

Computational methods for best mathematical model building using a new adequacy criterion

Valerii Kuzmin^{1,†}, Yuliia Petrova^{1,*,†} and Alina Osipchuk^{1,†}

¹National University "Kyiv Aviation Institute", Liubomyra Huzara Ave., 1, Kyiv, 03058, Ukraine

Abstract

The article considers the problem of approximating experimental data with a heterogeneous structure, in particular in cases where several observed values are in the same cross-section. Such a distribution of data can reduce the accuracy and reliability of mathematical models, which makes additional preprocessing methods necessary. The study analyzes the approximation of experimental points using several functional dependencies: a linear function, a quadratic parabola, a cubic polynomial, a 5th degree polynomial and a two-segment linear model. The article proposes various ways to prove the nonlinearity of the studied sample to exclude a linear model as a possible approximator. The effectiveness of the models is assessed by building cumulative curves of the approximation residuals and comparing the ratio of their ranges to the standard deviation. Attention is also paid to the predictive qualities of the selected mathematical models. The proposed approach can be used to increase the reliability of mathematical modeling in applied problems where experimental data are unevenly distributed or contain several values in a certain section.

Keywords

best mathematical model, adequacy criterion, data heterogeneous structure, least squares method, cumulative curve of residuals, predictive qualities of the mathematical models, approximation of the experimental points

1. Introduction

In modern research in various branches of science and technology, an important place is occupied by the construction of mathematical models based on experimental data. One of the key tasks is to increase the accuracy of approximation of such data, especially in cases where several experimental values observation are observed in individual sections. Such a situation complicates the construction of an adequate model and requires additional approaches to data processing [1, 2, 3, 4].

This work considers the problem of approximating experimental points using different types of approximation functions. At the initial stage, the original data are approximated using a linear function, a parabola, a cubic model, and a two-segment linear approximation. This approach allows us to compare the effectiveness of different models for describing the studied data set [1, 2, 5].

Special attention is paid to cases when several experimental points are present in individual cross-sectional data. To take this feature into account, an approach is proposed that takes place in the correlation field compression (scatter data set). After this data transformation dispersion is decreasing. An approximation is performed using the same methods, which allows us to estimate the impact of pre-processing on the accuracy of the models [6, 7, 8].

Additionally, cumulative curves of residuals are constructed to analyze the results, which allow us to assess the approximation quality. Cumulative curves are a second method for assessing approximation accuracy. For quantitative comparison, the ratio of the ranges of the cumulative curves of residuals to the standard deviation of the corresponding models is used. This makes it possible to determine the most appropriate approximation function for experimental data. This allows us to determine the

CMiGIN'26: Fifth International Workshop on Conflict Management in Global Information Networks, June 26–27, 2026, Krakow, Poland

*Corresponding author.

† These authors contributed equally.

✉ kuzmin_vn@i.ua (V. Kuzmin); panijulia.p@gmail.com (Y. Petrova); alina.osipchuk2012@gmail.com (A. Osipchuk)

ORCID 0000-0003-4461-9297 (V. Kuzmin); 0000-0002-3768-7921 (Y. Petrova); 0000-0002-3768-7921 (A. Osipchuk)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

best mathematical model [9, 10, 11, 12]. The results obtained can be used to improve the accuracy of mathematical modeling in problems where experimental data have a heterogeneous structure.

2. Literature review and problem statement

The building of mathematical models based on the approximation of experimental data remains one of the key areas of modern applied research. The classical approach to solving such problems is the usage of the least squares method (LSM), which ensures ease of implementation, statistical validity, and ease of results interpretation [1, 2, 5, 10, 11, 13]. However, as modern research shows, this method does not always provide sufficient adequacy in cases of complex data structure or the presence of abnormal observations.

In this regard, alternative approaches to approximation are actively being developed. In particular, modern studies propose improved weighted least squares methods, which improve the performance of models through adaptive weighting of deviations [14]. Recently, considerable attention is paid to methods that take into account the heteroscedasticity of data in approximation, since a violation of the assumption of constant variance can lead to ineffective estimates and distortion of the modeling results of the studied data [7, 15, 16, 17].

A separate area of research is the application of clustering methods to analyze data structure [11, 12, 15, 18, 19].

Modern research also emphasizes the need for a comprehensive assessment of approximation quality. It has been shown that minimizing the standard deviation alone does not ensure complete model compliance with experimental data, as local deviations or asymmetries in the error distribution can have a significant impact on modeling results. Therefore, along with classical criteria, additional indicators of adequacy are increasingly used, taking into account the balance of the location of experimental points in relation to the approximating function, maximum deviations, as well as the structure of the residuals.

Methods of data preprocessing and structuring occupy a special place in modern research. Clustering algorithms are used to identify hidden patterns and improve model stability, particularly k-means, which remains one of the most common approaches due to its computational efficiency [20, 21, 22, 23]. However, research indicates its sensitivity to outliers and initial conditions, stimulating the development of modified methods combined with approximation methods and preliminary experimental data analysis procedures [24, 25, 26, 27, 28].

Among the main criteria for building approximating models in modern research, the following can be distinguished [3, 4, 19, 20]:

1. Minimization of the sum of squared deviations;
2. Minimization of the sum of absolute deviations;
3. Minimization of the maximum deviation.

At the same time, the following approaches are used to assess the quality of approximation and increase its reliability:

1. Taking into account heteroscedasticity of data, in particular by minimization of maximum deviations.
2. Analysis of the balance of the number of observations relative to the approximating function.

It should be noted that the choice of the approximation criterion and the method of assessing its quality are always interrelated tasks [19].

Thus, the modern development of the applied approximation theory is characterized by the transition from classical methods to combined and robust approaches that take into account the features of real data, in particular their heterogeneity, the presence of outliers and the complex structure of dependencies, which is confirmed by modern research [7, 9, 12, 24].

Consequently, the evolution of mathematical modeling methods is characterized by a transition from the usage of classical approximation criteria alone to the development of comprehensive adequacy

Table 1
Experimental Data

T, s	A, μm	T, s	A, μm	T, s	A, μm	T, s	A, μm
50	19	100	30; 29; 33	125	28	155	41; 33
60	20; 21	105	35; 32	130	31; 32	160	40; 30; 37
70	17; 22	110	30; 28; 30	135	34; 25	165	32
80	25; 28	115	31; 36; 30	140	26; 33	170	35
90	21; 25; 31	120	36; 25; 28	145	31	180	38
95	25			150	36; 33		

criteria that take into account the characteristics of real data, their heterogeneity, the presence of outliers, and the structure of the error distribution. This determine the relevance of developing new adequacy criteria for constructing mathematical models capable of providing a more reliable and objective assessment of approximation quality [4, 7, 12, 29].

Despite the significant number of existing approaches to approximating experimental data, the issue of improving the reliability of models in cases of complex data set structure remains open [7, 13].

In particular, situations where several experimental values are available in some cross-sectional data remain insufficiently studied, which complicates the building of an adequate approximating function [7, 27].

In addition, the usage of traditional quality criteria, such as the standard deviation, does not always allow to detect local deviations of the model, which may lead to incorrect conclusions about its adequacy. This necessitates the usage of additional analysis methods, in particular the usage of cumulative curves and alternative indicators of the approximation quality [11, 12].

Therefore, it is important to explore approaches to data preprocessing, in particular, to test the original data for nonlinearity using various criteria, in particular using various methods of correlation field compression [10, 13].

3. Materials and methods

The purpose of this work is to increase the reliability of the approximation of experimental data by combining methods of data set preprocessing and the use of different classes of approximating functions with subsequent assessment of their quality based on cumulative curves of residuals and predictive properties.

For this study, experimental data representing the growth of ice crystals A (μm) as a function of time T (s) were selected [1] in Table 1. The data were obtained in the time range from 50 s to 180 s.

A feature of this data set is that in some time intervals (cross-section) there are several repeated length values. This indicates the presence of experimental scatter in the results, caused by both random measurement errors and possible physical features of the crystal growth process.

The experimental data and their approximation by a linear function are shown in Figure 1.

In the work of Draper and Smith's [1] these data were approximated by a straight line and the approximation was considered satisfactory (Fig. 1). But even a visual analysis of this data raises doubts that linear approximation is the best option. This can be confirmed by the fact that the first 5 points in the initial (60-79 s) section lie below the linear approximation, and the probability of this is at least $1/32$. And in the center (80-140 s) of the data, most of the experimental points lie above the linear approximation. To prove the impossibility of using linear approximation, a new test proposed in the work [24] was used.

3.1. The first method for proving nonlinearity

The test is based on constructing a cumulative curve of residuals relative to a linear approximation. In this case, the maximum range of the cumulative curve is calculated, which consists of the sum of the

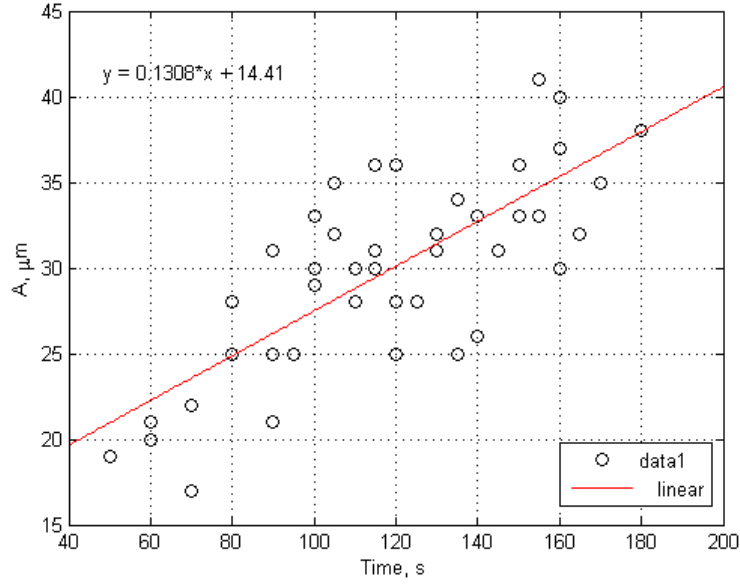


Figure 1: Experimental data and their approximation by a straight line.

maximum upper half-range and lower half-range relative to the zero line. Next, the standard deviation for the linear approximation was calculated. The ratio of the maximum range of the cumulative curve (R_{max}) to the standard deviation (S) shows the degree of adequacy of the applied approximation. This parameter (D) is called the maximum relative range.

$$D = \frac{R_{max}}{S}. \quad (1)$$

For this linear approximation, $R_{max} = 39.4066\mu\text{m}$; $S = 3.7216\mu\text{m}$; $D = 10.5886$.

Using Table 1 in [27] for 43 experimental values with a confidence probability of 0.975, it can be stated that the studied data are nonlinear.

3.2. The second method for proving nonlinearity

Since some cross-sectional data contain repeated observations, a second method of proving nonlinearity is proposed, which is based on compression of the correlation field of the data. Compression consists in averaging the data in cross-sections where repeated values are observed, namely replacing the original data with the same number of average values within that cross-section.

This resulted in a partially compressed correlation field.

A linear approximation for the transformed data by the LSM was performed.

Next, a cumulative curve of residual was constructed for the transformed data, and the standard deviation, maximum range of the cumulative curve, and maximum relative range were found: $R_{max} = 30.4693\mu\text{m}$; $S = 2.4404\mu\text{m}$; $D = 12.4852$.

Then, Table 1 from the [27] was similarly applied to 43 transformed data points, leading to the conclusion that the studied data are nonlinear with a confidence probability of 0.99. These calculations proved that the second method showed the nonlinearity of the data more reliably.

3.3. The third method for proving nonlinearity

It consists of ranking both parameters of the original data and establishing a rank correspondence between the parameters. As a result, new pairs of values are organized and a complete compression of the correlation field is obtained. Let's analyze this correlation field. Let's construct a straight line using the LSM and calculate the standard deviation.

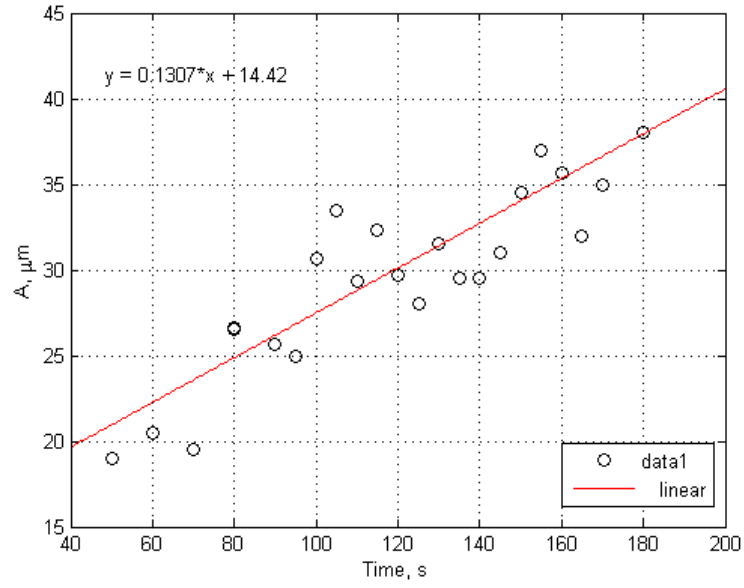


Figure 2: Partially transformed data according to the 2nd method and their approximation by a straight line.

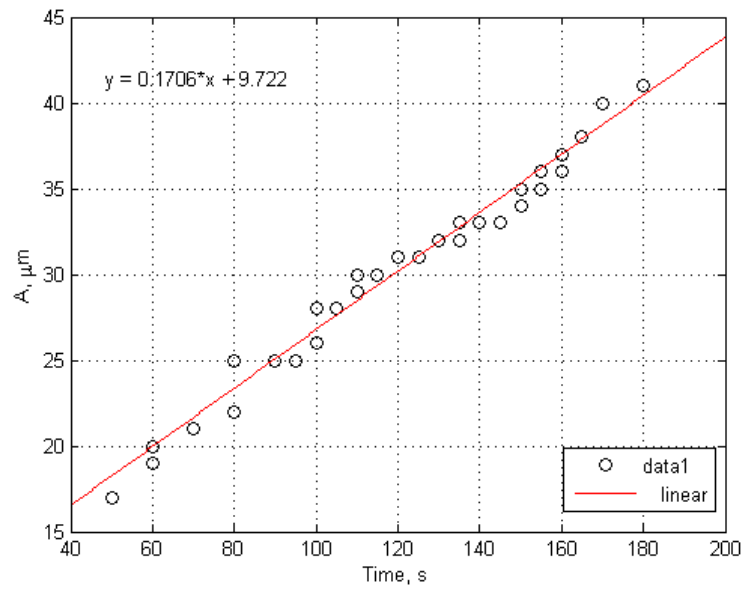


Figure 3: Transformed data according to the 3rd method and their approximation by a straight line.

Further, by analogy with the second method, cumulative curve of residuals is constructed and the standard deviation, the maximum range of the cumulative curve and the maximum average range are determined as follows : $R_{max} = 12.85\mu\text{m}$; $S = 0.8811\mu\text{m}$; $D = 14.5834$.

To prove the impossibility of using linear approximation, we use Table 1 in paper[27]. For 43 data points and the relative range obtained from the table, the resulting confidence level is 0.995. This further enhances the reliability of the proof. The results of constructing cumulative curves of residual for these three cases were not presented in the paper.

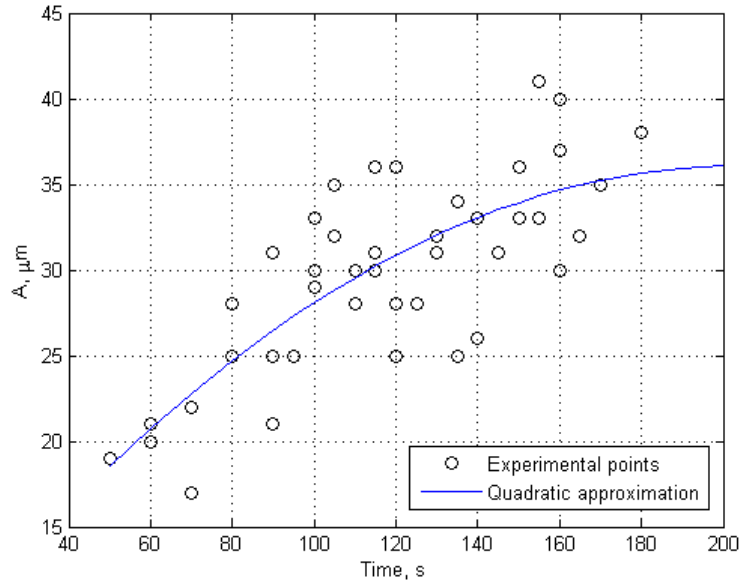


Figure 4: Experimental data and their approximation by a parabola.

3.4. Building of alternative models

Since linear approximation is finally recognized as unsatisfactory, it is necessary to consider several alternative methods of mathematical models building to ensure their maximum adequacy. Usually, the procedure for building, justification and selection an appropriate mathematical model is carried out in the following order:

1. The linear mathematical model is initially verified.
2. Since a simple linear model is not adequate, more complex versions of mathematical models should be considered.

Based on the steps described above, more adequate alternative mathematical models are further considered.

3.5. Parabolic model

At the first stage, approximation by a parabola of the second degree using the least squares method was carried out. The quadratic model (parabola) makes it possible to recognize the presence of curvature in the data. It is often used in cases where nonlinearity is observed. The model has the form

$$A = c + bT + aT^2, \quad (2)$$

where a, b, c is model parameters.

For our case:

$$A = 5.27 + 0.3020T - 0.0007397T^2, \quad (3)$$

As can be seen from Figure 4, the parabolic approximation more adequately describes the first five points, i.e. better than linear approximation. Although this model has a smaller standard deviation than the linear approximation, it is not an appropriate model for describing the data because it goes down after the time value of 204.05 seconds, that contradicts the physical nature of the data under study.

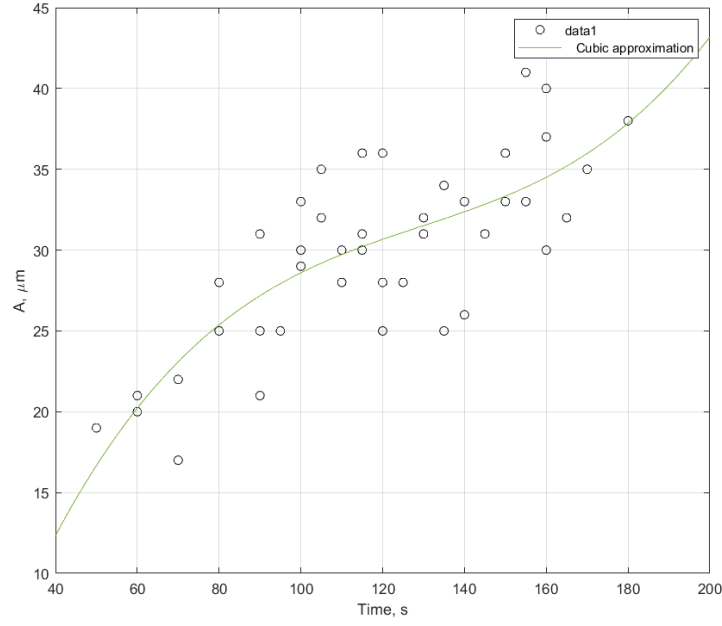


Figure 5: Experimental data and their approximation by a cubic model.

3.6. Cubic model

At the next stage, the mathematical model will be built in the cubic form by the LSM (Fig. 5), its equation taking the form:

$$A = -14.73 + 0.9046T - 0.006352T^2 + 1.6387 * 10^{-5}T^3, \quad (4)$$

The graph shows that the cubic approximation has two sections that do not correspond to the physical nature of crystal growth: namely, in the initial segment at $T = 0$, it has negative values for crystal growth A , and in the time segment after $T = 140$ s, the values for A increase sharply. Therefore, this approximating function also cannot be accepted as satisfactory.

3.7. Polynomial model

An even more complex model in the form of a 5th degree polynomial obtained using LSM will be used to increase the accuracy of the approximation. The resulting 5th degree polynomial equation has the form:

$$A = 269.47 - 13.359T + 0.2657T^2 - 0.00246T^3 + 1.0858 * 10^{-5}T^4 - 1.835210^{-8}T^5. \quad (5)$$

The graph of a 5th degree polynomial is shown in Fig. 6. The graph shows that the polynomial of the 5th degree has significant fluctuations outside the range of the given data, making it entirely inappropriate for forecasting. This indicates that this approximating function cannot be accepted too. In the opinion of the authors, it would be more rational to use two-segment linear approximation as the best mathematical model.

3.8. Building of a two-segment linear model

A two-segment linear model with a single switching point allows one to describe relationships that vary across different ranges of independent variable values. This approach is appropriate in cases where a process has different modes of operation or where a trend change is observed. The analytical description of the trend change is performed using a modular representation, namely, by the formula:

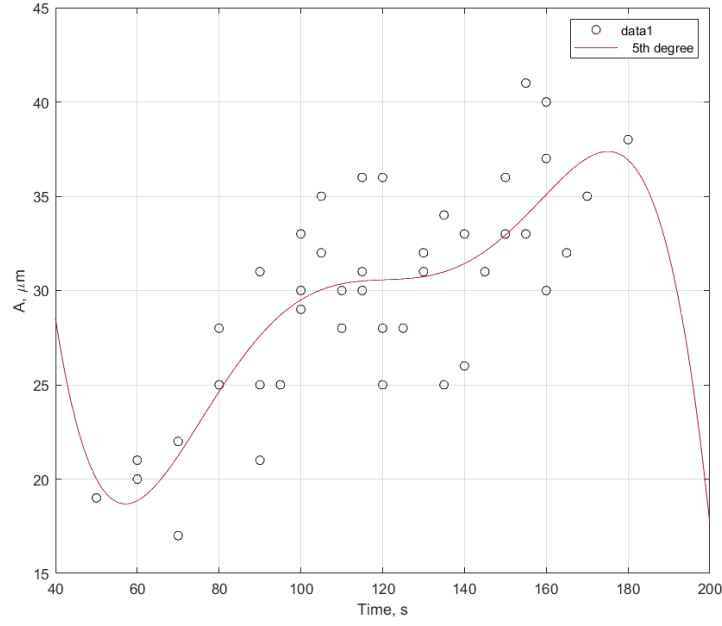


Figure 6: Experimental data and their approximation by a 5th degree polynomial.

$$A = a + bT + c(T - \dot{T})_+, \quad (6)$$

$$(T - \dot{T})_+ = \frac{(T - \dot{T}) + |(T - \dot{T})|}{2}. \quad (7)$$

where a, b, c – model parameters, $\dot{T}$ – is the abscissa of the switching point.

The choice of such mathematical model is driven by the need to construct a more adequate model, while still maintaining low complexity. Moreover, such a model has very good predictive properties.

An important point when using this model is the determination of the optimal abscise of the switching point.

Let's now proceed to calculation of the two-segment model for the original experimental data using the least-squares method. To do this, the following system of equations should be solved. To simplify the notation, we denote x as the independent variable, $\dot{x}$ as the abscissa of the switching point, and y as a function of x .

$$\begin{cases} \sum y_i = Na + b \sum x_i + c \sum (x_i - \dot{x})_+ \\ \sum y_i x_i = a \sum x_i + b \sum x_i^2 + c \sum x_i (x_i - \dot{x})_+ \\ \sum y_i (x_i - \dot{x})_+ = a \sum (x_i - \dot{x})_+ + b \sum x_i (x_i - \dot{x})_+ + c \sum (x_i - \dot{x})_+^2 \end{cases} \quad (8)$$

To correctly solve this system, it should be taken into account that $(x_i - \dot{x})_+^2 = (x_i - \dot{x})_+ \cdot (x_i - \dot{x})_+$.

Let's determine the abscissa of the optimal switching point. For this, 5 variants of the two-segment linear model for the abscissas of the switching points 80, 90, 100, 110 and 120 should be calculated. For each model variant, the standard deviation should be calculated and a parabola should be constructed for the obtained data using LSM (Fig. 7). Next, the minimum of this function is determine, which represents the optimal abscissa of the switching point.

The optimal abscissa for the switching point is $x_{min} = 100.84$.

It should be noted that determining the optimal abscissa of the switching point is an iterative process; however, the current study is limited to one iteration.

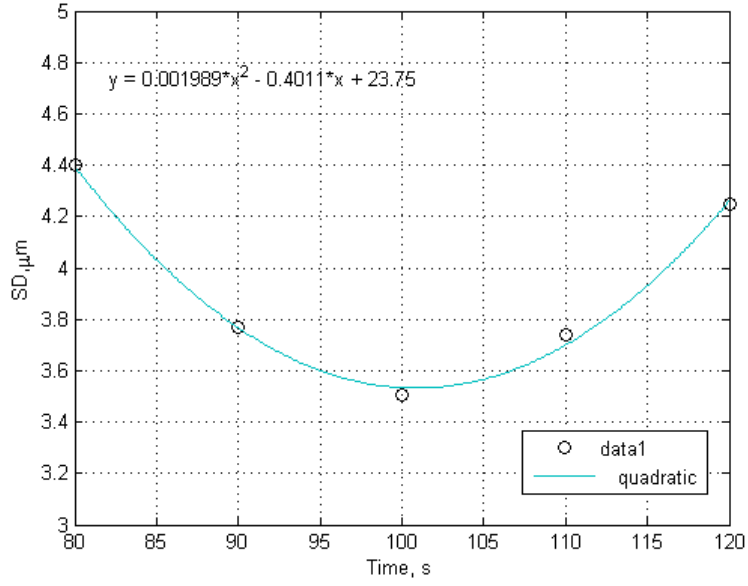


Figure 7: Parabola for determining the optimal switching point.

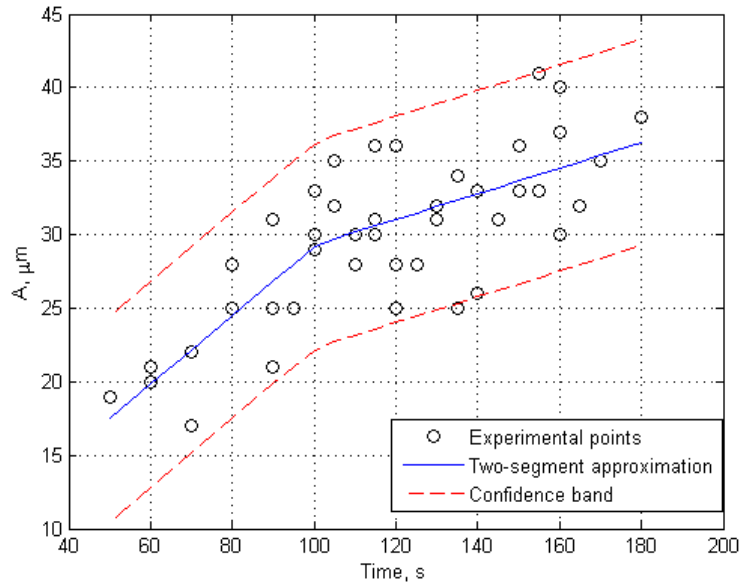


Figure 8: Experimental data and their approximation by two-segment linear regression with confidence intervals.

Next, system (8) was solved using the abscissa of the optimal switching point to find the coefficients a, b, c .

The two-segment linear regression equation for this case has the form:

$$A = 5.818 + 0.2335T - 0.1466(T - 100.84)_+. \quad (9)$$

Also, for this approximation, the figure shows a confidence band of ± 2 standard deviations. It should be noted that all experimental points fell within this band.

A preliminary analysis of the results of the studied models – namely, linear, parabolic and 5th degree polynomial – showed that due to their inherent shortcomings, these models cannot be used as the optimal choice.

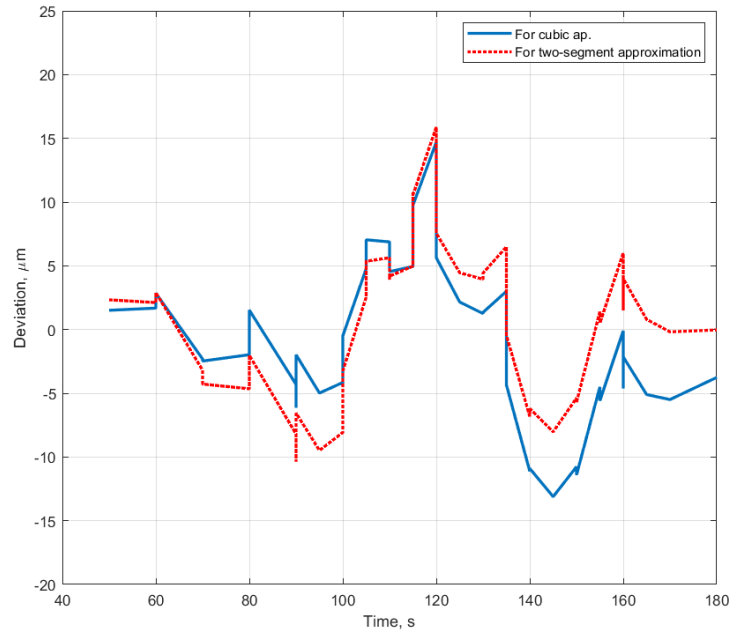


Figure 9: Cumulative curves of residuals for cubic model and two-segment linear regression for the experimental data set

Table 2

Statistical Accuracy Parameters of Different Approximation Types of Original Data

Type of approximating function	Standard deviation	The maximum range of the cumulative curve of residuals	Maximum relative range
Linear	3.7216	39.4066	10.5886
Parabola	3.6624	30.2115	8.2491
Cubic model	3.5576	26.2629	7.3821
5th degree polynomial	3.4219	18.2379	5.3298
Two-segment linear regression	3.5482	27.8186	7.8403

Further, a comparative analysis of the remaining models adequacy was carried out – for cubic model and two-segment linear regression. For these models, cumulative curves of residuals were constructed.

For all studied types of approximating functions, a summary table of statistical accuracy (adequacy) parameters was compiled.

Analysis of Table 2 shows that the best quantitative indicators of model adequacy are provided by cubic approximation and two-segment linear regression (with the exception of the 5th degree polynomial, which is not considered due to poor predictive properties). However, due to the poor predictive properties of cubic approximation, the two-segment linear approximation with the optimal switching point was chosen as the best approximation.

4. Conclusions

This study involved searching for the optimal mathematical model for the analyzed data by selecting the best approximating function.

The proof of nonlinearity was performed by three methods. The first involved testing the nonlinearity of the original data, while the second involved testing the transformed data by partially compressing

the correlation field. The third method utilized full compression of the correlation field. In all three cases, convincing evidence of the data nonlinearity was obtained, i.e., demonstrating the impossibility of using linear approximation for studied data set.

The following approximation functions were considered as alternative models: parabolic, cubic, fifth-degree polynomial, and two-segment linear regression. To construct the two-segment regression approximation function, the optimal switching point was determined. When considering the best models, the linear, parabolic, and fifth-degree polynomial approximations were discarded due to their predictive shortcomings.

A further comparative analysis of the cubic and two-segment linear regression models was carried out using a new test, which consists of construction a cumulative curve of residuals.

As a result, the optimal two-segment linear regression with the optimal switching point was selected as the best model, which showed very good adequacy and good predictive properties.

It was also shown that while the 5th degree polynomial has the smallest standard deviation of the approximation functions, its predictive properties are unsatisfactory; therefore its usage is not recommended. This once again demonstrates that models based on high-order polynomials are ineffective for such applications, and that two-segment regression should be used more frequently when constructing mathematical models.

It should be remembered that when choosing the best mathematical model, there must always be a compromise between the best parameters of the model's adequacy and its predictive properties.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

- [1] N. Draper, H. Smith, *Applied Regression Analysis*, John Wiley and Sons, New York, 1998.
- [2] D. C. Montgomery, E. A. Peck, G. G. Vining, *Introduction to Linear Regression Analysis*, Wiley, New York, 2021.
- [3] G. James, D. Witten, T. Hastie, R. Tibshirani, *An Introduction to Statistical Learning: With Applications in R*, 2 ed., Springer, New York, 2021. doi:10.1007/978-1-0716-1418-1.
- [4] M. Nguyen, *Regression Techniques for Data Analysis*, 1 ed., Springer, Cham, 2026. doi:10.1007/978-3-032-01835-9.
- [5] Y. Zuo, H. Zuo, Weighted least squares regression with the best robustness and high computability, *Axioms* 13 (2024).
- [6] N. Klein, Distributional regression for data analysis, *Annual Review of Statistics and Its Application* 11 (2024) 321–346. doi:10.1146/annurev-statistics-040722-053607.
- [7] H. White, Robust estimation of heteroscedastic regression models: overview and new proposals, *Statistical Papers* (2025).
- [8] V. Volkogon, T. Konrad, O. Kolganova, L. Tereshchenko, Situational synthesis of algorithmic support of regional transport systems based on expert system, in: *Proceedings of the International Workshop on Advances in Civil Aviation Systems Development*, volume 736 of *Lecture Notes in Networks and Systems*, 2023, pp. 321–333. doi:10.1007/978-3-031-38082-2_25.
- [9] V. Kuzmin, O. Shcherbyna, Y. Petrova, A. Osipchuk, New approaches to linear approximation in condition of heteroskedancity data, 2021, pp. 218–221. doi:10.1109/ACIT52158.2021.9548483.
- [10] T. Farrar, R. Blignaut, R. Luus, S. Steel, A review and comparison of methods of parameter estimation and inference for heteroskedastic linear regression models, *Journal of Applied Statistics* 52 (2025) 3091–3120. doi:10.1080/02664763.2025.2496719.
- [11] R. Esposito, M. Cerrato, M. Locatelli, Partitioned least squares, *Machine Learning* 113 (2024) 6839–6869. doi:10.1007/s10994-024-06582-3.
- [12] D. Birkes, Y. Dodge, *Alternative Methods of Regression*, John Wiley & Sons, New York, 1993.

- [13] M. Salibian-Barrera, Robust nonparametric regression: Review and practical considerations, *Econometrics and Statistics* 28 (2023) 75–101. doi:10.1016/j.ecosta.2023.04.004.
- [14] E. Xiao, Comprehensive k-means clustering, *Journal of Computer and Communications* 12 (2024).
- [15] J. F. Vera-Vera, J. A. Roldán-Nofuentes, A robust alternating least squares k-means clustering approach, *Mathematical Biosciences and Engineering* 21 (2024) 6839–6869.
- [16] O. Sushchenko, et al., Algorithm of determining errors of gimballed inertial navigation system, in: O. Gervasi, B. Murgante, C. Garau, D. Taniar, A. M. A. C. Rocha, M. N. Faginas Lago (Eds.), *Computational Science and Its Applications – ICCSA 2024 Workshops*, volume 14816 of *Lecture Notes in Computer Science*, Springer Nature Switzerland, Cham, 2024, pp. 206–218. doi:10.1007/978-3-031-65223-3_14.
- [17] O. Solomentsev, M. Zaliskyi, O. Kozhokhina, T. Herasymenko, Efficiency of data processing for uav operation system, in: 2017 IEEE 4th International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), 2017, pp. 27–31. doi:10.1109/APUAVD.2017.8308769.
- [18] K. D. Lawrence, J. L. Arthur, *Robust Regression: Analysis and Applications*, CRC Press, Boca Raton, 2019.
- [19] D. M. Himmelblau, *Process Analysis by Statistical Methods*, John Wiley and Sons, New York, 1970. 958.
- [20] R. S. Kenett, S. Zacks, *Modern Industrial Statistics: With Applications in R, MINITAB, and JMP*, 3 ed., Wiley, Hoboken, 2021.
- [21] V. Kuzmin, O. Shcherbyna, Y. Petrova, A. Osipchuk, New approaches to linear approximation in condition of heteroskedancity data, 2021, pp. 218–221. doi:10.1109/ACIT52158.2021.9548483.
- [22] O. Holubnychyi, et al., Self-organization technique with a norm transformation based filtering for sustainable infocommunications within CNS/ATM systems, in: I. Ostroumov, M. Zaliskyi (Eds.), *Proceedings of the 2nd International Workshop on Advances in Civil Aviation Systems Development*, volume 992 of *Lecture Notes in Networks and Systems*, Springer Nature Switzerland, Cham, 2024, pp. 262–278. doi:10.1007/978-3-031-60196-5_20.
- [23] M. Zaliskyi, Y. Petrova, M. Asanov, E. Bekirov, Statistical data processing during wind generators operation, *International Journal of Electrical and Electronic Engineering & Telecommunications* 8 (2019) 33–38. doi:10.18178/ijeetc.8.1.33-38.
- [24] V. N. Kuzmin, New statistical method for identification of nonlinearity of empirical data, in: *Proceedings of Fifth International Conference on Computer Data Analysis and Modeling*, volume A–M, 1998, pp. 159–164.
- [25] V. N. Kuzmin, Transformation of correlation fields of empirical data [in russian], in: 2014 Fifteenth International Scientific Mykhailo Kravchuk Conference, volume *Probability theory and mathematical statistics*, NTUU “KPI”, Kiev, Ukraine, 2014, p. 80.
- [26] V. Kuzmin, O. Solomentsev, M. Zaliskyi, Methodical approach for testing statistical data for linearity [in ukrainian], *Problems of Informatization and Management* 4 (2015) 63–67.
- [27] V. Ivannikova, M. Zaliskyi, O. Solomentsev, I. Ostroumov, N. Kuzmenko, Statistical data processing technologies for sustainable aviation: A case study of Ukraine, *Sustainability* 17 (2025) 5781. doi:10.3390/su17135781.
- [28] O. Solomentsev, M. Zaliskyi, T. Herasymenko, O. Kozhokhina, Y. Petrova, Efficiency of operational data processing for radio electronic equipment, *Aviation* 23 (2019) 71–77. doi:10.3846/aviation.2019.11849.
- [29] S. Smerichevska, L. Ivanenko, Z. Poberezhna, S. Smerichevskyi, T. Konrad, D. Dudko, Matrix-vector approach to assessing the competitiveness of logistics enterprises in the context of economic security and information risks, in: *CEUR Workshop Proceedings*, volume 4024, 2025, pp. 440–451.