



Jamil Al-Azzeh^{1,*}, Abdelwadood Mesleh², Maksym Zaliskyi³, Roman Odarchenko³ and Valeriyi Kuzmin³

- ¹ Electrical Engineering Department, Faculty of Engineering Technology, Al-Balqa Applied University, Amman 19117, Jordan
- ² Faculty of Artificial Intelligence, Al-Balqa Applied University, Amman 19117, Jordan
- ³ Department of Telecommunication and Radioelectronic Systems, National Aviation University,
 - 03058 Kyiv, Ukraine
- Correspondence: jamil.azzeh@bau.edu.jo

Abstract: The main purpose of mathematical model building while employing statistical data analysis is to obtain high accuracy of approximation within the range of observed data and sufficient predictive properties. One of the methods for creating mathematical models is to use the techniques of regression analysis. Regression analysis usually applies single polynomial functions of higher order as approximating curves. Such an approach provides high accuracy; however, in many cases, it does not match the geometrical structure of the observed data, which results in unsatisfactory predictive properties. Another approach is associated with the use of segmented functions as approximating curves. Such an approach of estimating the coordinates of the breakpoint between adjacent segments. This article proposes a new method for determining abscissas of the breakpoint for segmented regression, minimizing the standard deviation based on multidimensional paraboloid usage. The proposed method is explained by calculation examples obtained using statistical simulation and real data observation.

Keywords: mathematical model building; ordinary least squares; segmented regression; optimization of breakpoint abscissa; multidimensional paraboloid; accuracy increment

1. Introduction

Scientists use various models when studying different environmental phenomena. Mathematical models provide an opportunity to determine equations and dependencies to correlate the parameters of miscellaneous objects and processes. Mathematical models are built for various reasons, including the achievement of the best understanding of the objects under study, the possibility of mathematical analysis, and the possibility of conducting experimentation with the model in case it is difficult to repeat the experiment with the objects under study [1].

The process of mathematical model building contains several steps:

- (1) Experimental study and the measuring of the parameters of real-world systems and phenomena;
- (2) Collecting initial data for the model;
- (3) Mathematical formulations and fitting one or more models;
- (4) The statistical simulation of the model to validate it [2].

There are general rules for building mathematical models. These rules assume the following: (1) collecting background information for the phenomenon under study, (2) using simple models at the first stage, (3) determining all parameters and the quantities and correlations between them based on data analysis, (4) complicating the model based on the nature of the phenomenon under study, (5) estimating the efficiency of the model, and (6) others [3]. The efficiency analysis involves choosing the optimal mathematical model for the problem considered.



Citation: Al-Azzeh, J.; Mesleh, A.; Zaliskyi, M.; Odarchenko, R.; Kuzmin, V. A Method of Accuracy Increment Using Segmented Regression. *Algorithms* **2022**, *15*, 378. https://doi.org/10.3390/a15100378

Academic Editor: George Karakostas

Received: 2 September 2022 Accepted: 13 October 2022 Published: 17 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). There are various efficiency measures for mathematical models. Generally, researchers use the following parameters:

- (1) Accuracy—for the coincidence analysis of the output of a mathematical model with observed data;
- (2) Reliability—for the analysis of the precision of a mathematical model;
- (3) Transparency—for the analysis of choices and assumptions of the output expectations [4,5].

To analyze mathematical models, researchers can use additional criteria, such as model simplicity, calculation time, costs, depth level, and others.

The main parameters for the efficiency level of mathematical models in terms of accuracy analysis are standard deviation [6,7], the sum of absolute deviations between the model output and the observed data [8], a weighted sum of squared deviations [9,10], and the maximal deviation [11]. The criterion for these parameters is the minimum value of the estimated parameter [12,13].

This article contains seven sections. The first section discusses the background information for the problems of mathematical model building. The second section presents a literature review regarding the topic of research and presents the statement of the problem. The third section deals with the description of mathematical tools for segmented regression building while using ordinary least squares. The fourth section proposes the step-by-step procedure for accuracy increment during segmented regression usage. The fifth section concentrates on the analysis of the proposed method based on statistical simulations. The sixth section discusses the implementation of the proposed method in real data examples, and the seventh section presents the conclusions.

2. Literature Review and Statement of the Problem

Mathematical model building aims at decreasing the uncertainty level for the objects being studied [14,15]. The analysis of the level, location, and nature of uncertainty helps to obtain more reliable information and adequate knowledge [16,17].

To build mathematical models, researchers use methods from different sciences, such as mathematical analysis, probability theory, data science, regression analysis, mathematical statistics, recognition theory, applied geometry, and others [18].

This article concentrates on the techniques of regression analysis for mathematical model building, so corresponding methods are considered in detail. Regression analysis is used to determine the relationship between two or more variables [19] and is widely used to fit mathematical models to statistical data [20].

Regression analysis is frequently used in various applications due to its approximate ease of calculation, high accuracy, and good predictive properties, depending on the approximating function type usage. Regression analysis is applied to different fields in different capacities, for example, in:

- Medicine: to detect Parkinson's disease based on the analysis of finger-tapping data [21], to forecast the uptake of oxygen based on genes evaluation and to predict data on patient admission [22], and others;
- (2) Econometrics: to predict the audit opinion using six financial indicators [23], to determine the dependence of economic growth on the level of environmental pollution [24], to describe the trends of economical parameters in correlation with various factors [25,26], and others;
- (3) Transport systems: to determine the optimal periodicity of the implementation of operation processes [27,28], and to analyze possible routes and traffic intensity [29–31];
- (4) Aviation: to identify flight conditions and situations based on diagnostic parameter monitoring [32,33], and to predict the human state and decision making depending on various environmental factors [34,35];
- (5) Radar systems: to estimate the efficiency of signal detection [36], to determine the dependence of weather parameters on radar-received signals [37–39], and others;

- (6) Navigation systems: to build a mathematical model for the optimal selection of the navigation equipment [40–42], to establish the correlation between navigation equipment failures [43], to approximate operational data trends for the prediction of possible aviation events [44], and others;
- (7) Cybersecurity: to evaluate the efficiency of information web-resources functioning [45], to synthesize data-processing algorithms while detecting cyberattacks [46–48], to ensure high-level security against cyberattacks [49,50], and others;
- (8) Engineering and control: to describe nonlinear dynamic object behavior [51,52], to build the mathematical model for statistical parameters while designing control systems [53], to make decisions based on statistical information processing [54,55], and others;
- (9) Equipment maintenance: to build the mathematical model for diagnostic variable trends [56], and to determine the uncertainty level while conducting condition monitoring and maintenance preference analysis [57,58];
- (10) Reliability analysis: to describe the behavior of reliability parameters [59,60], to simulate statistically nonstationary random processes of failures occurrence [61,62], to describe the processes of technical condition deterioration in the trend of failure rate [63,64], and others.

Regression analysis usually starts with research on the possibility of using a linear regression model. In the case of an unsatisfactory level of accuracy, more complicated models are used [65]. These models are nonlinear regression models [66]. Nonlinear regression models suggest parabolic, hyperbolic, exponential, segmented, and other approximating functions [65,67]. Because of the complicated calculations required when using a nonlinear regression model, various software can be utilized [68].

There are various methods for increasing the accuracy and predictive properties of mathematical models. One approach is to use segmented regression [69,70]. In this case, it is necessary to determine the coordinates of the breakpoint between adjacent segments. This problem can be solved using various algorithms [69–75]. These algorithms use the maximum likelihood estimator [69,70], Bayesian changepoint models [71,72], inverted F test [73], random search method, the method of cumulative sums [74,75], and others. A comparative analysis showed some flaws in the algorithms for determining breakpoint coordinates. These flaws are related to a need for prior limitations, as well as the effectiveness of the obtained estimate in terms of robustness and bias. Additionally, the discussed algorithms do not give the possibility to obtain a single mathematical formula for breakpoint coordinates and require the usage of the iterative numerical method described in [76].

The considered literature review motivates authors to synthesize a new approach for calculating the optimal coordinates of breakpoints while using segmented regression and analyzing time series with nonstationary behavior. The building of a mathematical model based on segmented regression usage is of considerable importance because:

- 1. Using segmented regression gives the possibility to obtain a model with greater accuracy.
- 2. Segmented regression more correctly describes the geometrical structure of time series.
- 3. The obtained segmented models have effective predictive properties.

The research gap in the field of mathematical model building is associated with the absence of a step-by-step procedure for determining the optimal segmented regression model in case of multiple breakpoints in a dataset structure. At the same time, to solve such problems, the method of simple enumeration of the possible options is often used. However, such an approach does not provide mathematical formulations and requires a long computing time.

Therefore, the goal of this article is: (1) to describe the technique of segmented regression building and (2) to obtain mathematical equations for a step-by-step procedure of accuracy increment based on optimal breakpoints abscissas calculations.

Let us state the research problem mathematically. Let us present the statistical dataset in two arrays $Y = \{y_i\}$ and $X = \{x_i\}$, each with sample size *n*. *Y* is the dependent or response variable, while *X* is the independent or predictor variable. The relationship between the variables is determined by the function set $\phi_k(X, \vec{c}_{m,k})$, where *k* describes the quantity of the model being fitted to the dataset and $\vec{c}_{m,k}$ is a vector of *m* parameters for the *k*-th regression model. In this case, the regression model is determined by the equation [65]

$$Y = \phi_k(X, \vec{c}_{m,k}) + \Delta,$$

where Δ is an error, which can be described by a normal probability density function. Such an assumption allows the use of ordinary least squares (OLS). For example, in the case of linear regression, $\phi_0(X, \vec{c}_{m,0}) = c_{0,0} + c_{1,0}X$, where $c_{0,0}$ and $c_{1,0}$ are coefficients to be estimated.

This paper focuses on increasing the accuracy of mathematical models based on segmented regression usage. In this case, the function set $\phi_k(X, \overrightarrow{c}_{m,k}, x_{\text{br } q, k})$ depends on abscissas $x_{\text{br } q, k}$ of the breakpoints, where q is the quantity of breakpoints. The accuracy of the model using OLS is usually estimated by the standard deviation σ between the model output and the observed data. The standard deviation depends on the values of abscissas $x_{\text{br } q, k}$ of the breakpoint. Thus, this paper aims to solve the minimization problem that can be formulated as follows:

$$\left\{x_{\mathrm{br opt1}}, x_{\mathrm{br opt2}}, \ldots, x_{\mathrm{br opt } q}\right\} = \arg \min(\sigma(x_{\mathrm{br 1}}, x_{\mathrm{br 2}}, \ldots, x_{\mathrm{br } q})).$$

3. Segmented Regression Models

This section presents the basic mathematical equations for different segmented regression models. Authors mostly employ piecewise linear, linear-quadratic, and quadratic models.

1. Segmented linear regression (SLR)

This regression type is a sequential connection of q + 1 straight-line segments without discontinuities. The mathematical model of SLR is given as

$$\phi_1(X, \overrightarrow{c}_{m,1}, x_{\mathrm{br}\,q,1}) = c_{0,1} + c_{1,1}X + \sum_{i=1}^q c_{i+1,-1}(X - x_{\mathrm{br}\,-i})h(X - x_{\mathrm{br}\,-i})$$
(1)

where $h(X - x_{br i})$ is the Heaviside function. This function helps to obtain the single mathematical equation for the segmented model.

An example of a mathematical model of three-segmented linear regression has the form

$$p_1(X, \vec{c}_{m,1}, x_{\text{br} q,1}) = c_{0,1} + c_{1,1}X + c_{2,1}(X - x_{\text{br} 1})h(X - x_{\text{br} 1}) + c_{3,1}(X - x_{\text{br} 2})h(X - x_{\text{br} 2})$$

This model has two breakpoints, $x_{br 1}$ and $x_{br 2}$, and it requires the computation of four unknown coefficients: $c_{0, 1}, c_{1, 1}, c_{2, 1}$, and $c_{3, 1}$. These coefficients are estimated based on the OLS. The computation result can be presented in the form of matrix equations

$$C = \Omega^{-1} \Psi, C = \begin{pmatrix} c_{0,1} \\ c_{1,1} \\ c_{2,1} \\ c_{3,1} \end{pmatrix}, \Psi = \begin{pmatrix} \sum_{\substack{i=1 \\ p \\ i=1 \\ y_i x_i \\ y_i (x_i - x_{br-1}) \\ \sum_{\substack{\forall x_i > x_{br-1} \\ \forall x_i > x_{br-2}}} y_i (x_i - x_{br-2}) \end{pmatrix},$$

$$\Omega = \begin{pmatrix} n & \sum_{i=1}^{n} x_i & \sum_{\forall x_i > x_{br-1}} (x_i - x_{br-1}) & \sum_{\forall x_i > x_{br-2}} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \sum_{\forall x_i > x_{br-1}} x_i (x_i - x_{br-1}) & \sum_{\forall x_i > x_{br-2}} x_i (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-1}) & \sum_{i=1}^{n} x_i (x_i - x_{br-1}) & \sum_{i=1}^{n} (x_i - x_{br-1}) \\ \sum_{i=1}^{n} (x_i - x_{br-1}) & \sum_{i=1}^{n} (x_i - x_{br-1}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-1}) & \sum_{i=1}^{n} (x_i - x_{br-1}) & \sum_{i=1}^{n} (x_i - x_{br-1})^2 & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-1}) (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-1}) (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{i=1}^{n} (x_i - x_{br-2}) & \sum_{i=1}^{n} (x_i - x_{br-2}) \\ \sum_{$$

where $\forall x_i > x_{br \ 1 \ (2)}$ corresponds to all x_i greater than $x_{br \ 1 \ (2)}$.

2. Segmented quadratic regression (SQR)

This regression type is a sequential connection of q + 1 quadratic parabola segments without discontinuities. The mathematical model of SQR is given as

$$\phi_2(X, \overrightarrow{c}_{m,2}, x_{\text{br} q,2}) = c_{0,2} + c_{1,2}X + c_{2,2}X^2 + \sum_{i=1}^q c_{i+2,2}(X - x_{\text{br} i})^2 h(X - x_{\text{br} i}) \quad (2)$$

An example of a mathematical model of two-segmented quadratic regression has the form

$$\phi_2(X, \vec{c}_{m,2}, x_{\text{br } q,2}) = c_{0,2} + c_{1,2}X + c_{2,2}X^2 + c_{3,2}(X - x_{\text{br } 1})^2 h(X - x_{\text{br } 1})$$

This model has one breakpoint, x_{br}_1 , and it requires the computation of four unknown coefficients: $c_{0, 2}$, $c_{1, 2}$, $c_{2, 2}$, and $c_{3, 2}$. These coefficients are estimated based on the OLS. The computation result can be presented in the form of matrix equations

$$\Omega = \begin{pmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i^2 \\ \sum_{i=1}^{n} x_i^2$$

3. Segmented linear-quadratic regression (SLQR)

This regression type is a sequential connection of q + 1 straight lines and quadratic parabola segments without discontinuities. The mathematical model of SLQR is given as

$$\phi_{3}(X, \overrightarrow{c}_{m,3}, x_{\mathrm{br}\,q,3}) = c_{0,3} + c_{1,3}X + c_{2,3}X^{2}s(0) + \sum_{i=1}^{q} c_{i+2,3}(X - x_{\mathrm{br}\,i})^{s(i)+1}h(X - x_{\mathrm{br}\,i}), \tag{3}$$

where s(i) is an indicator function. If the segment is a straight line, s(i) = 0. If the segment is a quadratic parabola, s(i) = 1.

An example of a mathematical model of two-segmented linear-quadratic regression has the form

$$f_3(X) = c_{0,3} + c_{1,3}X + c_{2,3}X^2 - c_{3,3}(X - x_{br 1})^2 h(X - x_{br 1}).$$

This model has one breakpoint, x_{br}_1 , and it requires the computation of three unknown coefficients: $c_{0,3}$, $c_{1,3}$, and $c_{2,3}$. The feature of this model is the equality of adjacent coefficients for the transition between the quadratic parabola segment and the straight-line segment. Thus, $c_{3,3} = c_{2,3}$. The coefficients are estimated based on the OLS. The computation result can be presented in the form of matrix equations

$$C = \Omega^{-1} \Psi, \ C = \begin{pmatrix} c_{0,3} \\ c_{1,3} \\ c_{2,3} \end{pmatrix}, \ \Psi = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} y_i x_i \\ \sum_{i=1}^{n} y_i x_i^2 - \sum_{\forall x_i > x_{br-1}} y_i (x_i - x_{br-1})^2 \end{pmatrix},$$

$$\Omega = \begin{pmatrix} n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 - \sum_{\forall x_i > x_{br-1}} (x_i - x_{br-1})^2 \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 - \sum_{\forall x_i > x_{br-1}} x_i (x_i - x_{br-1})^2 \\ \sum_{i=1}^{n} x_i^2 - \sum_{\forall x_i > x_{br-1}} (x_i - x_{br-1})^2 & \sum_{i=1}^{n} x_i^3 - \sum_{\forall x_i > x_{br-1}} x_i (x_i - x_{br-1})^2 \\ \sum_{i=1}^{n} x_i^2 - \sum_{\forall x_i > x_{br-1}} (x_i - x_{br-1})^2 & \sum_{i=1}^{n} x_i^3 - \sum_{\forall x_i > x_{br-1}} x_i (x_i - x_{br-1})^2 \\ \sum_{i=1}^{n} x_i^2 - \sum_{\forall x_i > x_{br-1}} (x_i - x_{br-1})^2 & \sum_{i=1}^{n} x_i^3 - \sum_{\forall x_i > x_{br-1}} x_i (x_i - x_{br-1})^2 \\ \sum_{i=1}^{n} x_i^2 - \sum_{\forall x_i > x_{br-1}} (x_i - x_{br-1})^2 & \sum_{i=1}^{n} x_i^3 - \sum_{\forall x_i > x_{br-1}} x_i (x_i - x_{br-1})^2 \\ \sum_{i=1}^{n} x_i^2 - \sum_{\forall x_i > x_{br-1}} (x_i - x_{br-1})^2 & \sum_{i=1}^{n} x_i^3 - \sum_{\forall x_i > x_{br-1}} x_i (x_i - x_{br-1})^2 \\ \sum_{i=1}^{n} x_i^2 - \sum_{\forall x_i > x_{br-1}} (x_i - x_{br-1})^2 & \sum_{i=1}^{n} x_i^3 - \sum_{\forall x_i > x_{br-1}} x_i (x_i - x_{br-1})^2 \\ \sum_{i=1}^{n} x_i^2 - \sum_{\forall x_i > x_{br-1}} (x_i - x_{br-1})^2 & \sum_{i=1}^{n} x_i^3 - \sum_{\forall x_i > x_{br-1}} x_i (x_i - x_{br-1})^2 \\ \sum_{i=1}^{n} x_i^2 - \sum_{\forall x_i > x_{br-1}} (x_i - x_{br-1})^2 & \sum_{i=1}^{n} x_i^3 - \sum_{\forall x_i > x_{br-1}} x_i (x_i - x_{br-1})^2 \\ \sum_{i=1}^{n} x_i^2 - \sum_{\forall x_i > x_{br-1}} (x_i - x_{br-1})^2 & \sum_{i=1}^{n} x_i^2 - \sum_{\forall x_i > x_{br-1}} (x_i - x_{br-1})^2 \\ \sum_{i=1}^{n} x_i^2 - \sum_$$

4. Step-by-Step Procedure for Accuracy Increment during Segmented Regression Usage

The method of accuracy increment during segmented regression usage is associated with the estimation of breakpoint abscissas. The breakpoint is the point of connection between two neighboring segments.

The step-by-step procedure contains the following operations:

- 1. Choosing of the regression model and the quantity of segments. At this stage, the researcher analyzes the geometrical structure of the observed data presented graphically in the form of the dependence of *Y* on *X*. After that, based on their experience, the researcher must choose one of the models SLR, SQR, and SLQR. To substantiate the decision on segmented regression usage, the researcher can test the initial data for nonlinearity. The geometrical structure of the observed data also gives the ability to choose the quantity *q* of the breakpoints
- 2. Determining the possible range of values of the breakpoint abscissas. At this stage, the researcher subjectively chooses the discrete range for all breakpoints. The minimal quantity of discrete values should be greater than five. The result of this step is a two-dimensional array x_{br} with size $q \times w$, where w is the number of discrete values in the range of breakpoint abscissas.
- 3. Building a regression model. At this stage, based on the matrix equations presented in the previous section, the researcher calculates the unknown coefficients for the chosen regression model and all possible values in the array x_{br} .
- 4. Calculating the standard deviations. In the case of OLS usage, the accuracy of the model is determined by the standard deviation between the model output and the observed data, which can be presented as follows:

$$\sigma = \sqrt{\frac{1}{n-l}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2},$$

where *l* is the degree of freedom for the chosen regression model.

At this stage, it is necessary to determine the discrete multidimensional dependence $\sigma(x_{br 1}, x_{br 2}, ..., x_{br q})$ for all possible values in the array x_{br} .

Note that in the case of an alternative regression method (for example, least absolute deviations regression), similar calculations for corresponding accuracy measures should be completed.

- 5. Approximating the standard deviation dependence on the breakpoint abscissas by multidimensional paraboloid using OLS. The dimension of the paraboloid corresponds to the quantity *q* of breakpoints. It is possible to use one of two types of paraboloid:
 - (a) General:

$$\Sigma\Big(x_{\rm br\ 1}, \ x_{\rm br\ 2}, \ \dots, \ x_{\rm br\ q}\Big) = \alpha_0 + \sum_{i=1}^q \alpha_i x_{\rm br\ i}^2 + \sum_{i=1}^q \beta_i x_{\rm br\ i} + \sum_{\forall i < j} \gamma_{i, j} x_{\rm br\ i} x_{\rm br\ j},$$
(4)

(a) Simplified:

$$\Sigma(x_{\rm br\ 1}, x_{\rm br\ 2}, \ldots, x_{\rm br\ q}) = \alpha_0 + \sum_{i=1}^q \alpha_i x_{\rm br\ i}^2 + \sum_{i=1}^q \beta_i x_{\rm br\ i}, \qquad (5)$$

where A_i , β_i , and $\gamma_{i,j}$ are approximation coefficients. The simplified paraboloid (5) can be used in case of assumptions about $\gamma_{i,j} = 0$ for the general paraboloid (4).

The coefficients of Equations (4) and (5) are estimated based on OLS. Such a calculation is possible, because all of the values of the possible breakpoints in the two-dimensional array X_{br} with size $q \times w$ are known, and function $\Sigma(x_{br 1}, x_{br 2}, \ldots, x_{br q})$ values correspond to the standard deviations $\sigma(x_{br 1}, x_{br 2}, \ldots, x_{br q})$ obtained at the previous step.

Consider the case of a simplified paraboloid. According to OLS, it is necessary to solve the system of equations

$$\frac{\partial}{\partial \alpha_{0}} \sum_{i_{1}=1}^{w} \dots \sum_{i_{q}=1}^{w} \left(\sigma \left(x_{\text{br 1 } i_{1}}, x_{\text{br 2 } i_{2}}, \dots, x_{\text{br q } i_{q}} \right) - \left(\alpha_{0} + \sum_{j=1}^{q} \alpha_{j} x_{\text{br } j, i_{j}}^{2} + \sum_{j=1}^{q} \beta_{j} x_{\text{br } j, i_{j}} \right) \right)^{2} = 0,$$

$$\frac{\partial}{\partial \alpha_{1}} \sum_{i_{1}=1}^{w} \dots \sum_{i_{q}=1}^{w} \left(\sigma \left(x_{\text{br 1 } i_{1}}, x_{\text{br 2 } i_{2}}, \dots, x_{\text{br q } i_{q}} \right) - \left(\alpha_{0} + \sum_{j=1}^{q} \alpha_{j} x_{\text{br } j, i_{j}}^{2} + \sum_{j=1}^{q} \beta_{j} x_{\text{br } j, i_{j}} \right) \right)^{2} = 0,$$

$$\frac{\partial}{\partial \beta_{1}} \sum_{i_{1}=1}^{w} \dots \sum_{i_{q}=1}^{w} \left(\sigma \left(x_{\text{br 1 } i_{1}}, x_{\text{br 2 } i_{2}}, \dots, x_{\text{br q } i_{q}} \right) - \left(\alpha_{0} + \sum_{j=1}^{q} \alpha_{j} x_{\text{br } j, i_{j}}^{2} + \sum_{j=1}^{q} \beta_{j} x_{\text{br } j, i_{j}} \right) \right)^{2} = 0,$$

$$\dots$$

$$\frac{\partial}{\partial \alpha_{q}} \sum_{i_{1}=1}^{w} \dots \sum_{i_{q}=1}^{w} \left(\sigma \left(x_{\text{br 1 } i_{1}}, x_{\text{br 2 } i_{2}}, \dots, x_{\text{br q } i_{q}} \right) - \left(\alpha_{0} + \sum_{j=1}^{q} \alpha_{j} x_{\text{br } j, i_{j}}^{2} + \sum_{j=1}^{q} \beta_{j} x_{\text{br } j, i_{j}} \right) \right)^{2} = 0,$$

$$\frac{\partial}{\partial \beta_{q}} \sum_{i_{1}=1}^{w} \dots \sum_{i_{q}=1}^{w} \left(\sigma \left(x_{\text{br 1 } i_{1}}, x_{\text{br 2 } i_{2}}, \dots, x_{\text{br q } i_{q}} \right) - \left(\alpha_{0} + \sum_{j=1}^{q} \alpha_{j} x_{\text{br } j, i_{j}}^{2} + \sum_{j=1}^{q} \beta_{j} x_{\text{br } j, i_{j}} \right) \right)^{2} = 0,$$

Let us simplify the first equation in the system. After derivative calculation, it can be presented as follows:

$$-2\sum_{i_{1}=1}^{w}\dots\sum_{i_{q}=1}^{w}\left(\sigma\left(x_{\text{br }1\ i_{1}},\ x_{\text{br }2\ i_{2}},\dots,\ x_{\text{br }q\ i_{q}}\right)-\left(\alpha_{0}+\sum_{j=1}^{q}\alpha_{j}x_{\text{br }j,\ i_{j}}^{2}+\sum_{j=1}^{q}\beta_{j}x_{\text{br }j,\ i_{j}}\right)\right)=0$$

or

$$\sum_{i_1=1}^{w} \dots \sum_{i_q=1}^{w} \left(\alpha_0 + \sum_{j=1}^{q} \alpha_j x_{\text{br}\ j,\ i_j}^2 + \sum_{j=1}^{q} \beta_j x_{\text{br}\ j,\ i_j} \right) = \sum_{i_1=1}^{w} \dots \sum_{i_q=1}^{w} \sigma \left(x_{\text{br}\ 1\ i_1}, \ x_{\text{br}\ 2\ i_2}, \dots, \ x_{\text{br}\ q\ i_q} \right)$$

Making simplifications in the left side of equation, we can get

$$\sum_{i_1=1}^{w} \dots \sum_{i_q=1}^{w} \alpha_0 + \sum_{i_1=1}^{w} \dots \sum_{i_q=1}^{w} \sum_{j=1}^{q} \alpha_j x_{\mathrm{br}\ j,\ i_j}^2 + \sum_{i_1=1}^{w} \dots \sum_{i_q=1}^{w} \sum_{j=1}^{q} \beta_j x_{\mathrm{br}\ j,\ i_j} = \sum_{i_1=1}^{w} \dots \sum_{i_q=1}^{w} \sigma\Big(x_{\mathrm{br}\ 1\ i_1},\ x_{\mathrm{br}\ 2\ i_2},\ \dots,\ x_{\mathrm{br}\ q\ i_q}\Big).$$

Taking into account that

$$\sum_{i_1=1}^{w} \dots \sum_{i_q=1}^{w} \alpha_0 = w^q,$$

$$\sum_{i_1=1}^{w} \dots \sum_{i_q=1}^{w} \sum_{j=1}^{q} \alpha_j x_{\mathrm{br}\ j,\ i_j}^2 = w^{q-1} \sum_{j=1}^{q} \sum_{i_j=1}^{w} \alpha_j x_{\mathrm{br}\ j,\ i_j}^2,$$

$$\sum_{i_1=1}^{w} \dots \sum_{i_q=1}^{w} \sum_{j=1}^{q} \beta_j x_{\mathrm{br}\ j,\ i_j} = w^{q-1} \sum_{j=1}^{q} \sum_{i_j=1}^{w} \beta_j x_{\mathrm{br}\ j,\ i_j},$$

w

w

the first equation can be presented as follows:

$$\alpha_0 w^q + \alpha_1 w^{q-1} \sum_{i_1=1}^w x_{\text{br }1 \ i_1}^2 + \beta_1 w^{q-1} \sum_{i_1=1}^w x_{\text{br }1 \ i_1} + \alpha_2 w^{q-1} \sum_{i_2=1}^w x_{\text{br }2 \ i_2}^2 + \beta_2 w^{q-1} \sum_{i_2=1}^w x_{\text{br }2 \ i_2} + \dots + \alpha_q w^{q-1} \sum_{i_q=1}^w x_{\text{br }q \ i_q}^2 + \beta_q w^{q-1} \sum_{i_q=1}^w x_{\text{br }q \ i_q} = \sum_{i_1=1}^w \dots \sum_{i_q=1}^w \sigma \Big(x_{\text{br }1 \ i_1}, \ x_{\text{br }2 \ i_2}, \dots, \ x_{\text{br }q \ i_q} \Big)$$

Similar simplifications can be made for other equations in the system. Therefore, the computation result for paraboloid (5) can be presented in the form of matrix equations

$$C = \Omega^{-1} \Psi, C = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \beta_1 \\ \cdots \\ \alpha_q \\ \beta_q \end{pmatrix}, \Psi = \begin{pmatrix} \sum_{i_1=1}^{w} \cdots \sum_{i_q=1}^{w} \sigma \left(x_{\text{br 1} i_1}, x_{\text{br 2} i_2}, \cdots, x_{\text{br q} i_q} \right) \\ \sum_{i_1=1}^{w} \cdots \sum_{i_q=1}^{w} x_{\text{br 1} i_1}^2 \sigma \left(x_{\text{br 1} i_1}, x_{\text{br 2} i_2}, \cdots, x_{\text{br q} i_q} \right) \\ \cdots \\ \sum_{i_1=1}^{w} \cdots \sum_{i_q=1}^{w} x_{\text{br 1} i_1} \sigma \left(x_{\text{br 1} i_1}, x_{\text{br 2} i_2}, \cdots, x_{\text{br q} i_q} \right) \\ \cdots \\ \sum_{i_1=1}^{w} \cdots \sum_{i_q=1}^{w} x_{\text{br q} i_q}^2 \sigma \left(x_{\text{br 1} i_1}, x_{\text{br 2} i_2}, \cdots, x_{\text{br q} i_q} \right) \\ \sum_{i_1=1}^{w} \cdots \sum_{i_q=1}^{w} x_{\text{br q} i_q} \sigma \left(x_{\text{br 1} i_1}, x_{\text{br 2} i_2}, \cdots, x_{\text{br q} i_q} \right) \end{pmatrix},$$

$$\Omega = \begin{pmatrix} w^{q} & w^{q-1} \sum_{i_{1}=1}^{w} x_{br \ 1 \ i_{1}} & w^{q-1} \sum_{i_{1}=1}^{w} x_{br \ 1 \ i_{1}} & w^{q-1} \sum_{i_{1}=1}^{w} x_{br \ 1 \ i_{1}} & \dots & w^{q-1} \sum_{i_{q}=1}^{w} x_{br \ q \ i_{q}} & w^{q-1} \sum_{i_{q}=1}^{w} x_{br \ q \ q} & w^{q-1} \sum_{i_{q}=1}^{w} x_{br \ q} & u^{q-1} \sum_{i_{q}=1}^{w} x_{br \ q}$$

6. Calculating the coordinates of paraboloid optimum. To obtain the minimum standard deviation, it is necessary to determine the coordinates of the minimum multidimen-

sional paraboloid. To do this, the partial derivatives are calculated and equated to zero [77]:

$$\frac{\frac{\partial \Sigma(x_{\text{br }1}, x_{\text{br }2}, \dots, x_{\text{br }q})}{\partial x_{\text{br }1}} = 0,$$

$$\frac{\frac{\partial \Sigma(x_{\text{br }1}, x_{\text{br }2}, \dots, x_{\text{br }q})}{\partial x_{\text{br }2}} = 0,$$

$$\dots$$

$$\frac{\partial \Sigma(x_{\text{br }1}, x_{\text{br }2}, \dots, x_{\text{br }q})}{\partial x_{\text{br }q}} = 0.$$

This system for general paraboloid (4) can be presented in the form of *q* linear equations system. For paraboloid (5), the solution of the system is given as

$$x_{\text{br }i \text{ opt}} = -\frac{\beta_i}{2\alpha_i}$$

7. Calculating the coefficients of the model for the optimal case. The coefficients of SLR, SQR, or SLQR are computed for the optimal location of the breakpoints using OLS. The final model can be used for the explanation and prediction of the response variable.

Consider the simple example for proposed method. Let us use the dataset with a small sample size presented in [6]. These data describe the relationship between production lot size x and the average production cost per unit y (in dollars) and are given in Table 1.

Table 1. Relationship between production lot size *x* and the average production cost per unit *y*.

N⁰	X	Y	N⁰	X	Ŷ	ււ	X	Ŷ
1	100	9.73	5	180	5.87	9	260	4.02
2	120	9.61	6	200	4.98	10	280	4.46
3	140	8.15	7	220	5.09	11	300	3.82
4	160	6.98	8	240	4.79			

Consider this step-by-step procedure.

- 1. To describe the presented data, the SLR model with q = 1 breakpoint is chosen.
- 2. The possible breakpoint abscissas values are $x_{br} = \{160; 180; 200; 220; 240\}$. Therefore, in this case x_{br} is a two-dimensional array with size 1×5 .
- 3. There are five alternative SLR models for all possible values in the array x_{br} :

$$\phi_{1,1}(X) = 14.446 - 0.0619X + 0.0408(X - 160)h(X - 160),$$

$$\phi_{1,2}(X) = 15.825 - 0.05601X + 0.0397(X - 180)h(X - 180),$$

$$p_{1,3}(X) = 15.117 - 0.0502X + 0.03885(X - 200)h(X - 200),$$

$$\phi_{1,4}(X) = 14.324 - 0.0443X + 0.0373(X - 220)h(X - 220),$$

$$\phi_{1,5}(X) = 13.713 - 0.04002X + 0.0391(X - 240)h(X - 240).$$

- 4. The standard deviations for the obtained SLR models are $\sigma = \{0.5; 0.362; 0.316; 0.435; 0.53\}$.
- 5. Because of one breakpoint, this multidimensional paraboloid converts into simple parabola. The result of the calculation is

$$\Sigma(x_{\text{br 1}}) = 4.41715 - 0.04446x_{\text{br 1}} + 1.128 \cdot 10^{-4} x_{\text{br 1}}^2.$$

6. The optimal value of the breakpoint abscissa is

$$x_{\text{br 1 opt}} = -\frac{\beta_1}{2\alpha_1} = 197.031.$$

7. The optimal SLR model is calculated for the obtained breakpoint abscissa. The final equation is

$$\phi_{1 \text{ opt}}(X) = 15.256 - 0.0513X + 0.03945(X - 197.031)h(X - 197.031)$$

The standard deviation for the optimal SLR model is 0.313. The result of the model building using the SLR model is shown in Figure 1.



Figure 1. Dataset and final optimal SLR model.

5. Analysis of Proposed Method Based on Statistical Simulation

The analysis of the proposed method is performed using statistical simulation and real data examples. This section presents the statistical simulation results. During the simulation, a dataset with two breakpoints is generated using build-in software operators. The dataset is an additive mixture of deterministic components and random noise.

Assume that the deterministic component corresponds to an SLR model

$$\phi_1(X, c'_{m, 1}, x_{br q, 1}) = c_{0, 1} + c_{1, 1}X + c_{2, 1}(X - x_{br 1})h(X - x_{br 1}) + c_{3, 1}(X - x_{br 2})h(X - x_{br 2})$$

The random noise is distributed according to the Gaussian probability density function. The initial data for the simulation are as follows:

- (1) Sample size n = 120;
- (2) Sampling time $\delta = 1$ (for discrete representation of the deterministic component);
- (3) Predetermined parameters of the SLR model: $c_{0,1} = 220$, $c_{1,1} = -3$, $c_{2,1} = 5$, $c_{3,1} = -4$, $x_{br 1} = 25$, and $x_{br 2} = 70$ (such parameters correspond, for example, to the real process of deterioration occurrence when monitoring the values of voltage for the supply of electronic devices [63]);
- (4) Predetermined parameters of Gaussian noise: the expected value is equal to zero and the standard deviation equal to 20 (additionally, it is assumed that the noise values are independent random variables for any sampling time moment);
- (5) The quantity of simulations reiteration N = 1000.

Consider the calculation procedure of the proposed method for one of the generated datasets. Table 2 shows one of the generated datasets. Figure 2 presents three realizations of the generated datasets, and each realization is marked by circle, triangle, or diamond (the circles correspond to the data in Table 2).

X	Ŷ	X	Ŷ	X	Ŷ	X	Ŷ
1	198.903	31	130.969	61	225.826	91	196.18
2	194.804	32	173.305	62	235.963	92	199.174
3	230.13	33	177.699	63	225.268	93	186.254
4	241.929	34	135.858	64	235.382	94	159.026
5	205.046	35	146.671	65	240.457	95	180.743
6	207.058	36	166.772	66	226.665	96	186.322
7	221.116	37	191.347	67	264.917	97	198.164
8	196.142	38	151.133	68	208.282	98	155.179
9	185.149	39	180.641	69	238.465	99	172.515
10	168.836	40	176.754	70	247.729	100	177.046
11	140.462	41	224.396	71	220.332	101	148.359
12	164.657	42	141.499	72	256.255	102	212.448
13	181.903	43	196.572	73	250.635	103	170.645
14	189.763	44	179.661	74	244.736	104	171.808
15	193.573	45	166.336	75	216.803	105	139.09
16	153.473	46	175.007	76	225.626	106	162.471
17	173.226	47	192.467	77	241.812	107	178.876
18	173.601	48	162.748	78	248.445	108	147.776
19	164.416	49	175.224	79	228.271	109	168.604
20	144.779	50	208.317	80	193.772	110	177.17
21	165.477	51	179.932	81	215.457	111	151.077
22	156.022	52	202.743	82	209.727	112	153.421
23	191.786	53	183.889	83	223.962	113	125.35
24	124.953	54	182.191	84	202.548	114	135.484
25	144.006	55	204.996	85	206.732	115	152.601
26	181.289	56	212.034	86	238.368	116	111.133
27	131.828	57	192.96	87	214.105	117	131.803
28	148.114	58	240.106	88	204.29	118	142.927
29	189.118	59	230.511	89	185.714	119	151.265
30	159.11	60	188.666	90	184.075	120	145.401

Table 2. Example of obtained dataset.

To describe the obtained dataset, we choose the SLR model with three segments with q = 2 breakpoints. To simplify the calculations, we choose the quantity of discrete values within the range of possible breakpoints to be w = 5. According to the geometrical structure of the observed dataset (Figure 2), the ranges for two breakpoints are as follows:

$$x_{\text{br }1} = \{15; 20; 25; 30; 35\},$$

 $x_{\text{br }2} = \{60; 65; 70; 75; 80\}.$

The next step is to evaluate the unknown coefficients $c_{0, 1}$, $c_{1, 1}$, $c_{2, 1}$, and $c_{3, 1}$ for all possible values of the first and second breakpoints using OLS. As a result, 25 alternative SLR models are obtained.



Figure 2. Examples of generated additive mixture of SLR model and Gaussian noise.

After that, the standard deviations between the model output and the observed data for these SLR models are determined. Table 3 shows the computation results.

Table 3. Computation results for standard deviation	itation results for standard deviation
--	--

	Abscissa of the Second Breakpoint							
Standard Deviation	$x_{br2} = 60$	$x_{br2} = 65$	$x_{\rm br2}=70$	$x_{br2} = 75$	$x_{br2} = 80$			
	$x_{\rm br1} = 15$	22.947	21.042	19.928	19.911	20.948		
	$x_{\rm br1} = 20$	21.692	19.813	18.879	19.166	20.534		
The abscissa of the first breakpoint	$x_{\rm br1} = 25$	20.912	19.145	18.454	19.041	20.667		
breakpoint	$x_{\rm br1} = 30$	20.637	19.053	18.622	19.454	21.229		
	$x_{\rm br1} = 35$	20.818	19.443	19.249	20.239	22.051		

Even visual analysis of the data on the standard deviation (Table 3) indicates that the minimal standard deviation is located approximately near $x_{br 1} = 25$ and $x_{br 2} = 70$. To estimate the exact values of breakpoint abscissas, paraboloids (4) and (5) are built using OLS.

After the calculations, the following mathematical equations were obtained:

$$\Sigma(x_{\rm br\ 1}, \ x_{\rm br\ 2}) = 148.8 - 1.1295x_{\rm br\ 1} - 3.2925x_{\rm br\ 2} + 0.01008x_{\rm br\ 1}^2 + 0.02188x_{\rm br\ 2}^2 + 8.534 \cdot 10^{-3}x_{\rm br\ 1}x_{\rm br\ 2},$$

$$\Sigma(x_{\rm br\ 1}, x_{\rm br\ 2}) = 133.87 - 0.5321x_{\rm br\ 1} - 3.0791x_{\rm br\ 2} + 0.01008x_{\rm br\ 1}^2 + 0.02188x_{\rm br\ 2}^2.$$

Figures 3 and 4 show the visual presentation of paraboloids (4) and (5) for this numerical example, respectively.



Figure 3. General paraboloid (4) for presented example.



Figure 4. Simplified paraboloid (5) for presented example.

To determine the optimum coordinates for three-dimensional general paraboloid (4), it is necessary to solve the following system of two linear equations:

$$\begin{cases} \frac{\partial \Sigma(x_{\rm br\,1}, x_{\rm br\,2})}{\partial x_{\rm br\,1}} = 0,\\ \frac{\partial \Sigma(x_{\rm br\,1}, x_{\rm br\,2})}{\partial x_{\rm br\,2}} = 0. \end{cases}$$

In this case, the calculation gives the following solution:

$$x_{\text{br 1 opt}} = \frac{\beta_2 \gamma_{1,2} - 2\alpha_2 \beta_1}{4\alpha_1 \alpha_2 - \gamma_{1,2}^2},$$

$$x_{\text{br }2 \text{ opt}} = \frac{\beta_1 + 2\alpha_1 x_{\text{br }1 \text{ opt}}}{\gamma_{1,2}}$$

The general paraboloid (4) has a minimum standard deviation at the coordinates

$$x_{br \ 1 \ opt}^{(gen)} = 26.361,$$

 $x_{br \ 2 \ opt}^{(gen)} = 70.086.$

The simplified paraboloid (5) has a minimum standard deviation at the coordinates

$$x_{br \ 1 \ opt}^{(sim)} = 26.397,$$

 $x_{br \ 2 \ opt}^{(sim)} = 70.351$

The results of the calculation for paraboloids (4) and (5) almost coincide. The relative error for the first and second breakpoint abscissa is equal to 5.558% and 0.5014%, respectively.

After the calculation of the model's coefficients for the optimal case, the optimal SLR models for paraboloids (4) and (5) are obtained:

$$\phi_1(X, \vec{c}_{m,1}, x_{\text{br} q,1}) = 212.169 - 2.526X + 4.711(X - 26.361)h(X - 26.361) - -4.440(X - 70.086)h(X - 70.086),$$

$$\phi_1(X, \vec{c}_{m,1}, x_{\text{br }q,1}) = 212.022 - 2.509X + 4.678(X - 26.397)h(X - 26.397) - -4.443(X - 70.351)h(X - 70.351).$$

The obtained SLR models give almost the same standard deviations equal to 18.429 and 18.424, respectively.

Figure 5 shows the generated dataset and final optimal SLR models. Visual analysis shows the coincidence of both SLR models.



Figure 5. Generated dataset and final optimal SLR models.

We consider the general simulation results for all iterations. Repeating the simulation provides an opportunity to perform a complete statistical analysis of the breakpoint estimation during mathematical model building. An analysis was performed by plotting histograms and evaluating the numerical characteristics of the random variables. Figure 5 shows the histograms for the estimate of two breakpoint abscissas and the usage of different optimization options (general and simplified paraboloids). The parameter λ in Figure 6



is the quantity of breakpoint abscissa estimates, which are located in the corresponding grouping interval of the histogram.

Figure 6. Histograms for breakpoint abscissas for different optimization options: (**A**)—estimates of the first breakpoint for general paraboloid; (**B**)—estimates of the second breakpoint for general paraboloid; (**C**)—estimates of the first breakpoint for simplified paraboloid; (**D**)—estimates of the second breakpoint for simplified paraboloid.

Table 4 shows the numerical characteristics of the breakpoint abscissas estimates (mathematical expectation, standard deviation, range of change, and skewness).

Statistical Characteristic	General Paraboloid	Simplified Paraboloid
Mathematical expectation of x_{br1}	25.746	25.753
Standard deviation for x_{br1}	2.218	2.13
Minimum of <i>x</i> _{br1}	17.158	17.07
Maximum of x_{br1}	36.143	36.306
Skewness for x_{br1}	-0.055	0.054
Mathematical expectation of x_{br2}	70.113	70.35
Standard deviation for x_{br2}	2.397	2.274
Minimum of x_{br2}	62.138	63.486
Maximum of <i>x</i> _{br2}	84.242	83.238
Skewness for x_{br2}	0.583	0.536

 Table 4. Numerical characteristics of breakpoint abscissas estimates.

To describe the obtained estimates of breakpoint abscissa completely, it is necessary to fit the histogram by theoretical probability density function. Approximate assumptions can be made based on the graphical view of the histograms in Figure 6. The shape of the histogram can correspond to the Gaussian probability density function. Such an assumption can be proven using the chi-squared test with high confidence probability.

The breakpoint estimation bias has preferable values when the general paraboloid method is used. However, the benefit is negligible and averages 0.337% compared with the

simplified paraboloid method. The highest percentage of estimate bias (in relative values) is 3.012%. In the case of a long-term breakpoint, the simplified paraboloid method has, on average, a narrower range of change of breakpoint estimates.

Let us analyze the proposed method in comparison with the method of simple enumeration. To obtain the approximately 3% of breakpoint abscissas estimate bias, the method of simple enumeration requires at least 33 possible values for each breakpoint. Therefore, it is necessary to repeat computations for at least 1089 iterations in the case of two breakpoints. At the same time, the proposed method requires 25 iterations and additional calculations of the paraboloid optimum. Therefore, the proposed method reduces the computing time by at least 30 times compared to the method of simple enumeration.

A comparison of the simulation results for a range of initial data provides the ability to conclude approximately the same accuracy characteristics for SLR models based on general and simplified paraboloid usage. Therefore, in practical cases, the adoption of the simplified paraboloid method usage is more advantageous when creating a segmented regression model because of the reduction in computations and calculation time.

6. Real Data Example

Consider the example of real data on the number of earthquakes with a magnitude of 7 or higher by year, according to the United States Geological Survey [78]. Table 5 presents the corresponding data from 1922 to 2021.

i	X	Ŷ	i	X	Ŷ	i	X	Ŷ	i	X	Ŷ
1	1922	6	26	1947	13	51	1972	16	76	1997	16
2	1923	16	27	1948	12	52	1973	9	77	1998	12
3	1924	8	28	1949	13	53	1974	11	78	1999	18
4	1925	10	29	1950	8	54	1975	13	79	2000	15
5	1926	9	30	1951	6	55	1976	14	80	2001	16
6	1927	12	31	1952	9	56	1977	11	81	2002	13
7	1928	18	32	1953	6	57	1978	12	82	2003	15
8	1929	14	33	1954	9	58	1979	8	83	2004	16
9	1930	4	34	1955	5	59	1980	6	84	2005	11
10	1931	17	35	1956	19	60	1981	10	85	2006	11
11	1932	7	36	1957	7	61	1982	8	86	2007	18
12	1933	8	37	1958	6	62	1983	14	87	2008	12
13	1934	12	38	1959	13	63	1984	14	88	2009	17
14	1935	13	39	1960	11	64	1985	15	89	2010	24
15	1936	9	40	1961	9	65	1986	11	90	2011	20
16	1937	9	41	1962	17	66	1987	13	91	2012	16
17	1938	23	42	1963	9	67	1988	11	92	2013	19
18	1939	14	43	1964	15	68	1989	8	93	2014	12
19	1940	8	44	1965	8	69	1990	18	94	2015	19
20	1941	11	45	1966	10	70	1991	17	95	2016	16
21	1942	13	46	1967	20	71	1992	13	96	2017	7
22	1943	17	47	1968	14	72	1993	12	97	2018	17
23	1944	12	48	1969	15	73	1994	13	98	2019	10
24	1945	7	49	1970	15	74	1995	20	99	2020	9
25	1946	12	50	1971	13	75	1996	15	100	2021	19

Table 5. Quantity of earthquakes of magnitude 7 or higher by year.



Table 5 contains data observed from 1922 to 2021, where i is the number of observations, X is the year, and Y is the quantity of earthquakes.



To simplify the presentation and calculations, the first year of observation (1922) is assigned a zero point at the abscissa axis in the next computations. Thus, to return to the original data, it is necessary to add 1922 for the shifted abscissa axis.

According to the visual analysis of the dataset, let us assume that there are five breakpoints in this realization. The following are the ranges for these breakpoints:

$$x_{br 1} = \{17; 18; 19; 20; 21\},$$

$$x_{br 2} = \{35; 36; 37; 38; 39\},$$

$$x_{br 3} = \{46; 47; 48; 49; 50\},$$

$$x_{br 4} = \{55; 56; 57; 58; 59\} \text{ and }$$

$$x_{br 5} = \{82; 83; 84; 85; 86\}.$$

With such a range of variables, 3125 different SLR model options are available. The standard deviation is calculated for each case. As a result, a six-dimensional array $\sigma(x_{br 1}, x_{br 2}, x_{br 3}, x_{br 4}, x_{br 5})$ is generated. To approximate the obtained data, OLS is used on a six-dimensional optimization paraboloid. For simplicity, we used a simplified paraboloid as follows:

$$\begin{split} \Sigma(x_{\mathrm{br}\ 1}, \ x_{\mathrm{br}\ 2}, \ x_{\mathrm{br}\ 3}, \ x_{\mathrm{br}\ 4}, \ x_{\mathrm{br}\ 5}) &= -55431 - 3.085 x_{\mathrm{br}\ 1} + 4.832 x_{\mathrm{br}\ 2} + 79.227 x_{\mathrm{br}\ 3} + 33.495 x_{\mathrm{br}\ 4} \\ &+ 1251 x_{\mathrm{br}\ 5} + 0.0811 x_{\mathrm{br}\ 1}^2 - 0.0651 x_{\mathrm{br}\ 2}^2 - 0.8251 x_{\mathrm{br}\ 3}^2 - 0.2937 x_{\mathrm{br}\ 4}^2 - 7.445 x_{\mathrm{br}\ 5}^2. \end{split}$$

This simplified paraboloid has optimum standard deviation at the coordinates

$$\begin{aligned} x_{\text{br 1 opt}}^{(\text{sim})} &= 19.01, \ x_{\text{br 2 opt}}^{(\text{sim})} = 37.086, \ x_{\text{br 3 opt}}^{(\text{sim})} = 48.008, \\ x_{\text{br 4 opt}}^{(\text{sim})} &= 57.019, \ x_{\text{br 5 opt}}^{(\text{sim})} = 84.005. \end{aligned}$$

After the calculation of the model's coefficients for the optimal case of breakpoint locations using OLS, the final SLR model is obtained:

$$\phi_1(X, \vec{c}_{m, 1}, x_{\text{br } q, 1}) = 10.197 + 0.1356X - 0.3553(X - 19.01)h(X - 19.01) \\ + 0.8002(X - 37.086)h(X - 37.086) - 1.1212(X - 48.008)h(X - 48.008) \\ + 0.7777(X - 57.019)h(X - 57.019) - 0.4252(X - 48.008)h(X - 84.005).$$



The standard deviation for the obtained SLR model is equal to 3.799. Figure 8 shows the observed dataset and the final optimal SLR model.

Figure 8. Observed dataset and final optimal SLR model.

The method of simple enumeration for a given dataset gives approximately the same result as that shown in Figure 8. However, this method increases the computing time approximately twice. Polynomial regression using a seventh-order polynomial is characterized by a faster computation time; however, it gives unacceptable predictive properties.

The results of the mathematical model building can be used for solving prediction problems. Consider this problem for the observed dataset based on generally known results that have been extensively described in the literature (for example, [76,77]) and innovative methods that may be used in accordance with the properties of segmented regression models.

To predict the future trend, let us determine the range of the SLR model change. For this purpose, we used a straight line and OLS to approximate the upper and lower ordinates of the breakpoints. The lower line contains the zero point, and the second and fourth breakpoints. The upper line contains the first, third, and fifth breakpoints. The numerical values of the calculated equations are

$$Y^{(\text{lower})}(X) = 9.871 - 3.606 \cdot 10^{-3} X, \ Y^{(\text{upper})}(X) = 11.881 + 0.0592 X$$



The last segment of the SLR model is continued to the intersection point with the lower straight line. Figure 9 shows the visual representation of the trend prediction.

Figure 9. Prediction-based SLR model.

This method of prediction and the obtained SLR model allow us to anticipate that, through 2042, the average annual number of earthquakes with a magnitude of 7 or higher would decrease.

In general, the proposed method can be applied to different datasets and, in the case of using multidimensional optimization, to determine breakpoints.

7. Conclusions

This article presents a method of accuracy increment when segmented regression is used. The main problem for segmented regression model building is the estimation of the coordinates of the breakpoint between adjacent segments. To solve this problem, two types of multidimensional optimization paraboloids are used. The paraboloids contain information on standard deviations between the model output and the observed data for different sets of possible values of breakpoint abscissas. The minimum standard deviation of each paraboloid coincided with the optimal position of the breakpoints.

A step-by-step procedure for the proposed method was described by examples based on statistical simulation and real data observation.

Generally, the use of SLR, SQR, and SLQR models provides a mathematical model with high accuracy, more accurately describes the geometrical structure of the analyzed dataset, and has good predictive properties.

The results of this research can be used during mathematical model building for statistical data obtained in various branches of human activity.

Author Contributions: Conceptualization, J.A.-A. and M.Z.; methodology, J.A.-A. and V.K.; software, R.O.; validation, J.A.-A., A.M. and M.Z.; formal analysis, R.O.; investigation, J.A.-A.; resources, A.M.; data curation, M.Z.; writing—original draft preparation, J.A.-A. and M.Z.; writing—review and editing, J.A.-A. and M.Z.; visualization, R.O.; supervision, V.K.; project administration, J.A.-A.; funding acquisition, J.A.-A. and A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data used in the study are obtained in two ways: (1) using random number generator during statistical simulation; (2) using open source from The United States Geological Survey with citing corresponding link.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

- OLS Ordinary least squares
- SLQR Segmented linear-quadratic regression
- SLR Segmented linear regression
- SQR Segmented quadratic regression

References

- 1. Williams, H.P. Model Building in Mathematical Programming, 5th ed.; John Wiley & Sons: New York, NY, USA, 2013.
- Banwarth-Kuhn, M.; Sindi, S. How and why to build a mathematical model: A case study using prion aggregation. *J. Biol. Chem.* 2020, 295, 5022–5035. [CrossRef] [PubMed]
- Neumaier, A. Mathematical model building. In *Modeling Languages in Mathematical Optimization;* Applied Optimization; Kallrath, J., Ed.; Springer: Boston, MA, USA, 2004; Volume 88.
- 4. Bodner, K.; Fortin, M.-J.; Molnár, P.K. Making predictive modelling ART: Accurate, reliable, and transparent. *Ecosphere* 2020, *11*, e03160. [CrossRef]
- Ostroumov, I.V.; Kuzmenko, N.S. Accuracy estimation of alternative positioning in navigation. In Proceedings of the 4th IEEE International Conference on Methods and Systems of Navigation and Motion Control, Kiev, Ukraine, 18–20 October 2016; pp. 291–294.
- 6. Montgomery, D.C.; Peck, E.A.; Vining, G.G. *Introduction to Linear Regression Analysis*, 5th ed.; John Wiley & Sons: Hoboken, NJ, USA, 2012.

- 7. Zaliskyi, M.; Petrova, Y.; Asanov, M.; Bekirov, E. Statistical data processing during wind generators operation. *Int. J. Electr. Electron. Eng. Telecommun.* **2019**, *8*, 33–38. [CrossRef]
- 8. Giloni, A.; Padberg, M. Alternative methods of linear regression. Math. Comput. Model. 2002, 35, 361–374. [CrossRef]
- Kaufman, R. *Heteroskedasticity in Regression: Detection and Correction;* Sage Publications: Los Angeles, CA, USA, 2013. [CrossRef]
 Hassan, M.: Hossny, M.: Nahavandi, S.: Creighton, D. Heteroskedasticity variance index. In Proceedings of the 14th International
- Hassan, M.; Hossny, M.; Nahavandi, S.; Creighton, D. Heteroskedasticity variance index. In Proceedings of the 14th International Conference on Computer Modelling and Simulation, Cambridge, UK, 28–30 March 2012; Volume 2012, pp. 135–141.
- Cheng, F. Maximum deviation of error density estimators in censored linear regression. *Stat. Probab. Lett.* 2012, 82, 1657–1664. [CrossRef]
- 12. Ostroumov, I.; Kuzmenko, N.S. Accuracy assessment of aircraft positioning by multiple radio navigational AIDS. *Telecommun. Radio Eng.* **2018**, *77*, 705–715. [CrossRef]
- 13. Weisberg, S. Applied Linear Regression; John Wiley and Sons: New York, NY, USA, 2005.
- 14. Prokopenko, I.; Omelchuk, I.; Maloyed, M. Synthesis of signal detection algorithms under conditions of aprioristic uncertainty. In Proceedings of the 2020 IEEE Ukrainian Microwave Week, Kharkiv, Ukraine, 21–25 September 2020; pp. 418–423. [CrossRef]
- 15. Prokopenko, I. Robust methods and algorithms of signal processing. In Proceedings of the IEEE Microwaves, Radar and Remote Sensing Symposium, Kyiv, Ukraine, 29–31 August 2017; Volume 2017, pp. 71–74. [CrossRef]
- 16. Walker, W.; Harremoës, P.; Rotmans, J.; van der Sluijs, J.P.; Van Asselt, M.; Janssen, P.; Von Krauss, M.K. Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integr. Assess.* **2003**, *4*, 5–17. [CrossRef]
- 17. van Oijen, M. Bayesian methods for quantifying and reducing uncertainty and error in forest models. *Curr. For. Rep.* 2017, *3*, 269–280. [CrossRef]
- 18. Rawlings, J.O.; Pantula, S.G.; Dickey, D.A. *Applied Regression Analysis: A Research Tool*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2001.
- 19. Nachtsheim, C.; Neter, J.; Kutner, M.; Wasserman, W. *Applied Linear Statistical Models*, 5th ed.; McGraw-Hill: New York, NY, USA, 2005.
- 20. Atkinson, A.; Riani, M. Robust Diagnostic Regression Analysis; Springer: New York, NY, USA, 2000. [CrossRef]
- Sano, Y.; Kandori, A.; Miyoshi, T.; Tsuji, T.; Shima, K.; Yokoe, M.; Sakoda, S. Severity estimation of finger-tapping caused by Parkinson's disease by using linear discriminant regression analysis. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012; Volume 2012, pp. 4315–4318.
- Jin, R.; Si, L.; Srivastava, S.; Li, Z.; Chan, C. A knowledge driven regression model for gene expression and microarray analysis. In Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, USA, 30 August–3 September 2006; Volume 2006, pp. 5326–5329.
- 23. Jinyu, T.; Xin, Z. Apply multiple linear regression model to predict the audit opinion. In Proceedings of the ISECS International Colloquium on Computing, Communication, Control, and Management, Sanya, China, 8–9 August 2009; pp. 303–306. [CrossRef]
- Yan, X.; Du, X.; Jiang, Y. Research on the relationship between environmental pollution and economy. In Proceedings of the International Conference on Mechanic Automation and Control Engineering, Wuhan, China, 26–28 June 2010; Volume 2010, pp. 1809–1812. [CrossRef]
- 25. Ding, Y.; Yuechao, D. A new method multi-factor trend regression and its application to economy forecast in Jiangxi. *First Int. Workshop Knowl. Discov. Data Min.* **2008**, 2008, 63–67. [CrossRef]
- Yun-Ning, Z.; Qian, Y. Regression analysis of the real estate investment and economy increased. In Proceedings of the Third International Conference on Intelligent System Design and Engineering Applications, Hong Kong, China, 16–18 January 2013; Volume 2013, pp. 1099–1101.
- 27. Goncharenko, A. A multi-optional hybrid functions entropy as a tool for transportation means repair optimal periodicity determination. *Aviation* **2018**, *22*, 60–66. [CrossRef]
- 28. Goncharenko, A. aircraft operation depending upon the uncertainty of maintenance alternatives. *Aviation* **2017**, *21*, 126–131. [CrossRef]
- Dyvak, M.; Darmorost, I.; Shevchuk, R.; Manzhula, V.; Kasatkina, N. Correlation analysis traffic intensity of the motor vehicles and the air pollution by their harmful emissions. In Proceedings of the 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), Lviv-Slavske, Ukraine, 20–24 February 2018; pp. 855–858. [CrossRef]
- Li, Z.; Zhang, J.; Zhang, X. A dynamic model for aircraft route optimizing in airport surface management. In Proceedings of the 9th International Conference on Electronic Measurement & Instruments, Beijing, China, 16–19 August 2009; pp. 3-1068–3-1072. [CrossRef]
- Ostroumov, I.; Marais, K.; Kuzmenko, N. Aircraft positioning using multiple distance measurements and spline prediction. *Aviation* 2022, 26, 1–10. [CrossRef]
- Kuzmenko, N.S.; Kharchenko, V.P.; Ostroumov, I.V. Identification of unmanned aerial vehicle flight situation. In Proceedings of the IEEE 5th International Conference on Actual problems of Unmanned Aerial Vehicles Development, Kiev, Ukraine, 17–19 October 2017; pp. 116–120.

- Bezkorovainyi, Y.N.; Sushchenko, O.A. Improvement of UAV positioning by information of inertial sensors. In Proceedings of the IEEE 5th International Conference on Methods and Systems of Navigation and Motion Control, Kyiv, Ukraine, 16–18 October 2018; pp. 123–126.
- Shmelova, T.; Sikirda, Y.; Kasatkin, M. Modeling of the collaborative decision making by remote pilot and air traffic controller in flight emergencies. In Proceedings of the 5th International Conference on Actual Problems of Unmanned Aerial Vehicles Developments, Kyiv, Ukraine, 22–24 October 2019; Volume 2019, pp. 230–233. [CrossRef]
- 35. Shmelova, T.; Sechko, O. Application artificial intelligence for real-time monitoring, diagnostics, and correction human state. *CEUR Workshop Proc.* **2019**, *2488*, 185–194.
- Prokopenko, I.G.; Migel, S.V.; Prokopenko, K.I. Signal modeling for the efficient target detection tasks. *Int. Radar Symp.* 2013, 2, 976–982.
- 37. Averyanova, Y.; Averyanov, A.; Yanovsky, F. Polarization signal components estimate in weather radar. In Proceedings of the 12th International Conference on Mathematical Methods in Electromagnetic Theory, Odessa, Ukraine, 29 June–2 July 2008; pp. 360–362.
- Averyanova, Y.; Averyanov, A.; Yanovsky, F. The approach to estimating critical wind speed in liquid precipitation using radar polarimetry. In Proceedings of the International Conference on Mathematical Methods in Electromagnetic Theory, Kharkiv, Ukraine, 28–30 August 2012; pp. 517–520. [CrossRef]
- Yanovsky, F.J.; Prokopenko, I.G.; Prokopenko, K.I.; Russchenberg, H.W.J.; Ligthart, L.P. Radar estimation of turbulence eddy dissipation rate in rain. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Toronto, ON, Canada, 24–28 June 2002; pp. 63–65. [CrossRef]
- 40. Ostroumov, I.; Kuzmenko, N. Configuration analysis of European navigational aids network. In Proceedings of the 2021 Integrated Communications Navigation and Surveillance Conference, Virtual, 20 April–22 April 2021; pp. 1–9. [CrossRef]
- Ostroumov, I.V.; Kuzmenko, N.S.; Marais, K. Optimal pair of navigational Aids selection. In Proceedings of the IEEE 5th International Conference on Methods and Systems of Navigation and Motion Control, Kyiv, Ukraine, 16–18 October 2018; pp. 32–35.
- 42. Ostroumov, I.; Kuzmenko, N.S. Accuracy improvement of vor/vor navigation with angle extrapolation by linear regression. *Telecommun. Radio Eng.* **2019**, *78*, 1399–1412. [CrossRef]
- 43. Solomentsev, O.; Zaliskyi, M. Correlated failures analysis in navigation system. In Proceedings of the IEEE 5th International Conference on Methods and Systems of Navigation and Motion Control, Kyiv, Ukraine, 16–18 October 2018; pp. 41–44.
- 44. Ostroumov, I.; Kuzmenko, N. Compatibility analysis of multi signal processing in apnt with current navigation infrastructure. *Telecommun. Radio Eng.* **2018**, *77*, 211–223. [CrossRef]
- 45. Dyvak, M.; Melnyk, A.; Kovbasistyi, A.; Shevchuk, R.; Huhul, O.; Tymchyshyn, V. Mathematical modeling of the estimation process of functioning efficiency level of information web-resources. In Proceedings of the 10th International Conference on Advanced Computer Information Technologies, Deggendorf, Germany, 13–15 May 2020; pp. 492–496. [CrossRef]
- Zaliskyi, M.; Odarchenko, R.; Gnatyuk, S.; Petrova, Y.; Chaplits, A. Method of traffic monitoring for DDoS attacks detection in e-health systems and networks. *CEUR Workshop Proc.* 2018, 2255, 193–204.
- Gnatyuk, S. Critical aviation information systems cybersecurity, meeting security challenges through data analytics and decision support. In NATO Science for Peace and Security Series, D: Information and Communication Security; IOS Press Ebooks: Amsterdam, The Netherlands, 2016; Volume 47, pp. 308–316.
- Hu, Z.; Odarchenko, R.; Gnatyuk, S.; Zaliskyi, M.; Chaplits, A.; Bondar, S.; Borovik, V. Statistical techniques for detecting cyberattacks on computer networks based on an analysis of abnormal traffic behavior. *Int. J. Comput. Netw. Inf. Secur.* 2021, 12, 1–13. [CrossRef]
- Kalimoldayev, M.; Tynymbayev, S.; Gnatyuk, S.; Ibraimov, M.; Magzom, M. The device for multiplying polynomials modulo an irreducible polynomial, news of the national academy of sciences of the republic of Kazakhstan. *Ser. Geol. Tech. Sci.* 2019, 2, 199–205. [CrossRef]
- 50. Gnatyuk, S.; Akhmetov, B.; Kozlovskyi, V.; Kinzeryavyy, V.; Aleksander, M.; Prysiazhnyi, D. New secure block cipher for critical applications: Design, implementation, speed and security analysis. *Adv. Intell. Syst. Comput.* **2020**, *1126*, 93–104.
- Volianskyi, R.; Sadovoi, O.; Volianska, N.; Sinkevych, O. Construction of parallel piecewise-linear interval models for nonlinear dynamical objects. In Proceedings of the 2019 9th International Conference on Advanced Computer Information Technologies (ACIT), Ceske Budejovice, Czech Republic, 5–7 June 2019; pp. 97–100. [CrossRef]
- Chikovani, V.V.; Suschenko, O.A. Differential mode of operation for Coriolis vibratory gyro with ring-like resonator. In Proceedings of the IEEE 34th International Scientific Conference on Electronics and Nanotechnology, Kyiv, Ukraine, 15–18 April 2014; pp. 451–455. [CrossRef]
- Solomentsev, O.; Zaliskyi, M. Method of sequential estimation of statistical distribution parameters in control systems design. In Proceedings of the IEEE 3rd International Conference Methods and Systems of Navigation and Motion Control, Kyiv, Ukraine, 14–17 October 2014; pp. 135–138. [CrossRef]
- Sushchenko, O.; Bezkorovayniy, Y.; Golytsin, V. Processing of redundant information in airborne electronic systems by means of neural networks. In Proceedings of the IEEE 39th International Conference on Electronics and Nanotechnology, Kyiv, Ukraine, 16–18 April 2019; pp. 652–655. [CrossRef]
- Chikovani, V.; Sushchenko, O.; Tsiruk, H. Redundant information processing techniques comparison for differential vibratory gyroscope. *East. Eur. J. Enterp. Technol.* 2016, 4, 45–52. [CrossRef]

- Solomentsev, O.V.; Zaliskyi, M.U.; Zuiev, O.V.; Asanov, M.M. Data processing in exploitation system of unmanned aerial vehicles radioelectronic equipment. In Proceedings of the IEEE 2nd International Conference on Actual Problems of Unmanned Air Vehicles Developments, Kyiv, Ukraine, 15–17 October 2013; pp. 77–80. [CrossRef]
- 57. Goncharenko, A.V. Optimal UAV maintenance periodicity obtained on the multi-optional basis. In Proceedings of the IEEE 4th International Conference on Actual Problems of UAV Developments, Kyiv, Ukraine, 17–19 October 2017; pp. 65–68. [CrossRef]
- 58. Goncharenko, A. Development of a theoretical approach to the conditional optimization of aircraft maintenance preference uncertainty. *Aviation* **2018**, *22*, 40–44. [CrossRef]
- Solomentsev, O.; Zaliskyi, M.; Zuiev, O. Radioelectronic equipment availability factor models. In Proceedings of the 2013 Signal Processing Symposium (SPS), Serock, Poland, 5–7 June 2013; pp. 1–3. [CrossRef]
- 60. Dhillon, B.S. Reliability, Quality, and Safety for Engineers; CRC Press: Boca Raton, FL, USA, 2005; p. 216.
- 61. Nakagawa, T. Maintenance Theory of Reliability; Springer: London, UK, 2005; p. 270.
- 62. Ulansky, V.; Terentyeva, I. Availability assessment of a telecommunications system with permanent and intermittent faults. In Proceedings of the IEEE First Ukraine Conference on Electrical and Computer Engineering, Kyiv, Ukraine, 29 May–2 June 2017; pp. 908–911. [CrossRef]
- Taranenko, A.G.; Gabrousenko, Y.I.; Holubnychyi, A.G.; Slipukhina, I.A. Estimation of redundant radionavigation system reliability. In Proceedings of the IEEE 5th International Conference on Methods and Systems of Navigation and Motion Control, Kiev, Ukraine, 16–18 October 2018; pp. 28–31.
- Solomentsev, O.; Zaliskyi, M.; Nemyrovets, Y.; Asanov, M. Signal processing in case of radio equipment technical state deterioration. In Proceedings of the 2015 Signal Processing Symposium (SPSympo), Debe Village, Poland, 10–12 June 2015; pp. 1–5. [CrossRef]
- 65. Seber, G.A.F.; Wild, C.J. Nonlinear Regression; John Wiley and Sons: New York, NY, USA, 2003; p. 768.
- 66. Bates, D.M.; Watts, D.G. Nonlinear Regression Analysis and Its Applications; John Wiley and Sons: New York, NY, USA, 1988; p. 366.
- 67. Knuepling, F.; Allen, J. Testing between different types of switching regression models. J. Econ. Econom. Econ. Econom. Soc. 2015, 58, 30–63.
- 68. Huet, S.; Bouvier, A.; Poursat, M.-A.; Jolivet, E. Statistical tools for nonlinear regression. In *A Practical Guide with S-PLUS and R Examples*; Springer: New York, NY, USA, 2004; p. 232.
- 69. Tishler, A.; Zang, I. A new maximum likelihood algorithm for piecewise regression. *J. Am. Stat. Assoc.* **1981**, *76*, 980–987. [CrossRef]
- 70. Buteikis, A. Practical Econometrics and Data Science; Vilnius University: Vilnius, Lithuania, 2020.
- 71. Carlin, B.P.; Gelfand, A.E.; Smith, A.F.M. Hierarchical bayesian analysis of changepoint problems. J. R. Stat. Soc. Ser. C Applied Stat. 1992, 41, 389–405. [CrossRef]
- 72. Ferreira, P.E. A bayesian analysis of a switching regression model: Known number of regimes. J. Am. Stat. Assoc. 1975, 70, 370–374. [CrossRef]
- 73. Toms, J.D.; Lesperance, M.L. Piecewise regression: A tool for identifying ecological thresholds. *Ecology* **2003**, *84*, 2034–2041. [CrossRef]
- Solomentsev, O.; Zaliskyi, M.; Shcherbyna, O.; Kozhokhina, O. Sequential procedure of changepoint analysis during operational data processing. In Proceedings of the 2020 IEEE Microwave Theory and Techniques in Wireless Communications (MTTW), Riga, Latvia, 1–2 October 2020; Volume 1, pp. 168–171. [CrossRef]
- Solomentsev, O.; Zaliskyi, M.; Herasymenko, T.; Petrova, Y. Data processing method for deterioration detection during radio equipment operation. In Proceedings of the 2019 IEEE Microwave Theory and Techniques in Wireless Communications (MTTW), Riga, Latvia, 1–2 October 2019; Volume 1, pp. 1–4. [CrossRef]
- 76. Nocedal, J.; Wright, S.J. Numerical Optimization, 2nd ed.; Springer: New York, NY, USA, 2006.
- Reklaitis, G.V.; Ravindran, A.; Ragsdell, K.M. Engineering Optimization, Methods & Applications; John Wiley and Sons: New York, NY, USA, 1983; p. 688.
- 78. The United States Geological Survey. Available online: https://earthquake.usgs.gov (accessed on 1 July 2022).