided Statistical Estimation for Solution of the Initial-Boundary Value Problem

Let us determine a two-sided estimate of the solution to the problems in (1)–(3) and (7), which is given using linear regression on the layer boundary $x = x_0$ in (16).

We enter the notation:

$$g(t, x) = f(t, x) - f_*\left(1 - \frac{x}{x_0}\right) - 2(at + b)\frac{x}{x_0}. \tag{23}$$

Then, we obtain the following from Formula (14):

$$g(t, x) = -\frac{2}{x_0}\sum_{n=1}^{\infty} e^{-Dy_n^2 t}\left(\frac{f_*}{y_n} - \frac{(-1)^n}{y_n}(2b + at)\right)\sin(y_n x). \tag{24}$$

We use the estimate for the exponent $e^{-Dy_n^2 t}$

$$0 \leq e^{-Dy_n^2 t} \leq 1 \text{ for } t \geq 0 \tag{25}$$

and take into account that the function $g(t, x)$, according to Formula (24), is negative. Then, we obtain the following estimate:

$$\frac{2}{x_0} \sum_{n=1}^{\infty} \left(\frac{(-1)^n}{y_n}(b + at) - \frac{f_*}{y_n} \right) \sin(y_n x) \leq g(t, x) \leq 0.$$

If we sum the two series included in the last inequality [39], then we obtain:

$$-(2b + at)\frac{x}{x_0} - f_*(1 - \frac{x}{x_0}) \leq g(t, x) \leq 0. \tag{26}$$

Returning to the function $f(t, x)$, according to relation (23), based on inequalities (25) and (26), we obtain the two-sided estimate for the desired function $f(t, x)$:

$$(at + 2b)\frac{x}{x_0} \leq f(t, x) \leq f_*\left(1 - \frac{x}{x_0}\right) + 2(at + b)\frac{x}{x_0}. \tag{27}$$

Substituting the linear regression coefficients in the form of (18) and (19) into the obtained Inequality (27), we have:

$$\left[\frac{N\sum_{i=1}^{N} t_i F_{ei} - \left(\sum_{i=1}^{N} t_i\right)\left(\sum_{i=1}^{N} F_{ei}\right)}{N\left(\sum_{i=1}^{N} t_i^2\right) - \left(\sum_{i=1}^{N} t_i\right)^2} t + 2\frac{\left(\sum_{i=1}^{N} t_i^2\right)\left(\sum_{i=1}^{N} F_{ei}\right) - \left(\sum_{i=1}^{N} t_i\right)\left(\sum_{i=1}^{N} t_i F_{ei}\right)}{N\left(\sum_{i=1}^{N} t_i^2\right) - \left(\sum_{i=1}^{N} t_i\right)^2} \right] \frac{x}{x_0} \leq$$

$$\leq f(t, x) \leq f_*\left(1 - \frac{x}{x_0}\right) +$$

$$+2\left[\frac{N\sum_{i=1}^{N} t_i F_{ei} - \left(\sum_{i=1}^{N} t_i\right)\left(\sum_{i=1}^{N} F_{ei}\right)}{N\left(\sum_{i=1}^{N} t_i^2\right) - \left(\sum_{i=1}^{N} t_i\right)^2} t + \frac{\left(\sum_{i=1}^{N} t_i^2\right)\left(\sum_{i=1}^{N} F_{ei}\right) - \left(\sum_{i=1}^{N} t_i\right)\left(\sum_{i=1}^{N} t_i F_{ei}\right)}{N\left(\sum_{i=1}^{N} t_i^2\right) - \left(\sum_{i=1}^{N} t_i\right)^2} \right] \frac{x}{x_0}. \tag{28}$$

Considering the case $N \to \infty$, by using (22), we obtain an asymptotic relation:

$$\frac{t\sum_{i=1}^{\infty} t_i F_{ei}}{\sum_{i=1}^{\infty} t_i^2} \frac{x}{x_0} \leq \lim_{N \to \infty} f(t, x) \leq f_*\left(1 - \frac{x}{x_0}\right) + 2\frac{t\sum_{i=1}^{\infty} t_i F_{ei}}{\sum_{i=1}^{\infty} t_i^2} \frac{x}{x_0}. \tag{29}$$

The asymptotic inequality (29) shows that changes in the sample size affect the limits of two-sided estimates for the function $f(t, x)$, restricting inequality (28).

If the experimental values F_{ei} are of the same sign, then as the sample size increases. Both limits of the functional interval for $f(t, x)$ (29) increase for $\forall t \in [0, t_N]$ and $\forall x \in [0, x_0]$; that is, the two-sided estimate of the solution shifts to the increasing values $f(t, x)$.

If we present the coefficients of the linear regression in terms of the covariance and variance, (18) and (19), then the two-sided evaluation of the function $f(t, x)$ takes the form:

$$\left(\frac{\text{cov}(t, F_e(t))}{\sigma_t^2} \left(t - 2\frac{N}{\sum_{i=1}^{N} t_i} \right) + 2\frac{\sum_{i=1}^{N} F_{ei}}{\sum_{i=1}^{N} t_i} \right) \frac{x}{x_0} \leq f(t, x) \leq .$$

$$\leq f_* \left(1 - \frac{x}{x_0} \right) + \left(\frac{2\text{cov}(t, F_e(t))}{\sigma_t^2} \left(t - \frac{N}{\sum\limits_{i=1}^{N} t_i} \right) + \frac{2\sum\limits_{i=1}^{N} F_{ei}}{\sum\limits_{i=1}^{N} t_i} \frac{x}{x_0} \right) \frac{x}{x_0}.$$

Note that the increase in the covariance of the values t and $F_e(t)$ leads to a shift in the two-sided estimate for $f(t, x)$ to the increase, and the interval itself increases. At the same time, the increase in the variance of the values t shifts in the direction of the decrease in the values of $f(t, x)$, and the width of the two-sided estimate narrows.

6. Confidence Intervals for the Predicted Value of the Constructed Linear Regression Model for the Function $F(t)$

To establish a two-tailed critical region to obtain a solution to the initial-boundary value problem in (1)–(3) and (7), we first need to find confidence intervals for the linear regression $F(t)$ (4) using the methodology presented in [34].

Let a given level of significance be $\alpha = 0.05$. Based on the experimental data (Table 1), we compile a matrix of regressors $\mathbf{K}$. For this case, it has the following form:

$$\mathbf{K}^T = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ t_1 & t_2 & \ldots & t_N \end{pmatrix}. \tag{30}$$

The variance of the predicted value of the linear regression $F(t)$ is determined by the following formula:

$$\sigma_F^2 = \mathbf{f}_t^T (\mathbf{K}^T \mathbf{K})^{-1} \mathbf{f}_t \sigma_t^2, \tag{31}$$

where $\mathbf{f}_t^T = \begin{pmatrix} 1 & t \end{pmatrix}$.

Substituting the estimate of the mean square deviation σ_F (31) into the expression $\left| \widehat{F}(t) - F(t) \right| / \sigma_F$, we obtain the sought confidence interval. Therefore, the true value of the desired function at the lower boundary of the layer $F(t)$ varies with probable $1 - \alpha = 0.95$, within the limits [29]

$$\widehat{F}(t) - t_\alpha \sigma_t \sqrt{\mathbf{f}_t^T (\mathbf{K}^T \mathbf{K})^{-1} \mathbf{f}_t} \leq F(t) \leq \widehat{F}(t) + t_\alpha \sigma_t \sqrt{\mathbf{f}_t^T (\mathbf{K}^T \mathbf{K})^{-1} \mathbf{f}_t}, \tag{32}$$

where $t_\alpha = t_\gamma = t(\alpha; \nu)$ is the tabular value of Student's distribution at the significance level α and the number of degrees of freedom $\nu = N - 1$. Calculation of the information matrix is carried out, using the following methodology:

$$(\mathbf{K}^T \mathbf{K}) = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ t_1 & t_2 & \ldots & t_N \end{pmatrix} \cdot \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ . & . \\ 1 & t_N \end{pmatrix} = \begin{pmatrix} 1 + 1 + \ldots + 1 & t_1 + t_2 + \ldots + t_N \\ t_1 + t_2 + \ldots + t_N & t_{1_1}^2 + t_2^2 + \ldots + t_N^2 \end{pmatrix},$$

Then the information matrix has the form

$$(\mathbf{K}^T \mathbf{K}) = \begin{pmatrix} N & \sum\limits_{i=1}^{N} t_i \\ \sum\limits_{i=1}^{N} t_i & \sum\limits_{i=1}^{N} t_i^2 \end{pmatrix},$$

the determinant of which is obtained in the form

$$\det(\mathbf{K}^T \mathbf{K}) = N \sum_{i=1}^{N} t_i^2 - \left(\sum_{i=1}^{N} t_i \right)^2.$$

Then, we obtain the error matrix as follows:

$$
\left(\mathbf{K}^T\mathbf{K}\right)^{-1} = \begin{pmatrix} \dfrac{\sum\limits_{i=1}^{N} t_i^2}{N\sum\limits_{i=1}^{N} t_i^2 - \left(\sum\limits_{i=1}^{N} t_i\right)^2} & \dfrac{-\sum\limits_{i=1}^{N} t_i}{N\sum\limits_{i=1}^{N} t_i^2 - \left(\sum\limits_{i=1}^{N} t_i\right)^2} \\[2em] \dfrac{-\sum\limits_{i=1}^{N} t_i}{N\sum\limits_{i=1}^{N} t_i^2 - \left(\sum\limits_{i=1}^{N} t_i\right)^2} & \dfrac{N}{N\sum\limits_{i=1}^{N} t_i^2 - \left(\sum\limits_{i=1}^{N} t_i\right)^2} \end{pmatrix}.
$$

Since the vectors $\mathbf{f}_t$ are different at different points of the factor space, the length of the confidence intervals will also be different; that is, the accuracy of the predicted value of the linear regression $F(t)$ may differ at different points in time t.

The limits of the two-tailed critical region for the function $F(t)$ are found. Let

$$
F(t) = \widehat{F}(t) \pm \widetilde{F}(t), \tag{33}
$$

where, in the case of linear regression, which follows from Formula (32), we have

$$
\widehat{F}(t) = at + b, \quad \widetilde{F}(t) = A\sqrt{q_1 + q_2 t + q_3 t^2}. \tag{34}
$$

Here,

$$
q_1 = \frac{\sum\limits_{i=1}^{N} t_i^2}{N\sum\limits_{i=1}^{N} t_i^2 - \left(\sum\limits_{i=1}^{N} t_i\right)^2}, \quad q_2 = \frac{-2\sum\limits_{i=1}^{N} t_i}{N\sum\limits_{i=1}^{N} t_i^2 - \left(\sum\limits_{i=1}^{N} t_i\right)^2}, \quad q_3 = \frac{N}{N\sum\limits_{i=1}^{N} t_i^2 - \left(\sum\limits_{i=1}^{N} t_i\right)^2}; \tag{35}
$$

$A = t_\alpha \sigma_F$, where $t_\alpha = t_{\alpha;\nu}$ is the solution to Equation [33]:

$$
2 \int_{t_{\alpha;\nu}/2}^{\infty} \frac{\Gamma\left(\dfrac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\dfrac{\nu}{2}\right)} \left(1 + \frac{t'^2}{\nu}\right)^{-\frac{\nu+1}{2}} dt' = \alpha. \tag{36}
$$

where ν is the number of degrees of freedom, $\nu = N + 1$, and $\Gamma(x)$ is Gamma-function [32]. Note that "+" in Formula (33) refers to the upper limit of the critical region, and "−" refers to the lower limit.

Taking into account that N is a natural number, we present the relation (36) in the following form:

$$
\alpha = \begin{cases} 2\int\limits_{t_{\alpha;\nu}/2}^{\infty} \dfrac{(N-1)!!}{2^{N/2}\sqrt{\pi N}\left(\dfrac{N}{2}-1\right)!} \left(1 + \dfrac{t'^2}{N}\right)^{-\frac{N+1}{2}} dt', & \text{if } N \text{ is even number} \\[2.5em] 2\int\limits_{t_{\alpha;\nu}/2}^{\infty} \dfrac{(N-1)!!}{\sqrt{2\pi N}\,(N-2)!!} \left(1 + \dfrac{t'^2}{N}\right)^{-\frac{N+1}{2}} dt', & \text{if } N \text{ is odd number} \end{cases} \tag{37}
$$

where $(2n)!! = 2^n n!$, $(2n+1)!! = 2^{n+1}\Gamma\left(n+\frac{3}{2}\right)$ [37].

It follows from Formula (37) that, if the number of measurements N increases, then the values of the integrand function decrease, and in order for the area under the integrand function to remain constant (namely, $\alpha/2$), the integration interval must increase; accordingly, the value $t_{\alpha;\nu}$ should decrease.

In our analysis of the influence of the parameter N on the coefficients q_i, $i = 1, 2, 3$ (35), if the number of measurements increased, then these coefficients decreased. Moreover,

$$\lim_{N\to\infty} q_1(N) = \lim_{N\to\infty} q_2(N) = 0, \; \lim_{N\to\infty} q_3 = \frac{1}{\sum\limits_{i=1}^{N} t_i^2}$$

Then, at a constant value of σ_t, an increase in the number of experiments N on the same interval $t \in [0; t_N]$ leads to a narrowing of the confidence interval (32), where the experimental data fall with probability $(1 - \alpha)$. And an increase in the mean square deviation of the variable t, that is, σ_t, leads to an increase in the confidence interval for the linear regression function $F(t)$ for the same number of measurements. We note that the influence of the variance of the variable t on the confidence interval is similar to the influence of σ_t, but the growth by/to the same number slows down the rate of growth of this interval.

7. Establishing the Two-Tailed Critical Region for Solution to the Initial-Boundary Value Problem

To establish the two-tailed critical region for the solution to the problem (1)–(3) and (7), we substitute a representation of the function $F(t)$ (33) into relationship (15). Then, we have:

$$f_F^{\pm}(t,x) = c_* \left(1 - \frac{x}{x_0}\right) + \widehat{F}(t)\frac{x}{x_0} \pm \widetilde{F}(t)\frac{x}{x_0} - \frac{2}{x_0}\sum_{n=1}^{\infty}\left[\frac{(-1)^n}{y_n}\widehat{F}(t) \pm \frac{(-1)^n}{y_n}\widetilde{F}(t) + \right.$$

$$+ e^{-Dy_n^2 t}\left[\frac{c_*}{y_n} - 2\frac{(-1)^n}{y_n}\left(\widehat{F}(t)\Big|_{t=0} \pm \widetilde{F}(t)\Big|_{t=0}\right) + \right.$$

$$(-1)^n Dy_n \int \widehat{F}(t)e^{Dy_n^2 t}dt\Big|_{t=0} \pm (-1)^n Dy_n \int \widetilde{F}(t)e^{Dy_n^2 t}dt\Big|_{t=0} -$$

$$\left.\left.- \frac{(-1)^n}{y_n}Dy_n^2 \int \widehat{F}(t)e^{Dy_n^2 t}dt \mp \frac{(-1)^n}{y_n}Dy_n^2 \int \widetilde{F}(t)e^{Dy_n^2 t}dt\right] \sin(y_n x). \tag{38}$$

Considering Formula (16) in the case of imposing linear regression as a condition on the lower boundary of the layer, Formula (38) can be represented as:

$$f_F^{\pm}(t,x) = f(t,x) \pm \widetilde{F}(t)\frac{x}{x_0} \mp \frac{2}{x_0}\sum_{n=1}^{\infty}\left[\frac{(-1)^n}{y_n}\widetilde{F}(t) + e^{-Dy_n^2 t}\left[2\frac{(-1)^{n+1}}{y_n}\widetilde{F}(t)\Big|_{t=0} + \right.\right.$$

$$\left.\left.+ (-1)^n Dy_n \int \widetilde{F}(t)e^{Dy_n^2 t}dt\Big|_{t=0} - \frac{(-1)^n}{y_n}Dy_n^2 \int \widetilde{F}(t)e^{Dy_n^2 t}dt\right] \sin(y_n x). \tag{39}$$

Here, $\widetilde{F}(t)\Big|_{t=0} = A\sqrt{q_1}$.

Note that the undefined integrals $\int \widetilde{F}(t)e^{-Dy_n^2 t}dt\Big|_{t=0}$ and $\int \widetilde{F}(t)e^{-Dy_n^2 t}dt$ are not taken in elementary functions, so we find their approximate expressions. Given that the function $\widetilde{F}(t)$ is continuously differentiable the required number of times, we expand it into a Taylor series on the interval of given experimental measurements $t \in [0, t_N]$ in the vicinity of some point τ in this interval:

$$\widetilde{F}(t) = A\sqrt{q_1 + q_2 t + q_3 t^2} =$$

$$= A\sqrt{q_1 + q_2\tau + q_3\tau^2} + A\frac{t - \tau}{1!}\frac{(q_2 + 2q_3\tau)}{2\sqrt{q_1 + q_2\tau + q_3\tau^2}} +$$

$$+ A\frac{(t - \tau)^2}{2!}\frac{4q_3 q_1 - q_2^2}{4(q_1 + q_2\tau + q_3\tau^2)^{\frac{3}{2}}} + \ldots + R_n(t). \tag{40}$$

Let $\tau = t - \Delta$. In order for the Taylor series to be convergent, the condition $|t - \tau| < 1$ or $|\Delta| < 1$ must be satisfied [40]. We examined the function $\widetilde{F}(t)$, calculated by Formula (40) depending on the value Δ, limiting ourselves to the first three terms of the Taylor series. Then, the total residual term of the series was calculated according to the following formula:

$$R_3(\xi) = A \frac{(t - \tau)^3}{3!} \frac{3(q_2^2 - 4q_3q_1)}{8} \frac{(q_2 + 2q_3\xi)}{(q_1 + q_2\xi + q_3\xi^2)^{\frac{5}{2}}}, \; \xi \in [\tau, t] \quad (41)$$

In Table 2, the exact values of the function and the values of its expansion into the Taylor series (40) are given, with a limitation of three terms, calculated for $\Delta = 0.001, 0.0015, 0.002, 0.0025, 0.003$ on the interval $t \in [0; 3.2]$, and the corresponding total residual terms R_3, calculated according to Formula (41), where ξ is the middle of the interval $[\tau, t]$ shown in Figure 1. Curves 1–5 correspond to the values $\xi(t) = t - 0.001, t - 0.0015, t - 0.002, t - 0.0025$, and $t - 0.003$.

Table 2. Exact and approximate values of the function $\widetilde{F}(t)$ for different values of Δ.

| t | $\widetilde{F}_{exact}(t)$ | $\widetilde{F}_{Taylor}(t)$ | | | | |
		$\Delta = 0.001$	$\Delta = 0.0015$	$\Delta = 0.002$	$\Delta = 0.0025$	$\Delta = 0.003$
0	0.106747163	0.106747305	0.106747483	0.106747731	0.10674805	0.1067484
0.16	0.102730307	0.102730467	0.102730667	0.102730947	0.102731307	0.1027317
0.32	0.106747439	0.106747582	0.106747761	0.106748011	0.106748334	0.1067487
0.48	0.117980745	0.117980851	0.117980983	0.11798117	0.11798141	0.1179817
0.64	0.134635943	0.134636014	0.134636103	0.134636229	0.134636391	0.1346366
0.8	0.154974703	0.15497475	0.154974809	0.154974891	0.154974997	0.1549751
0.96	0.177736943	0.177736974	0.177737013	0.177737067	0.177737138	0.1777372
1.12	0.202105477	0.202105498	0.202105524	0.202105561	0.202105609	0.2021057
1.28	0.227564864	0.227564878	0.227564897	0.227564923	0.227564956	0.227565
1.44	0.253787019	0.25378703	0.253787043	0.253787062	0.253787086	0.2537871
1.6	0.28055815	0.280558158	0.280558168	0.280558182	0.2805582	0.2805582
1.76	0.307735018	0.307735024	0.307735031	0.307735041	0.307735055	0.3077351
1.92	0.335218954	0.335218958	0.335218964	0.335218972	0.335218983	0.335219
2.08	0.362940206	0.36294021	0.362940215	0.362940221	0.362940229	0.3629402
2.24	0.390848283	0.390848286	0.39084829	0.390848295	0.390848301	0.3908483
2.4	0.418905846	0.418905849	0.418905852	0.418905856	0.418905861	0.4189059
2.56	0.447084753	0.447084755	0.447084758	0.447084761	0.447084765	0.4470848
2.72	0.475363425	0.475363426	0.475363428	0.475363431	0.475363435	0.4753634
2.88	0.503725059	0.50372506	0.503725062	0.503725064	0.503725067	0.5037251
3.04	0.532156392	0.532156393	0.532156394	0.532156396	0.532156399	0.5321564
3.2	0.560646819	0.56064682	0.560646821	0.560646823	0.560646825	0.5606468

Note that varying the parameters within wide limits leads to a similar result: $|R_3| < 10^{-10}$ (Figure 1). Therefore, we accept that taking into account three terms of the series (40) approximates the function $\widetilde{F}(t)$ in the form in (34) sufficiently well.
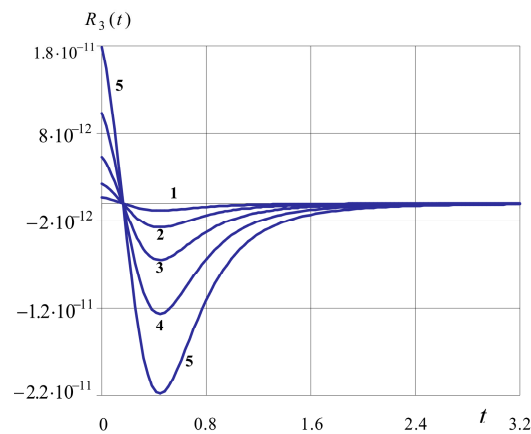
Figure 1. Graphs of total residual terms of the Taylor series (40) for different values of ξ.

If

$$\widetilde{F}(t) \approx A_3 t^2 + \widetilde{b}t + \widetilde{p},$$

where $\widetilde{p} = A_1 - A_2 a + A_3 a^2$, $\widetilde{b} = (A_2 - 2A_3 a)$, then we obtain

$$(-1)^n y_n e^{-Dy_n^2 t} \int \widetilde{F}(t)\, e^{Dy_n^2 t}\, dt = \frac{(-1)^n}{Dy_n} \left[A_3 t^2 + \widetilde{b}t + \widetilde{p} - \frac{(2A_3 t + \widetilde{b})}{Dy_n^2} + \frac{2A_3}{D^2 y_n^4} \right];$$

$$(-1)^n y_n \int \widetilde{F}(t)\, e^{Dy_n^2 t}\, dt \bigg|_{t=0} = \frac{(-1)^n}{Dy_n} \left[\widetilde{p} - \frac{\widetilde{b}}{Dy_n^2} + \frac{2A_3}{D^2 y_n^4} \right].$$

Accordingly, we obtain the following limits of the two-tailed critical region:

$$f_F^{\pm}(t, x) = f(t, x) \pm 2A \sqrt{q_1 + q_2 t + q_3 t^2}\, \frac{x}{x_0} \pm$$

$$\pm \frac{2}{x_0} \sum_{n=1}^{\infty} \frac{(-1)^n}{y_n} e^{-Dy_n^2 t} \left(2A \sqrt{q_1} + A_3 t^2 + \widetilde{b}t - 2\frac{(A_3 t + \widetilde{b})}{Dy_n^2} \right) \sin(y_n x). \qquad (42)$$

Because

$$\lim_{t \to \infty} f_F^{\pm}(t, x) = \lim_{t \to \infty} f(t, x) \pm 2A \sqrt{q_1 + q_2 t + q_3 t^2}\, \frac{x}{x_0},$$

then the functions $f_F^{\pm}(t, x)$ do not have a steady-state regime (time asymptotics), i.e., for long times of the transfer process, the longer the time t, the greater the width of the two-tailed critical region.

The coefficient A has the greatest influence on the width of the two-tailed critical region. Moreover, the smaller the value $|A|$, the narrower the two-tailed critical region for the same reliability coefficient β. And the greater the variance of the predicted value of the linear regression $F(t)$, the wider the two-tailed critical region.

8. The Influence of Statistical Characteristics of the Sample on Reliable Intervals and the Two-Tailed Critical Region of the Solution to the Initial-Boundary Value Problem

Let us investigate the influence of the statistical characteristics of the sample of experimental data on the desired function at the lower boundary of the layer using specific examples. Experimental data were given for both uniform and non-uniform division of the time interval of the study [41]. Consider the cases of large and small samples, which are characterized by large or small variance. Six samples of experimental data and maximum width of the two-tailed critical region of the solution for each Sample are presented in Tables 3–14. Graphs of linear regression, solutions to the initial-boundary value problem, their confidence intervals and two-tailed critical regions for each are shown in Figures 2–13.

Figures 2, 4, 6, 8, 10 and 12 show the linear regression for the corresponding sample (Figure a) and the solution to the initial-boundary value problem in (1)–(3) and (7) (Figure b). In Figure a, the function $F(t)$ is marked with a solid line, its confidence intervals $F^{\pm}(t)$ are marked with dashed–dotted lines, and the experimental data presented in the corresponding table are marked with green dots. For the large time intervals in Figures 2b, 4b, 6b, and 10b, the solutions of the original initial-boundary value problem $f(t, x)$ (solid lines) at times $t = 0.1, 0.5$, and 2 (curves 1–3) are shown, and for small time intervals, in Figures 8b and 12b, the solutions are shown at time points $t = 0.1, 0.5$, and 1 (curves 1–3). Curves (dashed lines) with the index "+" are calculated for the upper limits of confidence intervals $f_c^+(t, x)$, and those with the index "−" are calculated for the lower ones $f_c^-(t, x)$.

In Figures 3, 5, 7, 9, 11 and 13, the solutions to the initial-boundary value problem $f(t, x)$ are shown, normalized to the value of the function at the upper boundary of the layer $x = 0$ and the corresponding two-tailed critical regions. For large time intervals $[0, t_N]$ (Figures 3, 5, 7 and 11), graphs were calculated for small ($t = 0.1, 0.5$ (curves 1 and 2 in Figure a)) and large ($t = 1, 2$ (curves 3 and 4 in Figure b)) moments of time. For small time intervals $[0, t_N]$ (Figures 9 and 13), graphs were calculated for small ($t = 0.1, 0.3$ (curves 1 and 2 in Figure a)) and large ($t = 0.5, 1$ (curves 3 and 4 in Figure b)) moments of time.

In Tables 4, 6, 8, 10, 12 and 14, the maximum widths of the two-tailed critical regions of the solution to the initial-boundary value problem in (1)–(3) and (7) $\Delta_c(t, x) = \frac{1}{f_*}|f_c^+(t, x) - f_c^-(t, x)|$ are given for different moments of time t for six samples.

I. Large sample, large time interval, large variance

Let the sample of experimental data $F(t)$ with size $N = 36$ have the form presented in Table 3.

Table 3. Experimental data with large variance for large time interval.

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
t_i	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.1
$F(t_i)$	0	0.121	0.2527	0.4609	0.6212	0.7142	0.9401	0.9771	1.0312	1.0922	1.1052	1.1197
	t_{13}	t_{14}	t_{15}	t_{16}	t_{17}	t_{18}	t_{19}	t_{20}	t_{21}	t_{22}	t_{23}	t_{24}
t_i	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2	2.1	2.2	2.3
$F(t_i)$	1.1254	1.1741	1.2005	1.2129	1.2539	1.4017	1.4387	1.6109	1.7208	1.7562	1.7707	2.1208
	t_{25}	t_{26}	t_{27}	t_{28}	t_{29}	t_{30}	t_{31}	t_{32}	t_{33}	t_{34}	t_{35}	t_{36}
t_i	2.4	2.5	2.6	2.7	2.8	2.9	3	3.1	3.2	3.3	3.4	3.5
$F(t_i)$	2.2022	2.4131	2.4799	2.7115	2.7307	2.7471	2.7517	2.7641	2.8615	2.9029	2.9413	2.9716

The sample range is $R = [0, 2.9716]$; the sample variance is $S_F^2 = 0.893931523259907$.

Let us construct a regression according to the data in Table 3 using the method of least squares. Based on the type of correlation field, we make an assumption about the linear nature of the time dependence of the sought function on the lower boundary of the body $f(t)|_{x_0}$. According to sample data, the coefficients of linear regression (4) are $a = 0.835387001287001$ and $b = 0.16861996996997$.

Now, we find confidence intervals with reliability values of $\beta = 0.95$ for the coefficients a and b [32]: $a \in (0.800735; 0.922897)$ and $b \in (0.036363; 0.256763)$.
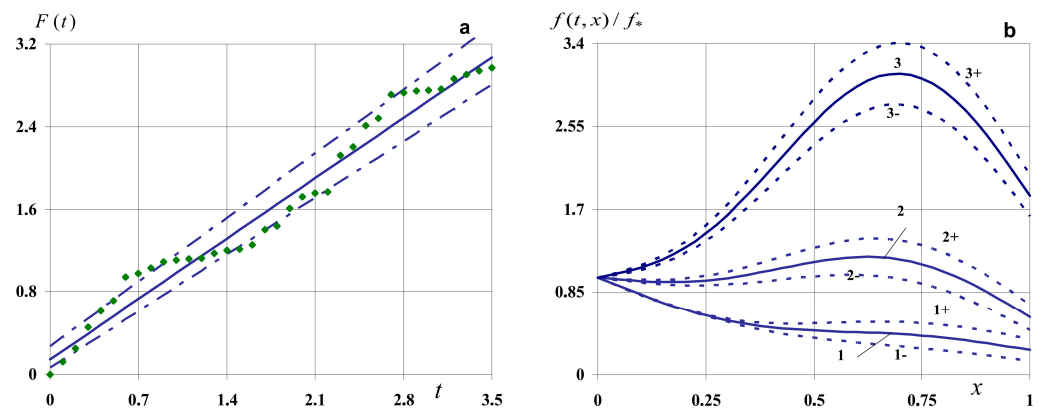
Figure 2. Linear regression $F(t)$ (**a**) and solutions to the initial-boundary value problem $f(t,x)$ in different moments (**b**) and their confidence intervals.

The limits of the two-tailed critical region $f_F^{\pm}(t,x)$, calculated by Formula (42) for this sample, are specified as follows:

The lower limit is

$$0.835387001\,t + 0.16861997 - 1.81478\sqrt{0.1066066 - 0.0900901\,t + 0.0257400\,t^2} \le F(t);$$

the upper limit is

$$F(t) \le 0.835387001\,t + 0.16861997 + 1.81478\sqrt{0.1066066 - 0.0900901\,t + 0.0257400\,t^2};$$
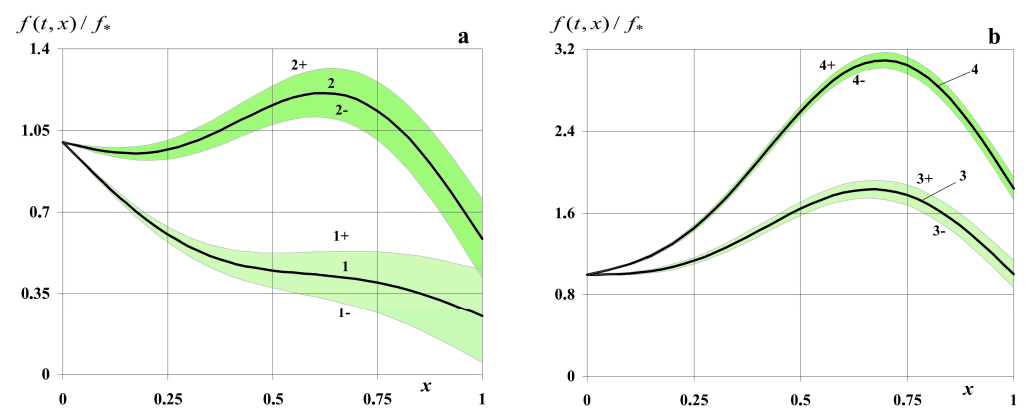
Figure 3. Solutions to the initial-boundary value problem $f(t,x)$ and corresponding two-tailed critical regions for small (**a**) and large moments of time (**b**).

Table 4. Maximum width of the two-tailed critical region of the solution for Sample I.

t	0.1	0.8	1.4	1.7	2.2	2.7	3.4
$\displaystyle\max_{x \in [0,x_0]} \Delta_F(t,x)$	0.0364039	0.29142238	0.22693402	0.2153051	0.23436629	0.29142238	0.40364039

Note that, for short periods of the transfer process, which are described by the initial-boundary value problem (1)–(3) and (7), the function $f(t,x)$ monotonically decreases (curve 1, Figure 2b). As t increases, the function values increase, and its maximum is formed at the point $x_{\max} = 0.625$ (curve 2, Figure 2b). Further, the behavior of the function $f(t,x)$ does not change, but its values increase significantly in the entire body region (curve 3, Figure 2b).

As the duration of the transfer process increases, not only do the values of $f(t,x)$ increase, but the width of the confidence interval $\Delta_F(t,x)$ increases as well (Figure 2b). Thus, with an increase in t from 0.5 to 2, the maximum width of the confidence interval increases approximately by one and a half times: $\left.\max\limits_{x\in[0,x_0]}\Delta_c(t,x)\right|_{t=2} / \left.\max\limits_{x\in[0,x_0]}\Delta_c(t,x)\right|_{t=0.5} = 1.598$.

As opposed to confidence intervals, the two-tailed critical region of the solution to the problem first decreases (Figure 3), reaching its narrowest values in the middle of the time interval $[0, t_N]$. Then, the two-tailed critical region expands symmetrically to the narrowing in the initial time interval (Table 4).

Thus, for short durations, the width of this region, $\Delta_F(t,x) = \frac{1}{f_*}\left|f_F^+(t,x) - f_F^-(t,x)\right|$, decreases to 17%: $\left.\max\limits_{x\in[0,x_0]}\Delta_F(t,x)\right|_{t=0.5} / \left.\max\limits_{x\in[0,x_0]}\Delta_F(t,x)\right|_{t=0.1} = 0.833$ (Figure 3a). For medium durations, the difference between $\left.\max\limits_{x\in[0,x_0]}\Delta_F(t,x)\right|_{t=1}$ and $\left.\max\limits_{x\in[0,x_0]}\Delta_F(t,x)\right|_{t=2}$ decreases by 16.6% (Figure 3b).

II.   Large sample, large time interval, small variance

Consider a sample of experimental data $F(t)$ with a size $N = 36$, which is presented in Table 5.

Table 5. Experimental data with small variance for large time interval.

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
t_i	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.1
$F(t_i)$	0	0.153	0.188	0.207	0.251	0.295	0.306	0.374	0.378	0.42	0.432	0.451
	t_{13}	t_{14}	t_{15}	t_{16}	t_{17}	t_{18}	t_{19}	t_{20}	t_{21}	t_{22}	t_{23}	t_{24}
t_i	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2	2.1	2.2	2.3
$F(t_i)$	0.464	0.465	0.502	0.508	0.512	0.519	0.561	0.562	0.565	0.568	0.58	0.586
	t_{25}	t_{26}	t_{27}	t_{28}	t_{29}	t_{30}	t_{31}	t_{32}	t_{33}	t_{34}	t_{35}	t_{36}
t_i	2.4	2.5	2.6	2.7	2.8	2.9	3	3.1	3.2	3.3	3.4	3.5
$F(t_i)$	0.589	0.609	0.628	0.654	0.656	0.659	0.662	0.668	0.672	0.674	0.677	0.694

The sample range is $R = [0.\ 0.694]$; the sample variance is $S_F^2 = 0.174264028380218$.

The confidence intervals with reliability $\beta = 0.95$ were calculated for the coefficients $a = 0.155924066924067$ and $b = 0.218493993993994$ in the following confidence limits: $a \in (0.136688; 0.17516)$ and $b \in (0.179346560289016; 0.257641427698971)$.
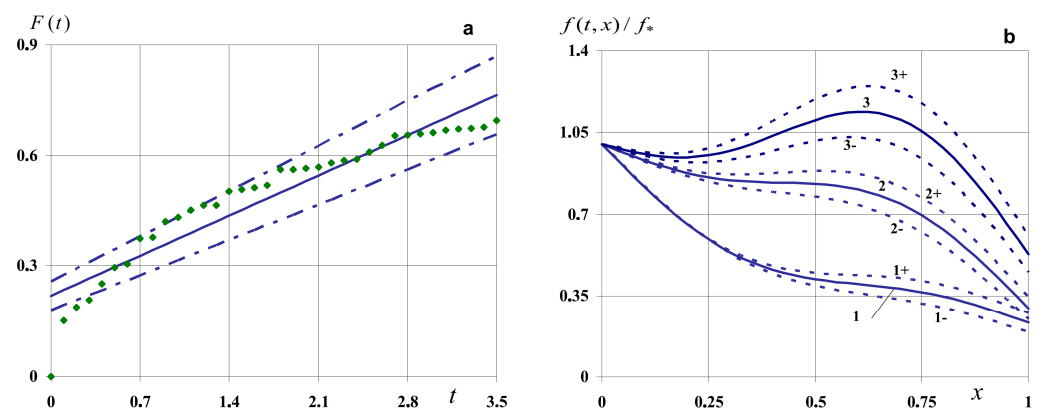

Figure 4. Linear regression $F(t)$ (**a**) and solutions to the initial-boundary value problem $f(t,x)$ at different moments (**b**), as well as their confidence intervals.

The limits of two-tailed critical region $f_F^{\pm}(t, x)$ for this sample are as follows. The lower limit is

$$0.155924\,t + 0.218494 - 0.353774783\sqrt{0.1066066 - 0.0900901\,t + 0.0257400\,t^2} \leq F(t);$$

the upper limit is

$$F(t) \leq 0.155924\,t + 0.218494 + 0.353774783\sqrt{0.1066066 - 0.0900901\,t + 0.0257400\,t^2}.$$
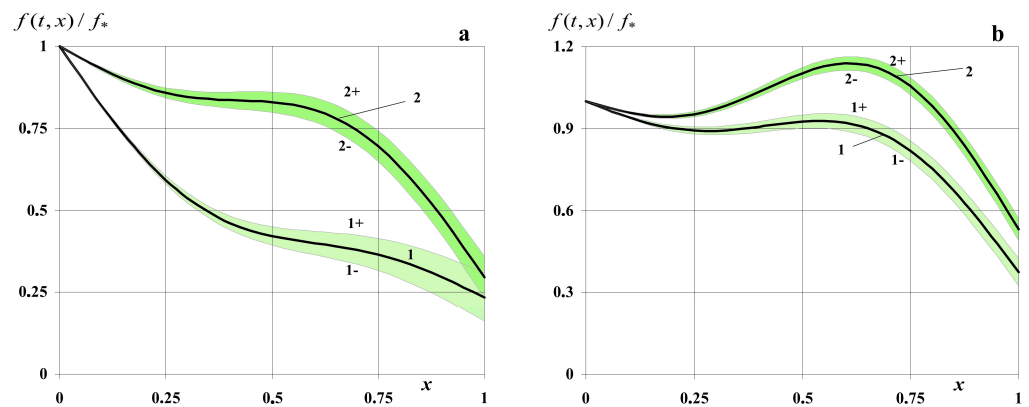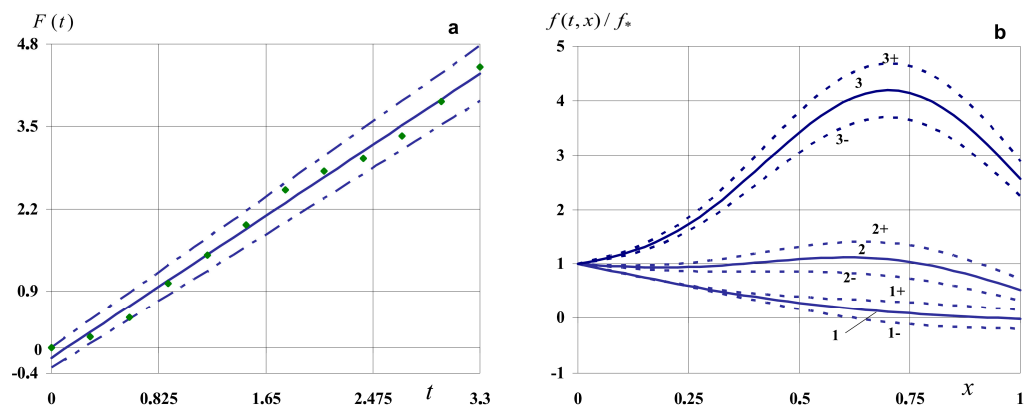

Figure 5. Solutions to the initial-boundary value problem $f(t, x)$ and corresponding two-tailed critical regions for small (**a**) and large moments of time (**b**).

Table 6. Maximum width of the two-tailed critical region of the solution for Sample II.

t	0.1	0.8	1.4	1.7	2.2	2.7	3.4
$\max\limits_{x \in [0, x_0]} \Delta_F(t, x)$	0.15002469	0.1083156	0.08434663	0.0800244	0.08710905	0.108315603	0.15002469

Note that, as the time of the transfer process increases, the values of the function $f(t, x)$ increase in the entire region of the body (Figure 4b). For short and medium durations of time, $f(t, x)$ is a monotonically decreasing function. At the moment of time $t = 0.7$, the local maximum of the function begins to form at the point $x_{\max} = 0.5$. As t increases, this maximum grows, shifts to the lower boundary of the layer ($x_{\max}|_{t=2} = 0.6$), and becomes global (Figure 4b).

As time progresses, the width of the confidence interval $\Delta_c(t, x)$ increases (Figure 4b). For example, $\max\limits_{x \in [0, x_0]} \Delta_c(t, x)\Big|_{t=2} \Big/ \max\limits_{x \in [0, x_0]} \Delta_c(t, x)\Big|_{t=0.5} = 1.598$.

For this sample, the two-tailed critical region of the problem solution also narrows at first (Figure 5), and then, in the second half of the time interval $[0, t_N]$, it expands (Table 6). Thus, the width of this region $\Delta_F(t, x)$ for short durations decreases to 17%, namely, $\max\limits_{x \in [0, x_0]} \Delta_F(t, x)\Big|_{t=0.5} \Big/ \max\limits_{x \in [0, x_0]} \Delta_F(t, x)\Big|_{t=0.1} = 0.833$ (Figure 5a). For long durations, the difference between $\max\limits_{x \in [0, x_0]} \Delta_F(t, x)\Big|_{t=1}$ and $\max\limits_{x \in [0, x_0]} \Delta_F(t, x)\Big|_{t=2}$ decreases by 16.6% (Figure 5b). There is also symmetry with respect to the widths $\Delta_F(t, x)$ at the time intervals $[0, t_N/2]$ and $[t_N/2, t_N]$.

For Samples I and II, there is growth of the confidence interval $\Delta_c(t, x)$ and a decrease in the two-tailed critical region $\Delta_F(t, x)$. The corresponding growth is the same.

III.　Small sample, large time interval, large variance

Now, the sample of experimental data $F(t)$ with the size $N = 12$ takes the form given in Table 7.

Table 7. Experimental data with large variance for large time interval.

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
t_i	0	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3	3.3
$F(t_i)$	0	0.1697	0.4738	1.0209	1.4735	1.9513	2.4972	2.7988	2.9986	3.3457	3.8914	4.4371

The sample range is $R = [0, 4.4371]$; the sample variance is $S_F^2 = 1.47815809975753$.

The confidence intervals with reliability $\beta = 0.95$ are calculated for the coefficients $a = 1.361503496503$ and $b = -0.158314102564102$ in the following confidence limits: $a \in (1.27876771583259; 1.4442392771744)$ and $b \in (-0.319489689069116; 0.00286148394091157)$.

Figure 6. Linear regression $F(t)$ (**a**) and solutions to the initial-boundary value problem $f(t, x)$ at different moments and their confidence intervals (**b**).

The limits of two-tailed critical region $f_F^{\pm}(t, x)$ for this sample are as follows. The lower limit is

$$1.3615035\, t - 0.1583141 - 3.253404\sqrt{0.2948718 - 0.2564103\, t + 0.0777001\, t^2} \leq F(t);$$

the upper limit is

$$F(t) \leq 1.3615035\, t - 0.1583141 + 3.253404\sqrt{0.2948718 - 0.2564103\, t + 0.0777001\, t^2}.$$

Table 8. Maximum width of the two-tailed critical region of the solution for Sample III.

t	0.1	0.8	1.4	1.7	2.2	2.5	3.2
$\max\limits_{x \in [0,x_0]} \Delta_F(t,x)$	0.61692272	0.44338969	0.35257482	0.34312913	0.38806521	0.44338969	0.61692272

For the experimental data on the values of the desired function at the lower boundary of the layer, which are presented in Table 7, the function $f(t, x)$ also monotonically decreased with short durations of the transfer process (curve 1, Figure 6b). As t increases, the value of the function grows. A maximum was formed at the point $x_{\max} = 0.6$ (curve 2, Figure 6b), which grew over time and shifted to the lower boundary of the layer. In particular, for $t = 2$, the maximum value of the function reached $x_{\max} = 0.7$. For long durations, the values of the function $f(t, x)$ increased significantly in the entire region of the body (curve 3, Figure 6b).
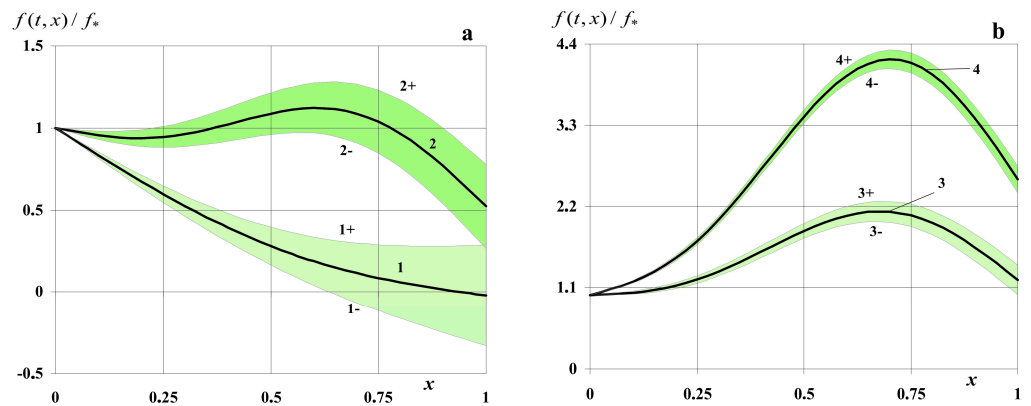
Figure 7. Solutions to the initial-boundary value problem $f(t,x)$ and corresponding two-tailed critical regions for small (**a**) and large moments of time (**b**).

As time increases, the width of the confidence interval $\Delta_c(t,x)$ increases (Figure 6b). For example, $\max\limits_{x\in[0,x_0]}\Delta_c(t,x)\Big|_{t=2} / \max\limits_{x\in[0,x_0]}\Delta_c(t,x)\Big|_{t=0.5} = 1.619$.

For this sample, the two-tailed critical region of the problem solution also decreases over time (Figure 7), and then expands (Table 8). Thus, the width of this region $\Delta_F(t,x)$, for small short durations, decreases to 17%: $\max\limits_{x\in[0,x_0]}\Delta_F(t,x)\Big|_{t=0.5} / \max\limits_{x\in[0,x_0]}\Delta_F(t,x)\Big|_{t=0.1} = 0.83$

(Figure 7a). For long durations, the difference between $\max\limits_{x\in[0,x_0]}\Delta_F(t,x)\Big|_{t=1}$ and $\max\limits_{x\in[0,x_0]}\Delta_F(t,x)\Big|_{t=2}$ decreases by 10.6% (Figure 7b).

Symmetry with respect to the widths of $\Delta_F(t,x)$ at time intervals $[0, t_N/2]$ and $[t_N/2, t_N]$ is also present.

IV.   Small sample, small time interval, large variance

Now, consider the sample of experimental data $F(t)$ with the size $N = 12$ and a small time interval, as presented in Table 9.

Table 9. Experimental data with large variance for large time interval.

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
t_i	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1
$F(t_i)$	0	0.0162	0.5809	1.1439	2.2175	3.2737	3.6345	4.1847	4.8021	5.1542	5.5129	5.9291

The sample range is $R = [0, 5.9291]$; the sample variance is $S_F^2 = 2.18380650788105$.

The confidence intervals with reliability $\beta = 0.95$ are calculated for the coefficients $a = 5.9880034965035$ and $b = -0.255926923076922$ in such confidence limits $a \in (5.34665817078342; 6.62934882222358)$ and $b \in (-0.672390063074964; 0.16053621692112)$.

The limits of two-tailed critical region $f_F^\pm(t,x)$ for this sample are as follows:

The lower limit is

$$5.988003\,t - -0.25593 - 3.253404\,\sqrt{0.2948718 - 0.2564103\,t + 0.0777001\,t^2} \leq F(t);$$

the upper limit is

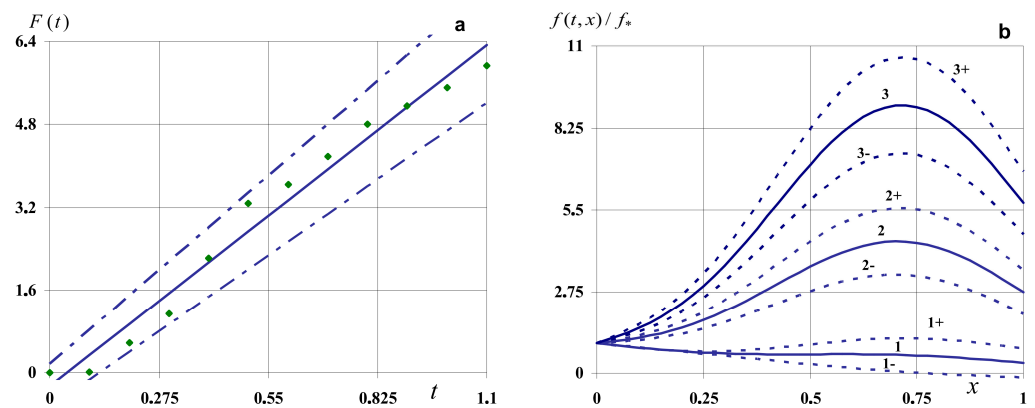$$F(t) \leq 5.988003\,t - 0.25593 + 3.253404\,\sqrt{0.2948718 - 0.2564103\,t + 0.0777001\,t^2}.$$

Figure 8. Linear regression $F(t)$ (**a**) and solutions to the initial-boundary value problem $f(t, x)$ at different moments and their confidence intervals (**b**).
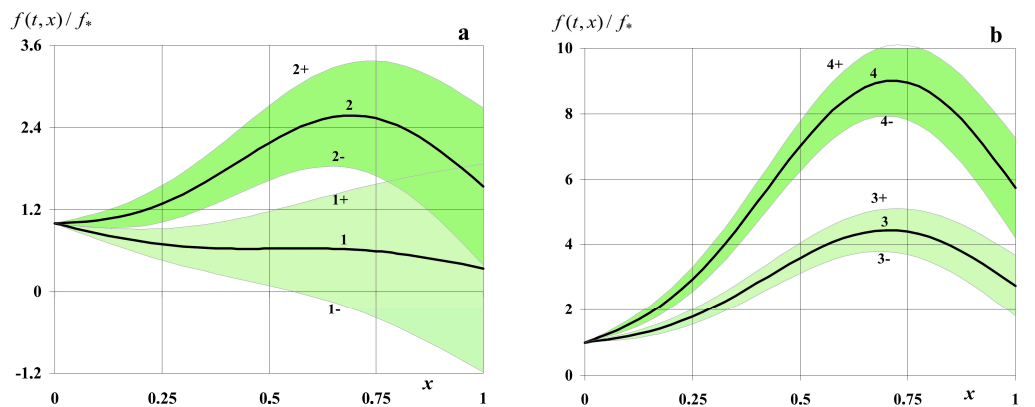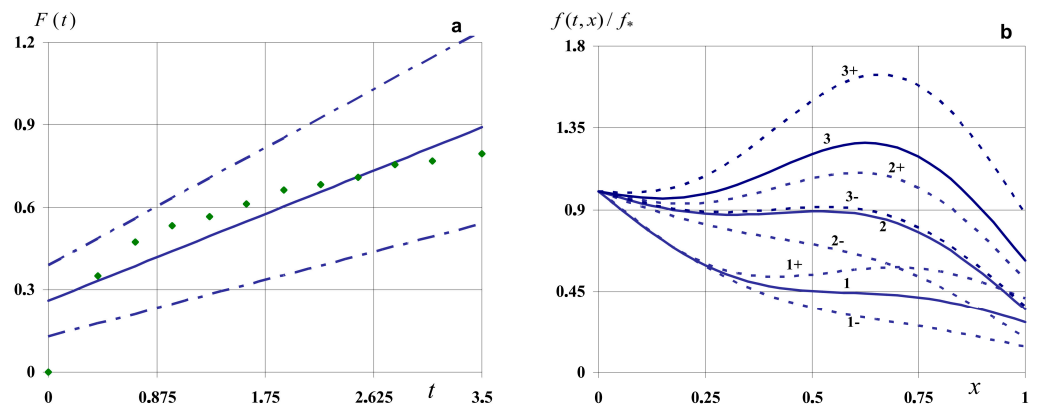

Figure 9. Solutions to the initial-boundary value problem $f(t, x)$ and corresponding two-tailed critical regions for small (**a**) and large moments of time (**b**).

Table 10. Maximum width of the two-tailed critical region of the solution for Sample IV.

t	0.1	0.3	0.5	0.6	0.7	0.8	1
$\max\limits_{x \in [0, x_0]} \Delta_F(t, x)$	3.04895999	2.29132303	1.87514347	1.87514347	2.023403671	2.29132303	3.04895999

For the experimental data on the values of the desired function at the lower boundary of the layer, which are presented in Table 9, formation of the local maximum of the function $f(t, x)$ begins for low durations. The $\max\limits_{x \in [0, x_0]} f(t, x)$ forms in the middle of the body increase with time and shift to the lower boundary of the layer. An example is $x_{\max}|_{t=0.1} = 0.575$ and $\max\limits_{x \in [0, x_0]} f(t, x) = 0.635718339$ (curve 1, Figure 8b), $x_{\max}|_{t=0.5} = 0.7$ and $\max\limits_{x \in [0, x_0]} f(t, x) = 4.4428612$ (curve 2, Figure 8b), in addition to $x_{\max}|_{t=1} = 0.725$ and $\max\limits_{x \in [0, x_0]} f(t, x) = 9.007544363$ (curve 3, Figure 8b).

The width of the confidence interval $\Delta_c(t, x)$ increases as the duration of the process increases (Figure 8b). For example, $\max\limits_{x \in [0, x_0]} \Delta_c(t, x)\Big|_{t=1} \Big/ \max\limits_{x \in [0, x_0]} \Delta_c(t, x)\Big|_{t=0.5} = 1.44$.

For this sample, the two-tailed critical region of the solution to the problem also narrows over time (Figure 9), and then increases symmetrically at the intervals $[0, t_N/2]$ and $[t_N/2, t_N]$ (Table 10). Thus, the width of this region $\Delta_F(t, x)$ for short durations decreases to 25%: $\max\limits_{x \in [0, x_0]} \Delta_F(t, x)\Big|_{t=0.3} \Big/ \max\limits_{x \in [0, x_0]} \Delta_F(t, x)\Big|_{t=0.1} = 0.751$ (Figure 9a). For medium dura-

tions, the difference between $\max\limits_{x\in[0,x_0]}\Delta_F(t,x)\bigg|_{t=0,5}$ and $\max\limits_{x\in[0,x_0]}\Delta_F(t,x)\bigg|_{t=1}$ increases by 62.6% (Figure 9b).

V.   Small sample, large time interval, small variance

The sample of experimental data $F(t)$ with the size $N = 12$ was obtained at large time interval, as shown in Table 11.

Table 11. Experimental data with small variance for large time interval.

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
t_i	0	0.4	0.7	1	1.3	1.6	1.9	2.2	2.5	2.8	3.1	3.5
$F(t_i)$	0	0.351	0.474	0.532	0.565	0.612	0.662	0.682	0.709	0.756	0.768	0.794

The sample range is $R = [0.\ 0.794]$; the sample variance is $S_F^2 = 0.223456914022016$.

The confidence intervals with reliability $\beta = 0.95$ were calculated for the coefficients $a = 0.180180811808118$ and $b = 0.26010024600246$, with confidence limits $a \in (0.116880945418082; 0.243480678198154)$ and $b \in (0.130502883575815; 0.389697608429105)$.

Figure 10. Linear regression $F(t)$ (**a**) and solutions to the initial-boundary value problem $f(t,x)$ at different moments, along with their confidence intervals (**b**).

The limits of the two-tailed critical region $f_F^{\pm}(t,x)$ for this sample were as follows: The lower limit is

$$0.18018081\,t + 0.26010025 - 0.49182535\sqrt{0.3093481 - 0.2583026\,t + 0.0738007\,t^2} \le F(t);$$

the upper limit is

$$F(t) \le 0.18018081\,t + 0.26010025 + 0.49182535\sqrt{0.3093481 - 0.2583026\,t + 0.0738007\,t^2}.$$

Table 12. Maximum width of the two-tailed critical region of the solution for Sample IV.

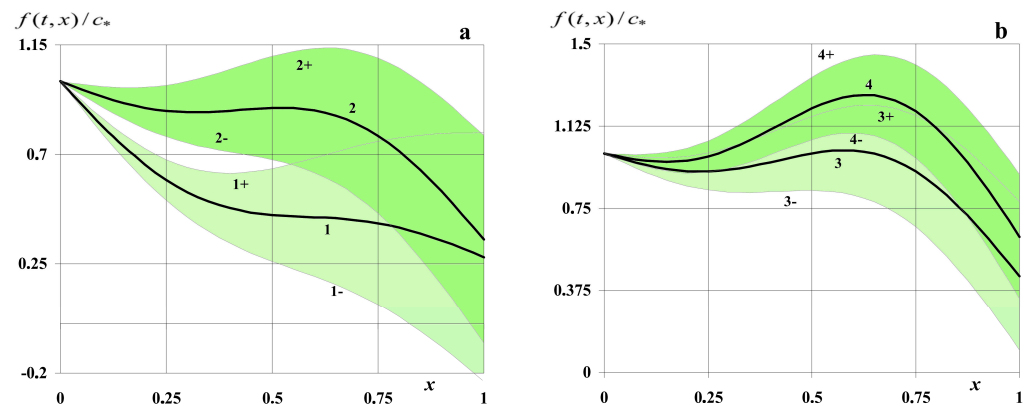| t	0.2	0.8	1.1	1.7	2.3	2.7	3.3
$\max\limits_{x\in[0,x_0]}\Delta_F(t,x)$	0.97894091	0.74249383	0.64888217	0.55414815	0.6232863	0.74249383	0.97894091

Figure 11. Solutions to the initial-boundary value problem $f(t,x)$ and corresponding two-tailed critical regions for small (**a**) and large moments of time (**b**).

For the experimental data on the values of the desired function at the lower boundary of the layer, which are presented in Table 11, the function $f(t,x)$ monotonically decreased for short durations of the transfer process (curve 1, Figure 10b). The local maximum formed at the point $x_{\max} = 0.525$ (curve 2, Figure 10b), and, as in previous cases, grew over time and shifted to the lower boundary of the layer (curve 3, Figure 10b) — for example, $x_{\max}|_{t=1} = 0.625$ and $\max\limits_{x \in [0,x_0]} f(t,x) = 1.267372491$ (curve 3, Figure 10b).

As the time increased, the width of the confidence interval $\Delta_c(t,x)$ increased (Figure 10b). For example, $\max\limits_{x \in [0,x_0]} \Delta_c(t,x)\Big|_{t=2} \Big/ \max\limits_{x \in [0,x_0]} \Delta_c(t,x)\Big|_{t=0.5} = 1.59$.

For this sample, the two-tailed critical region of the problem solution also narrowed over time (Figure 11), and then increased symmetrically at the intervals $[0, t_N/2]$ and $[t_N/2, t_N]$ (Table 12). Thus, the width of this region $\Delta_F(t,x)$ for short durations decreased to 17%: $\max\limits_{x \in [0,x_0]} \Delta_F(t,x)\Big|_{t=0.5} \Big/ \max\limits_{x \in [0,x_0]} \Delta_F(t,x)\Big|_{t=0.1} = 0.836$ (Figure 11a). For medium durations, the difference between $\max\limits_{x \in [0,x_0]} \Delta_F(t,x)\Big|_{t=1}$ and $\max\limits_{x \in [0,x_0]} \Delta_F(t,x)\Big|_{t=2}$ increased by 16.07% (Figure 11b).

VI.   Small sample, small time interval, small variance

The sample of experimental data $F(t)$ with the size $N = 12$ was obtained at the large time interval given in Table 13.

Table 13. Experimental data with small variance for small time interval.

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
t_i	0	0.105	0.2	0.331	0.427	0.55	0.632	0.8	0.868	0.951	1.01	1.1
$F(t_i)$	0	0.0186	0.0206	0.0241	0.0315	0.0333	0.0351	0.0373	0.0408	0.0429	0.0443	0.0461

The sample range is $R = [0, 0.0461]$; the sample variance is $S_F^2 = 0.11354814757706$.

The confidence intervals with reliability $\beta = 0.95$ were calculated for the coefficients $a = 0.034084888694602$ and $b = 0.0114076655203205$, with confidence limits $a \in (0.0258221071878995; 0.0423476702013045)$ and $b \in (0.00577685424483104; 0.0170384767958099)$.
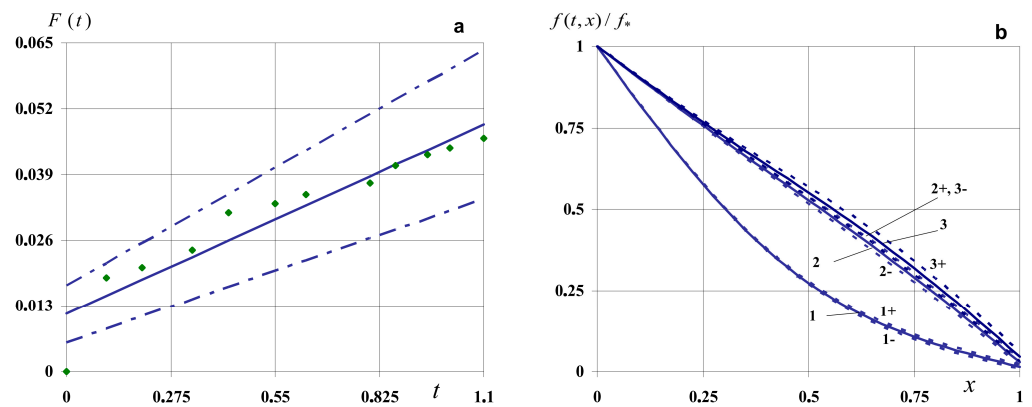
Figure 12. Linear regression $F(t)$ (**a**) and solutions to the initial-boundary value problem $f(t,x)$ at different moments and their confidence intervals (**b**).

The limits of the two-tailed critical region $f_F^{\pm}(t,x)$ for this sample are as follows: The lower limit is

$$0.034084888694602\,t + 0.0114076655203205 - 0.0620065727181975\times$$

$$\times\sqrt{0.306259616566363 - 0.769325309060361\,t + 0.663735977332974\,t^2} \leq F(t);$$

the upper limit is

$$F(t) \leq 0.034084888694602\,t + 0.0114076655203205 + 0.0620065727181975\times$$

$$\times\sqrt{0.306259616566363 - 0.769325309060361\,t + 0.663735977332974\,t^2}.$$
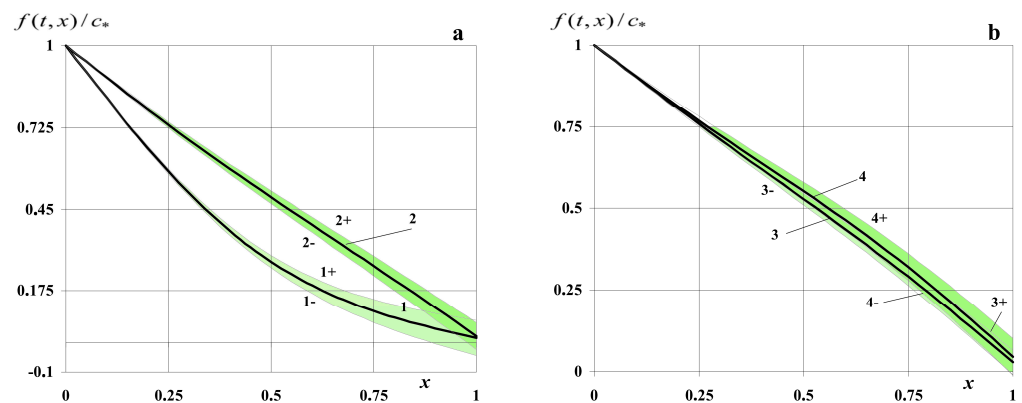
Figure 13. Solutions to the initial-boundary value problem $f(t,x)$ and corresponding two-tailed critical regions for small (**a**) and large moments of time (**b**).

Table 14. Maximum width of the two-tailed critical region of the solution for Sample IV.

t	0.1	0.3	0.5	0.6	0.7	0.8	1
$\max\limits_{x\in[0,x_0]} \Delta_F(t,x)$	0.12050392	0.09112286	0.07332212	0.07173896	0.07575928	0.08458784	0.11159884

Regarding the experimental data on the values of the desired function at the lower boundary of the layer, presented in Table 13, the function $f(t,x)$ monotonically decreased over the entire time interval $[0, t_N]$ (Figure 12b). Moreover, time $t = 1$ was already in a steady state: $t = t_{st}$. As the time approached t_{st}, the rate of growth of the function $f(t,x)$ slowed down significantly (Figure 12b).

The width of the confidence interval $\Delta_c(t,x)$ for this sample was significantly smaller than for the previously considered ones. As the duration of the process increased, the width of the confidence interval increased slightly (Figure 12b). For example, $\left.\max\limits_{x\in[0,x_0]}\Delta_c(t,x)\right|_{t=1} \Big/ \left.\max\limits_{x\in[0,x_0]}\Delta_c(t,x)\right|_{t=0.5} = 1.43$.

For this sample, the two-tailed critical region of the problem solution was significantly narrower than in the case of other samples. But, also for sample *VI*, the width of the two-tailed critical region also narrowed over time (Figure 13) and then increased symmetrically at intervals $[0, t_N/2]$ and $[t_N/2, t_N]$ (Table 14). Thus, the width of this region $\Delta_F(t,x)$ decreased to 24% for short durations, namely, $\left.\max\limits_{x\in[0,x_0]}\Delta_F(t,x)\right|_{t=0.3} \Big/ \left.\max\limits_{x\in[0,x_0]}\Delta_F(t,x)\right|_{t=0.1} = 0.756$

(Figure 13a). For medium durations, the difference between $\left.\max\limits_{x\in[0,x_0]}\Delta_F(t,x)\right|_{t=0,5}$ and

$\left.\max\limits_{x\in[0,x_0]}\Delta_F(t,x)\right|_{t=1}$ decreased by 34.3% (Figure 13b).

Note that, for the samples considered for all moments of time $t\in[0,t_N]$, the confidence interval and the two-tailed critical region for the function $f(t,x)$ were symmetric, i.e., $|f(t,x)-f_c^-(t,x)| = |f_c^+(t,x)-f(t,x)|$ and $|f(t,x)-f_F^-(t,x)| = |f_F^+(t,x)-f(t,x)|$ for $\forall t\in[0,t_N], \forall x\in[0,x_0]$.

We also note that, for all six samples, the point $x = x_{\max}^c$ at which the width of the confidence interval for the function $f(t,x)$ was the largest was the same, but it may have differed at different moments in time. Thus, $x_{\max}^c = 0.8$ for $t = 0.1$, $x_{\max}^c = 0.725$ for $t = 0.5$, and $t = 2$ (Figures 2b, 4b, 6b, 8b, 10b and 13b). At the same time, the largest width of the two-tailed critical region was observed at the lower boundary of the layer, i.e., $x_{\max}^t = 1$, and remained unchanged throughout the entire duration $t \in [0, t_N]$ of the process under investigation (Figures 3, 5, 7, 9, 11 and 13). The point $x = x_{\min}$ at which the widths of the confidence interval and the two-tailed critical region were equal to zero $\Delta = 0$ was at the upper boundary of the layer, since the value of the sought function at this boundary was known and constant over time.

9. Numerical Analysis of the Solution Depending on Statistical Characteristics of the Sample

We visualized and analyzed the solution to the initial-boundary value problem in (1)–(3) and (7) depending on the coefficients of the problem and the statistical parameters. Calculations for the function $f(t,x)$ were carried out using Formula (17), with basic values of the problem parameters: $t = 1$, $f_* = 1$, $D = 1$, and $N = 36$. Also, Sample 2, with a large size, small variance, and a long interval, was chosen as the basis sample.

In Figure 14, graphs of the function $f(t,x)/f_*$ are shown at different time points for $D = 0.05$ (Figure a) and $D = 1$ (Figure b). Here, curves 1–5 correspond to time points $t = 0.1, 0.5, 1, 2,$ and 3.

The influence of the coefficient D on the solution to the initial-boundary value problem (1)–(3) and (7) is shown in Figure 15, where curves 1–5 correspond to the values $D = 0.02$, 0.05, 0.1, 0.5, and 1. Here and in the future, the curves in Figure a are calculated at the moment $t = 0.1$, while the curves in Figure b are calculated at the moment $t = 1$. Figure 16 illustrates the behavior of the problem solution depending on the value of the desired function at the boundary of the layer $x = 0$. The values $f_* = 0.2, 0.5, 1, 2, 3$ correspond to curves 1–5.

Figure 17 demonstrate graphs of the function $f(t,x)/f_*$ for different values of the coefficient of linear regression a, $a > 0$. In Figure a ($t = 0.1$), the values $a = 0, 0.16, 1, 2,$ and 3, and the values $a = 0, 0.1, 0.16, 0.2, 0,5$ in Figure b ($t = 1$) correspond to curves 1–5, where $b = 0.22$. Figure 18 illustrates the behavior of the problem solution in (1)–(3) and (7) for negative values of the coefficient a. Here, the values $a = -4, -3, -2, -1, -0.5$ correspond to curves 1–5, $b = 0.22$. In Figures 19 and 20, the dependence of the solution on positive (Figure 19) and negative (Figure 20) values of the linear regression coefficient b is shown.

Curves 1–5 correspond to the values $b = 0, 0.22, 0.5, 1$, and 1.5 in Figure 19 and $b = -3, -2, -1, -0.5$, and -0.1 in Figure 20; $a = 0.16$.
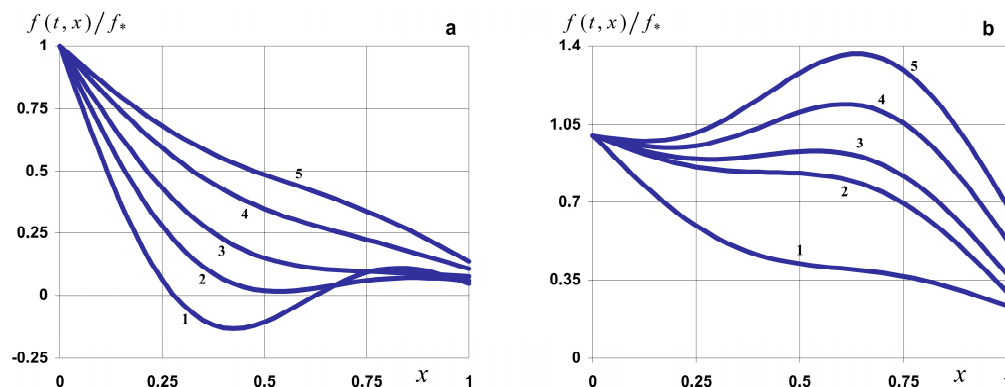

Figure 14. Graphs of the function $f(t,x)/f_*$ at different moments for $D = 0.05$ (**a**) and $D = 1$ (**b**).
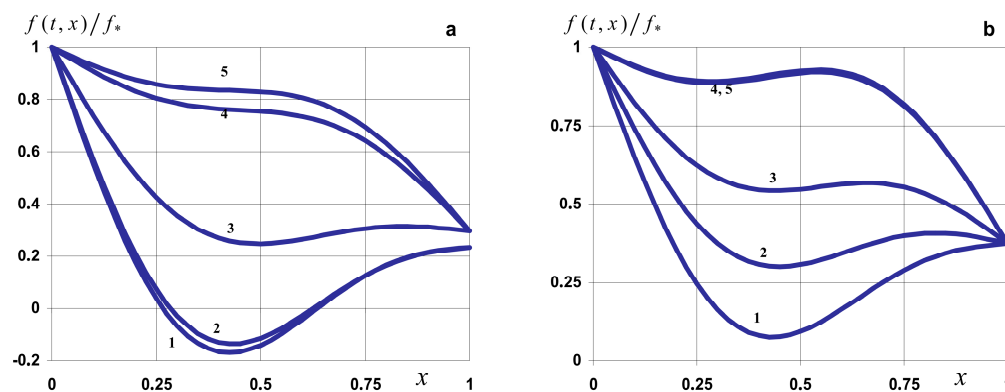

Figure 15. Graphs of the function $f(t,x)/f_*$ for different values of the coefficient D at the moments $t = 0.1$ (**a**) and $t = 1$ (**b**).
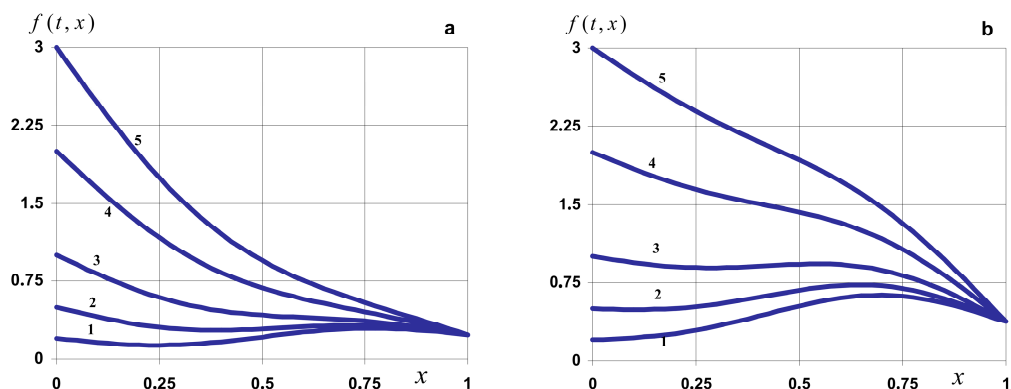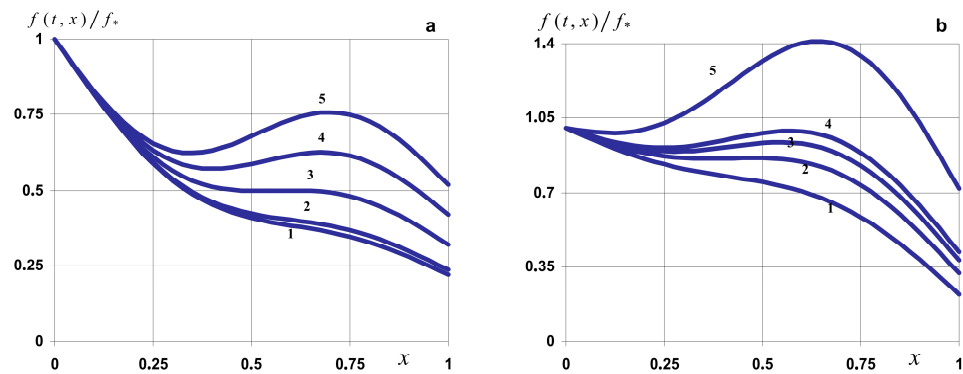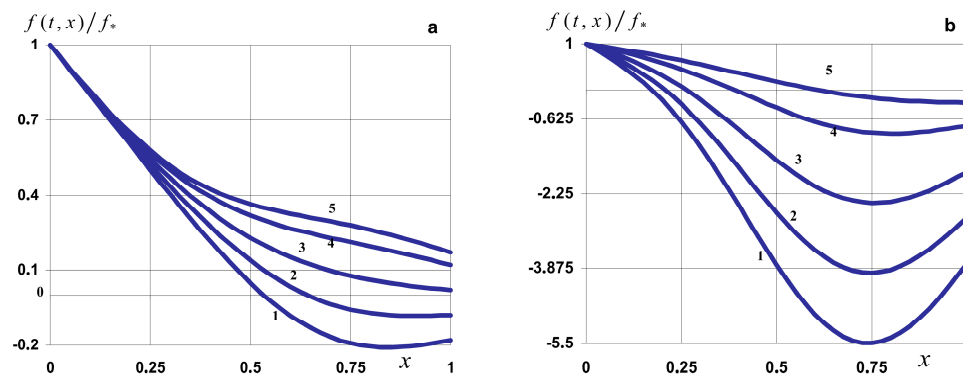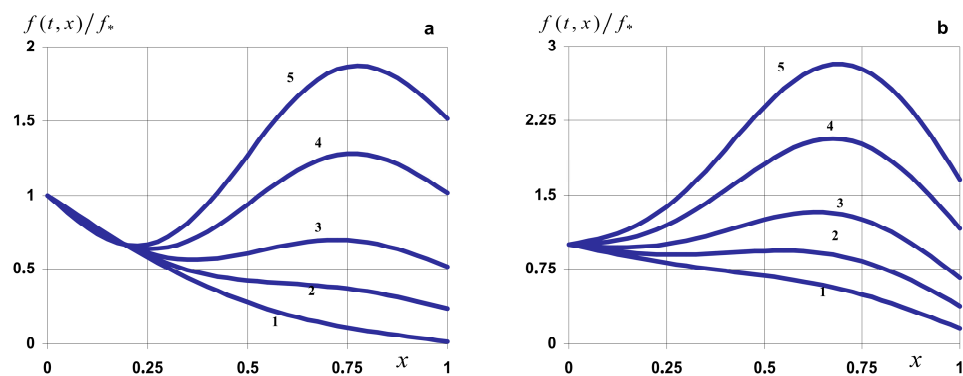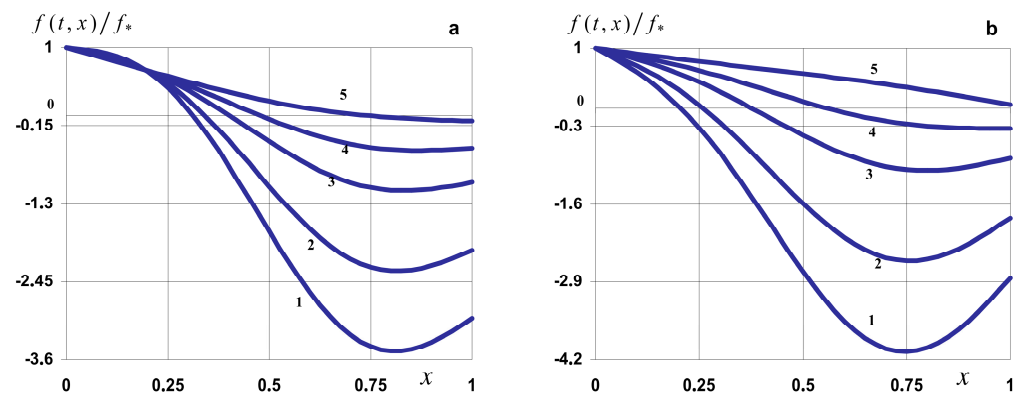

Figure 16. Graphs of the function $f(t,x)/f_*$ for different values of the parameter f_* at the moments $t = 0.1$ (**a**) and $t = 1$ (**b**).

Note that, as the duration of the process described by the initial-boundary value problem (1)–(3) and (7) increased, the function $f(t,x)$ increased in the entire area of the body (Figure 14). Moreover, for small values of the coefficient D and small t, there were local maximum and minimum functions $f(t,x)$. For example, $\min\limits_{x \in [0,x_0]} f(t,x)\Big|_{\substack{t = 0.1 \\ D = 0.05}} = -0.130214511$ at the point

$x_{\min} = 0.45$ and $\max\limits_{x \in [0,x_0]} f(t,x) \bigg|_{\substack{t=0.1 \\ D=0.05}} = 0.10893835$ at the point $x_{\max} = 0.825$ (curve 1,

Figure 14a). Over time, these local minimums and maximums leveled off (curves 2 and 3, Figure 14a), and the function monotonically decreased (curves 4 and 5, Figure 14a). For the coefficient $D \geq 1$, the behavior of the function $f(t,x)$ differed. For short durations, it monotonically decreased (curve 1, Figure 14b), and a local maximum began to form in the lower half of the layer (curve 2, Figure 14b) which further increased, shifted to the lower boundary of the body and became global (curves 3–5, Figure 14b).

Figure 17. Graphs of the function $f(t,x)/f_*$ for different values of the linear regression coefficient a, $a > 0$ at the moments $t = 0.1$ (**a**) and $t = 1$ (**b**).

Figure 18. Graphs of the function $f(t,x)/f_*$ for different values of the linear regression coefficient a, $a < 0$ at the moments $t = 0.1$ (**a**) and $t = 1$ (**b**).

Figure 19. Graphs of the function $f(t,x)/f_*$ for different values of the linear regression coefficient b, $b > 0$ at the moments $t = 0.1$ (**a**) and $t = 1$ (**b**).

Figure 20. Graphs of the function $f(t,x)/f_*$ for different values of the linear regression coefficient b, $b < 0$ at the moments $t = 0.1$ (**a**) and $t = 1$ (**b**).

The influence of the coefficient D on the behavior and values of the solution of the initial-boundary value problem can be significant (Figure 15). However, for small time intervals, a change in small values of D, and for large time intervals, a change in large values of D, have almost no effect on the function $f(t,x)$ in short intervals (curves 1 and 2 in Figure 15a and curves 4 and 5 in Figure 15b). We note that, the larger the value of the coefficient D, the greater the growth observed in the values of $f(t,x)$ becomes (Figure 15).

Small values of f_* are characterized by the presence of the local maximum at point $x_{max} = 0.775$ for small values of t (curve 1, Figure 16a) and at point $x_{max} = 0.7$ for large values of t (curve 1 in Figure 16b). Here, the value of the maximum increases approximately twice: $\left. \max_{x \in [0,x_0]} f(t,x) \right|_{\substack{t=1 \\ f_*=0.2}} / \left. \max_{x \in [0,x_0]} f(t,x) \right|_{\substack{t=0.1 \\ f_*=0.2}} = 2.13$. With the increasing values of f_*, for small times, the function $f(t,x)$ becomes concave (curves 3–5 in Figure 16a), and for large times, it becomes convex (curves 3–5 in Figure 16b).

The value of the linear regression coefficient a, in the region of its positive values (Figure 17) and its negative values (Figure 18), has almost no effect on the function $f(t,x)$ in the vicinity of the layer boundary $x = 0$. Here, the size of such an interval is $[0, 0.2]$ for short durations (Figures 17a and 18a), which decreases with time, and for $t = 1$, it is $[0, 0.06]$ (Figures 17b and 18b). In the rest of the layer, the influence of the coefficient a on the solution to the initial-boundary value problem is significant. Moreover, the higher the values of this coefficient, the higher the values of the function $f(t,x)$ (Figures 17 and 18). An increase in the coefficient a on the positive semi-axis of the real numbers changes the behavior from a monotonically decreasing function for small a to the formation of a local maximum in the second half of the layer for large a (curves 4 and 5 in Figure 17a and curves 2–5 in Figure 17b). And for long durations, this maximum of $f(t,x)$ becomes global (curve 5, Figure 17b). For negative small values of the coefficient a, there is a minimum of the function $f(t,x)$ in the lower part of the body. For example, for short durations, the minimum point is $x_{min} = 0.8$ (curve 1 in Figure 18a), and for long durations, $x_{min} = 0.75$ (curve 1, Figure 18b). An increase in the negative values of the coefficient a leads to an increase in $f(t,x)$ at the entire interval and a change in the behavior of the function to a monotonically decreasing one (curves 3–5 in Figure 18a and curve 5 in Figure 18b).

The value of the free term of the linear regression b, in the region of its positive values (Figure 19) and its negative values (Figure 20), has a weak effect on the function $f(t,x)$ in the vicinity of the layer boundary $x = 0$. The size of such an interval is $[0, 0.225]$ for short durations (Figures 17a and 18a), which decreases with time, and for $t = 1$, it is $[0, 0.1]$ (Figures 19b and 20b). In the rest of the body region, the coefficient b significantly affects the behavior and values of the solution to the initial-boundary value problem (1)–(3) and (7). As for the parameter a, the higher the coefficient b becomes, the higher values the function $f(t,x)$ reaches (Figures 19 and 20). The increase in the coefficient b on the

positive semi-axis of the real numbers changes the behavior of $f(t,x)$ from a monotonically decreasing function for small b (curves 1 and 2 in Figure 19a and curve 1 in Figure 19b) to the formation of a local maximum in the second half of the layer for large b (curves 3–5 in Figure 19a and curves 2–5 in Figure 19b). Moreover, for long durations times, this maximum of $f(t,x)$ becomes global (curves 3–5, Figure 19b). Similarly to changes in the coefficient a, the effect of the parameter b is observed, if $b < 0$, on the function $f(t,x)$. Thus, for negative small values of the coefficient b, there is a minimum of the function $f(t,x)$ in the lower part of the body. For example, for short durations, the minimum point is $x_{\min} = 0.8$ (curve 1 in Figure 20a), and for long durations, $x_{\min} = 0.75$ (curve 1, Figure 20b). An increase in the negative values of the coefficient b leads to an increase in $f(t,x)$ over the entire interval and a change in the behavior of the function to a monotonically decreasing one (curve 5 in Figure 20a and curves 3–5 in Figure 20b).

Taking into consideration the results obtained in Section 3, we can say that the larger the sampling size N is in the same time interval $[0, t_N]$, the closer the function $f(t,x)$ approaches to monotonically decreasing. The higher the covariance $\text{cov}(t, F_e(t))$ becomes, the larger the function $f(t,x)$ becomes. For a larger dispersion of σ_t^2 a decrease in the values of the solution to the problem occurs in the entire body region, and an increase in the values of $f(t,x)$ for $\forall x \in [0, x_0]$ for larger values of dispersion σ_F^2 can also be observed. An increase in the correlation coefficient R also leads to an increase in the function $f(t,x)$ and the formation of a local or global maximum in the lower part of the layer.

For all the considered samples, the solution to the origin initial-boundary value problem increases with time from the monotonically decreasing behavior of the function at small time intervals, and a local maximum forms in the lower part of the layer (Figure 21). For small time intervals of the process described by the problem in (1)–(3) and (7), and for large samples, the value of the sample variance does not significantly affect the value of $f(t,x)$ (curves 1 and 2 in Figure 21). Thus, the maximum difference between the values of $f(t,x)|_{\substack{sampleI \\ t=0.1}}$ and $f(t,x)|_{\substack{sampleII \\ t=0.1}}$ is up to 8%. For samples with small sizes (Samples III, V, and VI), both a large interval $[0, t_N]$ and a small sample variance can lead to a significant increase in the values of the function $f(t,x)$ (curve 5, Figure 21a). From the comparison of Figures 21a and 21b, it follows that the growth rate of the values of the function $f(t,x)$ for samples over large time intervals and with large variance (curves 1 and 3 in Figure 21b) is much higher than the growth rates of other samples (curves 2, 5 and 6 in Figure 21b). The significantly larger values of curves 4 (Sample IV) in Figure 21a,b compared to the other curves can be explained by the much larger experimental values of $f_{x_0 i}$ for $\forall i = 1, \dots, N$, than those presented in the other samples.
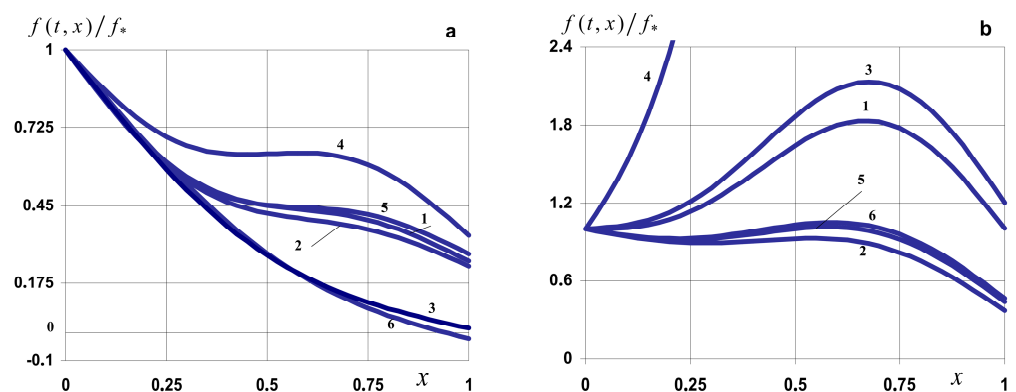

Figure 21. Graphs of the function $f(t,x)/f_*$ for samples presented in Section 7 at the moments $t = 0.1$ (**a**) and $t = 1$ (**b**).

Therefore, all the statistical characteristics considered in the paper sufficiently affect the behavior of the solution of the original initial-boundary value problem, as well as

the two-sided statistical estimation of the desired function, its reliable intervals, and its two-sided critical regions. Before studying the solution to the transfer initial-boundary value problem in the presence of experimental data on the boundary, it is necessary to investigate the statistical characteristics of the sample in detail.

10. Conclusions

Herein, we have proposed a mathematical model of the transfer processes in the layer described by second-order partial differential equations, when a constant value of the desired function is given at the upper boundary of the layer and only experimental data on the desired function at certain moments of time are known at the lower boundary. The linear regression, which is considered as a boundary condition, is constructed using the experimental data with the least squares method. The solution to the formulated initial-boundary value problem is obtained by means of finite integral Fourier transform.

The influences of statistical characteristics of the sample of experimental data such as sample size, covariance, variance in the time variable, and variance in the results of the measurements, as well as the correlation coefficient, on the coefficients of linear regression constructed on these data are analyzed.

The two-sided statistical estimation of the solution to the initial-boundary value problem is determined in terms of the coefficients of linear regression, which are analyzed in relation to the influence of the sample size and the covariance.

Confidence intervals for the coefficients of the regression equation with a given level of reliability are established. The corresponding confidence intervals for the desired function are determined based on the obtained solution of the initial-boundary value problem, and their limits are determined according to the confidence intervals for the linear regression coefficients. It is shown that, the larger the absolute values of the sought function are, the greater the width of the confidence interval is. For short durations, the influence of the confidence interval for the angle of inclination of the regression equation is practically imperceptible, while the influence of the confidence interval for the free term of the regression is many times greater, especially in the vicinity of the lower boundary of the layer. It was also noted that the smaller the variance of the sample of experimental data, the smaller the width of the confidence interval for solution to the formulated initial-boundary value problem. The formula for determining the two-tailed critical region based on Fisher's test was obtained and analyzed.

The results were obtained as part of the work under a grant from the Ministry of Education and Science of Ukraine (project number 0123U101691).

References

1. Kovács, R.; Rogolino, P.; Oliveri, F. Mathematical aspects in non-equilibrium thermodynamics. *Symmetry* **2023**, *15*, 929. [CrossRef]
2. Rogers, D.; Beck, T.; Rempe, S. An information theory approach to nonlinear, nonequilibrium thermodynamics. *J. Stat. Phys.* **2011**, *145*, 385–409. [CrossRef] [PubMed]
3. Fahrmeir, L.; Kneib, T.; Lang, S.; Marx, B. *Regression: Models, Methods and Applications*; Springer: New York, NY, USA, 2021; 746p.
4. Waldmann, E. Quantile regression: A short story on how and why. *Stat. Model.* **2018**, *18*, 203–218. [CrossRef]
5. Gallardo, D.; Bourguignon, M.; Galarza, C.; Gómez, H. A Parametric Quantile Regression Model for Asymmetric Response Variables on the Real Line. *Symmetry* **2020**, *12*, 1938. [CrossRef]

6. Abboud, C.; Bonnefon, O.; Parent, E.; Soubeyrand, S. Dating and localizing an invasion from post-introduction data and a coupled reaction–diffusion–absorption model. *J. Math. Biol.* **2019**, *79*, 765–789. [CrossRef] [PubMed]
7. Soubeyrand, S.; Roques, L. Parameter estimation for reaction - diffusion models of biological invasions. *Popul. Ecol.* **2014**, *56*, 427–434. [CrossRef]
8. Faugeras, B.; Maury, O. Modeling fish population movements: From an individual-based representation to an advection-diffusion equation. *J. Theor. Biol.* **2007**, *247*, 837–848. [CrossRef] [PubMed]
9. Hooten, M.B.; Garlick, M.J.; Powell, J.A. Computationally Efficient Statistical Differential Equation Modeling Using Homogenization. *J. Agric. Biol. Environ. Stat.* **2013**, *18*, 405–428. [CrossRef]
10. Albni, R.A.S.; Albani, V.V.L.; Gomes, L.E.S.; Migon, H.S.; Silva Neto, A.J. Bayesian inference and wind field statistical modeling applied to multiple source estimation. *Environ. Pollut.* **2023**, *321*, 121061. [CrossRef]
11. Papaïx, J.; Soubeyrand, S.; Bonnefon, O.; Walker, E.; Louvrier, J.; Klein, E.; Roques, L. Inferring mechanistic models in spatial ecology using a mechanistic-statistical approach. In *Statistical Approaches for Hidden Variables in Ecology*; Wiley: Hoboken, NJ, USA, 2022; pp. 69–95. [CrossRef]
12. Potts, J.R.; Schlägel, U.E. Parametrizing diffusion-taxis equations from animal movement trajectories using step selection analysis. *Methods Ecol. Evol.* **2020**, *11*, 1092–1105. [CrossRef]
13. Lu, X.; Williams, P.J.; Hooten, M.B.; Powell, J.A.; Womble, J.N.; Bower, M.R. Nonlinear reaction–diffusion process models improve inference for population dynamics. *Environmetrics* **2020**, *31*, e2604. [CrossRef]
14. Hefley, T.J.; Hooten, M.B.; Russell, R.E.; Walsh, D.P.; Powell, J.A. When mechanism matters: Bayesian forecasting using models of ecological diffusion. *Ecol. Lett.* **2017**, *20*, 640–650. [CrossRef] [PubMed]
15. Scharr, H.; Black, M.; Haussecker, H.W. Image statistics and anisotropic diffusion. In Proceedings of the Ninth IEEE International Conference on Computer Vision, Nice, France, 13–16 October 2003; pp. 840–847. [CrossRef]
16. Zhu, S.C.; Mumford, D. Prior learning and Gibbs reaction-diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **1997**, *19*, 1236–1250. [CrossRef]
17. Zhang, B.; Zhou, Y.; Zhang, X.; Wang, Z.; Yang, W.; Ban, Y. Experimental Study on Grouting Diffusion Law of the Different Crack Widths in Tunnel Lining. *KSCE J. Civ. Eng.* **2023**, *27*, 1789–1799. [CrossRef]
18. Aminul, M.; Hoque, M.A. Assessment onnon-linear models for demonstrating heavy metals release behavior from solid waste block. *J. Solid Waste Technol. Manag.* **2016**, *42*, 157–168. [CrossRef]
19. Kari, O.P.; Puttonen, J.; Skantz, E. Reactive transport modelling of long-term carbonation. *Cem. Concr. Compos.* **2014**, *52*, 42–53. [CrossRef]
20. Tarighat, A. Stochastic modeling and calibration of chloride content profile in concrete based on limited available data. *Int. J. Civ. Eng.* **2012**, *10*, 309–316.
21. Kawaharada, A.; Iima, M. An application of data-based construction method of cellular automata to physical phenomena. *J. Cell. Autom.* **2018**, *13*, 441–459.
22. Golmohammadi, J.; Ebert-Uphoff, I.; He, S.; Deng, Y.; Banerjee, A. High-dimensional dependency structure learning for physical processes. In Proceedings of the 2017 IEEE International Conference on Data Mining (ICDM), New Orleans, LA, USA, 18–21 November 2017; pp. 883–888. [CrossRef]
23. Sigrist, F.; Künsch, H.R.; Stahel, W.A. Stochastic partial differential equation based modelling of large space-time datasets. *J. R. Stat. Society. Ser. B Stat. Methodol.* **2015**, *77*, 3–33. [CrossRef]
24. Klyatskin, V.I. *Dynamics of Stochastic Systems*; Elsevier Science: Amsterdam, The Netherlands, 2005; 205p. [CrossRef]
25. Gusev, S.A. Estimation of the coefficients in the parabolic equation by the statistical simulation of diffusion trajectories. *Russ. J. Numer. Anal. Math. Model.* **2003**, *18*, 297–305. [CrossRef]
26. Tonaki, Y.; Kaino, Y.; Uchida, M. Parameter estimation for a linear parabolic SPDE model in two space dimensions with a small noise. *Stat. Inference Stoch. Process* **2024**, *27*, 123–179. [CrossRef]
27. Nakakita, S.H.; Uchida, M. Inference for convolutionally observed diffusion processes. *Entropy* **2020**, *22*, 1031. [CrossRef] [PubMed]
28. Wei, S.; Panaretos, V.M. Empirical evolution equations. *Electron. J. Stat.* **2018**, *12*, 249–276. [CrossRef]
29. Flath, H.P.; Wilcox, L.C.; AkçElik, V.; Hill, J.; Van Bloemen Waanders, B.; Ghattas, O. Fast algorithms for bayesian uncertainty quantification in large-scale linear inverse problems based on low-rank partial hessian approximations. *SIAM J. Sci. Comput.* **2011**, *33*, 407–432. [CrossRef]
30. Malara, G.; Spanos, P.D.; Jiao, Y. Efficient calculation of the response statistics of two-dimensional fractional diffusivesy stems. *Probabilistic Eng. Mech.* **2020**, *59*, 103036. [CrossRef]
31. Nye, T.M.W.; White, M.C. Diffusion on some simple stratified spaces. *J. Math. Imaging Vis.* **2014**, *50*, 115–125. [CrossRef]
32. Bakhrushin, V.E. *Methods of Data Analysis*; KPU: Zaporizhzhia, Ukraine, 2011; 268p. (In Ukrainian)
33. Grigelionis, B.; Prohorov, Y.V.; Sazonov, V.V.; Statulevičius, V. *Probability Theory and Mathematical Statistics*; De Gruyter: Berlin, Germany, 2020; 624p.
34. Sneddon, I. *The Use of Integral Transforms*; Tata Mc Graw-Hill: New York, NY, USA, 1979; 539p.
35. Abramowitz, M.; Stegun, I. *Handbook of Mathematical Functions*; National Bureau of Standards: Washington, DC, USA, 1972; 1046p.
36. Apostol, M. *Equations of Mathematical Physics*; Cambridge Scholars Publishing: Newcastle upon Tyne, UK, 2018; 250p.
37. Gumbel, E.J. *Statistics of Extremes*; Dover Publications: Chicago, IL, USA, 2004; 400p.

38. Vuchkov, I.; Boyadzhieva, L.; Solakov, E. *Applied Linear Regression Analysis*; Financy i Statistika: Moscow, Russia, 1987; 238p. (In Russian)
39. Prudnikov, A.P.; Brychkov, Y.A.; Marichev, O.I. *Integrals and Series. Elementary Functions*; Nauka: Moscow, Russia, 1981; 800p. (In Russian)
40. Ferrier, M. *Handbook of Mathematical Analysis*; Larsenand Keller Education: New York, NY, USA, 2023; 240p.
41. Kartashov, M.V. *Probability, Processes, Statistics*; Kyiv University Publ.: Kyiv, Ukraine, 2007; 504p. (In Ukrainian)