Zhadyra Artykova

Mathematical Modeling and Process Optimization of Composite Polymer Stabilizers

Technical Report No. 59
February 2026

Department of International Outreach

Title: Mathematical Modeling and Process Optimization of Composite Polymer Stabilizers

Authors: Zhadyra Artykova

Abstract

This paper examines current mathematical modeling methods used for the development and optimization of composite polymer stabilizers. A review of key challenges related to the stability of polymer materials is provided, along with a discussion of modern computational approaches, including molecular dynamics, finite element analysis, thermodynamic modeling, and machine learning. The necessity of an interdisciplinary approach integrating chemistry, materials science, and computational technologies is justified. Perspectives on the further development of modeling methods to enhance the efficiency and stability of polymer stabilizers are presented.

Introduction

Drilling fluid behavior in complex thermobaric conditions and under the influence of mineralized formation waters can be studied and described using mathematical modeling, an efficient scientific analysis and forecasting technique. Rheological and filtration parameters can be computed ahead of time using mathematical models, which helps choose the fluid composition more wisely and minimizes the need for expensive experimental research. When developing and using new polymer stabilizers, modeling is particularly important because it makes it possible to assess their efficacy, identify the limits of reagent performance under actual drilling conditions, and determine the ideal concentrations and usage conditions. Therefore, mathematical modeling provides a theoretical foundation for enhancing the chemical and technological systems of drilling fluids based on modified polymers in addition to being a tool for optimizing technological processes and enhancing drilling operations safety.

The advancement of polymer materials with high stability and operational reliability is necessary for the growth of sectors like biomedicine, water treatment, and oil and gas. Composite polymer stabilizers reduce environmental hazards, improve process efficiency, and increase equipment longevity. Mathematical modeling, which allows precise material property prediction and composition optimization, is one of the main techniques for their development.

This study's objective is to examine contemporary methods for mathematical modeling polymer stabilizers and talk about how they might be used to forecast the mechanical, chemical, and thermal stability of materials [1, 2].

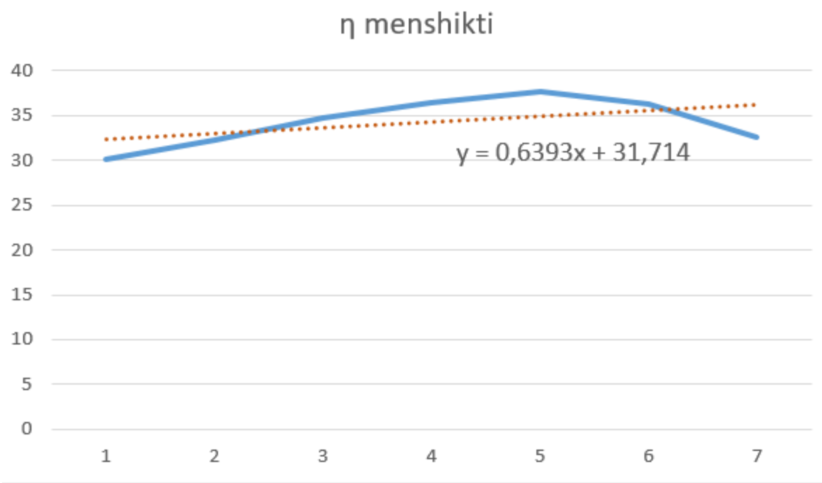
Methods and research

The relationship between the parameters under investigation was evaluated using regression analysis. A design matrix was created for this purpose, enabling a methodical examination of the impact of individual elements and their interactions. The coefficient of determination (R^2), which represents the percentage of variance explained by the model and, consequently, the suitability of the regression equation obtained, is a crucial statistical indicator of this analysis. The methods and research presented were carried out as part of the Bolashak program at Constructor University.

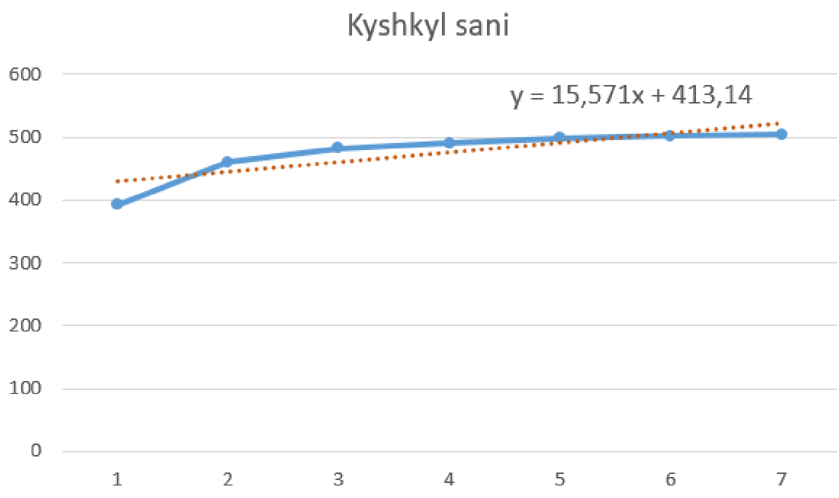
Results and discussion

Mathematical processing of the experimental data consisted in constructing the answer in the form of z = f(x1, x2, x3). The investigated parameter z was the changes in physico-chemical properties during the copolymerization of acrylonitrile and vinylsulfonic acids, their hydrolysis and further modification.

Analyzing the data obtained during hydrolysis and modification of the copolymer using gossypol resin and fatty acids, graphs were constructed for comparison with trend dependence, and a regression ratio was determined for each factor. (Pictures 1, 2).



Picture 1– Diagram for the eigenvalue factor η



Picture 2 – Diagram for acid number factor

Analysis was carried out according to the copolymerization data of acrylonitrile and vinylsulfonic acid (Figure 3).

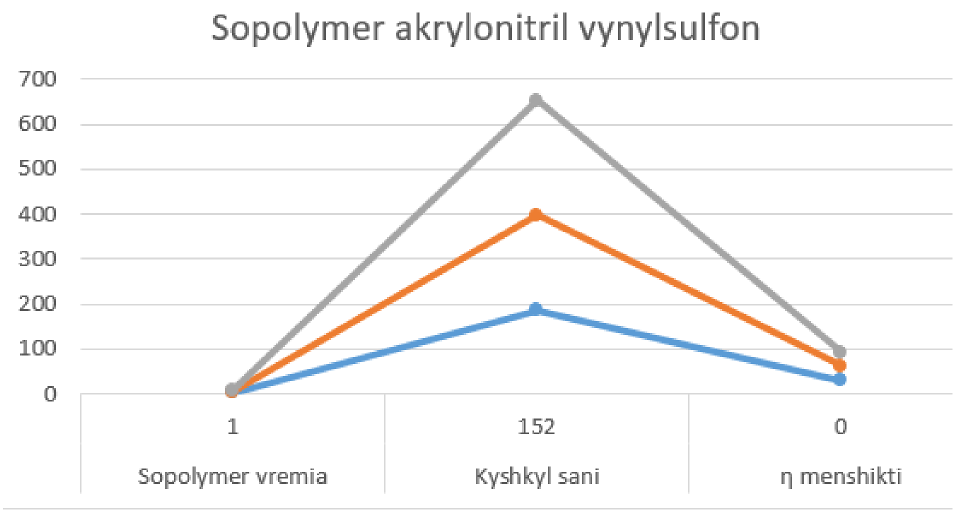


Figure 3 – Acrylonitrile and vinylsulfonic acid copolymerization details

Regression analysis was performed according to the copolymerization data of acrylonitrile and vinylsulfonic acid (Table 1).

Table 1 – Regression analysis according to the copolymerization data of acrylonitrile and vinylsulfonic acid

	a11	a12	a13
a11	1		
a12	0,9956703	1	1

a13	0,7829549	0,758626294	

The design matrix is a statistical analysis tool that allows determining the degree of dependency between variables, evaluating their interrelationships, and using the obtained data for forecasting or decision-making. Each cell in the table shows the dependency between two specific variables [3, 4].

The design matrix helps identify how changes in one variable can affect another, and based on this, make predictions, optimize processes, and make informed decisions.

The calculated coefficients of the design matrix indicate the strength of the relationship between the factors. The larger the absolute value of the coefficients, the closer the corresponding factor is to the result. The analysis of the matrix consists of two stages:

1. If there are elements in the first column of the calculated design matrix where $|r^{xy}| < 0.5$, the factors that meet this condition are removed from the model.
2. When considering the correlation of the calculated coefficients of the factors in the design matrix, it is necessary to assess their independence from each other. This is a required condition for regression analysis.

According to our calculations, the values of all parameters are at an adequate level, as shown in Table 2.

Table 2 – Regression Statistics

Multiple R	0,758626294
R-squared	0,575513853
Adjusted R-squared	0,36327078
Standard Error	34,76366004
Observations	4

Another important indicator is R-squared, also known as the coefficient of determination. It represents the proportion of the variance in the response variable that can be explained by the predictor variable [5].

The value of R-squared can range from 0 to 1. A value of 0 indicates that the response variable cannot be explained at all by the predictor variable. A value of 1 indicates that the response variable can be fully explained without error by the predictor variable.

One of the important indicators is the value located at the intersection of the "Y-intercept" row and the "coefficients" column. Here, the value of the "Y-intercept" is crucial. According to the calculations, this value is 151.43 (Table 3).

Table 3 – Coefficients

	Coefficients
Y-intercept	151,4289433
Variable X1	2,157608561

Additionally, the value at the intersection of the "Variable X1" row and the "Coefficients" column shows the level of dependency between Y and X. The coefficient of 2.158, obtained from the calculations, is considered a very high indicator of effect.

The calculations were performed using the "least squares method." The least squares method is a statistical procedure used to predict the behavior of dependent variables accurately. The essence of the least squares method is to find the closest approximation to reality among all linear functions. This can be done by searching for the function with the smallest deviation, or more precisely, by finding the minimal sum of squared deviations of the Y values from the regression equation obtained during the process.

The regression equation is written as follows:

$$Y=2.158x+151,429 \quad (1.1)$$

MatLab is a popular tool used for working with matrix data, virtualization, and mathematics. The MATLAB language is high-level and is used for modeling mechanical and chemical-technological processes. It includes an application package as well as an integrated development environment (IDE). MATLAB allows for efficient numerical computation, data visualization, and algorithm development, making it widely used in various fields such as engineering, science, and economics [6].

The regression equation results were obtained by solving through the values of the factors using programming code in the MatLab programming environment.

MatLab provides commands and functions for creating three-dimensional graphs. The values of elements in a numerical array are considered as points on a plane defined by the X and Y coordinates, with the Z coordinates representing the height above the plane. These functions and commands perform operations like connecting points in a cross-section (using the plot3 function) and creating mesh surfaces (using the mesh and surf functions). The surface drawn with the mesh function is a mesh surface, where its cells have a background color, and their boundaries can have a color defined by the EdgeColor property of the surface graphical object [7].

To graphically display the results of the regression equation, we use the functions of the MatLab programming environment. Using the mesh function, the following type of plot was obtained (Figures 4, 5, 6).

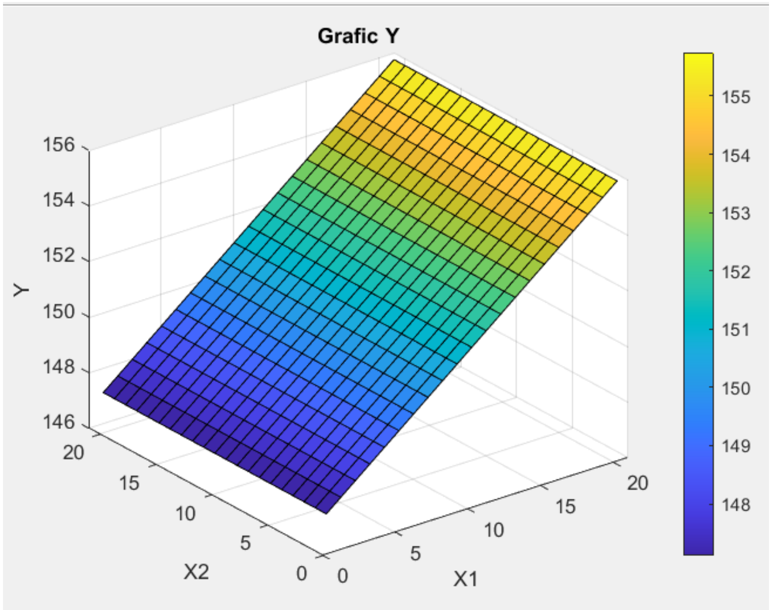


Figure 4 – Results of the regression equation using the surf function

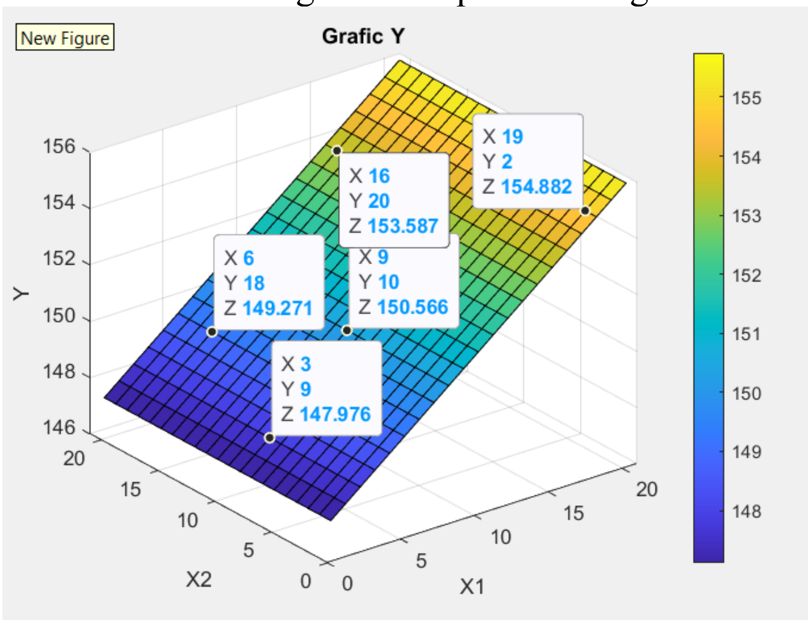


Figure 5 – Results of the regression equation with the coordinates of the given point

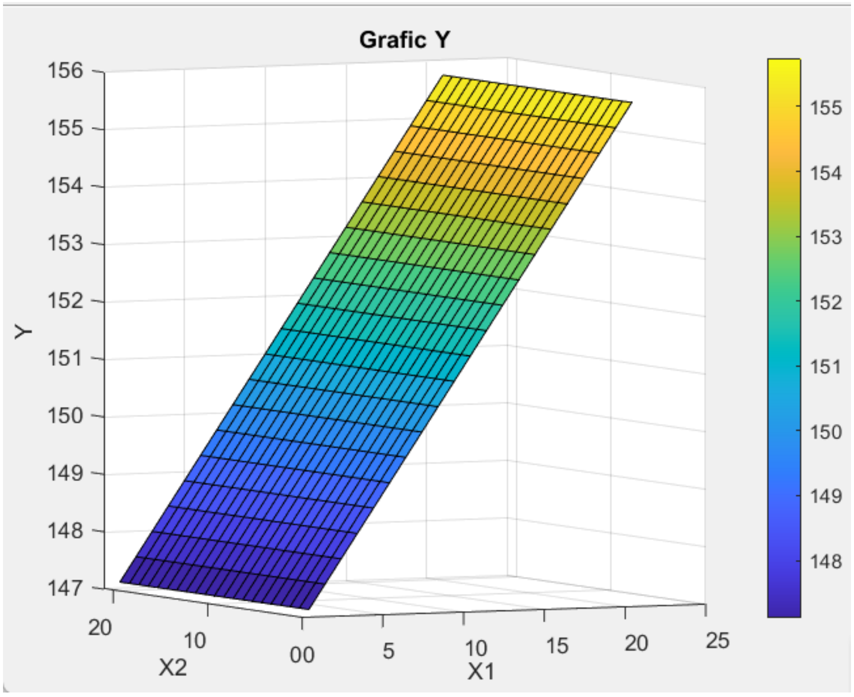


Figure 6 – Visualization of the calculation results with step reduction

The residual values in the calculations are very small and satisfy all given conditions. We determine the standard errors, i.e., we calculate the deviations of the observed values from the regression line. These calculated values also satisfy the conditions.

During the data analysis, we calculate the following parameters:

df – degrees of freedom of the regression, which is equal to the number of regression coefficients.

$$df = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1-1) + (s_2^2/n_2)^2/(n_2-1)} \qquad (1.2)$$

SS_ι – the sum of the squared residuals is calculated using the following formula:

$$SS_\iota = \sum_{i=1}^{a}\sum_{j=1}^{n}(X_{ij} - \bar{X}_\iota)^2 = \sum_{i=1}^{a}(n_i - 1)s_i^2 = (n_1 - 1)s_1^2 + \cdots + (n_i - 1)s_i^2 \qquad (1.3)$$

MS_a – the mean squared residuals (also known as the mean squared error, MSE) are calculated using the following formulas:

$$MS_a = \frac{SS_a}{a-1} \qquad (1.4)$$

$$MS_l = \frac{SS_l}{a-1} \quad (1.5)$$

The F-statistic is calculated as the ratio of the Mean Square of the regression (MS_regression) to the Mean Square of the residuals (MS_residual). This statistic shows how well the regression model fits the data compared to a model without independent variables. Essentially, it tests whether the regression model is generally useful.

The formula for the F-statistic is:

$$F = \frac{MS_a}{MS_l} \quad (1.6)$$

The calculation results are as follows (Table 4):

Table 4 – Calculation Results

	df	*SS*	*MS*	*F*	*F*
Regression	1	3276,075882	3276,975882	2,71157897	0,241373706
Residual	2	2417,024118	1208,512059		
Total	3	5694			

Each individual coefficient is interpreted as the average increase in the response variable for each unit increase of the corresponding predictor variable, assuming all other predictor variables remain constant. The accuracy of the regression analysis is checked according to Student's criteria [8].

Student's t-test is used to determine the statistical significance of the differences between means. To calculate this criterion, the mean value of the parameter from the first group is subtracted from the mean value of the parameter from the second group, and the result is divided by the sum of the squared errors. The latter is necessary to scale the t-criterion to the required measure.

Conditions for using the t-statistic:

- The data must be numerical;
- The data should follow a normal distribution;
- The Student's t-test can only be used to test the hypothesis of the difference between the means of two groups (if we are comparing multiple groups, then in most cases, we use analysis of variance (ANOVA));
- The variances of the two samples we are comparing should be approximately equal.

The data in the t-statistics column of the table represent the values of Student's t-test and are calculated using the following formula:

$$t = \frac{X_1 - X_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (1.7)$$

Table 5 – Regression Analysis

	Coefficients	Standard Error	t- statistic	p-Value
Y Intercept	151,4289433	34,76131416	4,356249095	0,048865385
Variable X1	2,157608561	1,310272205	1,646687271	0,241373706

The calculated values of the t-criterion in the table are analyzed. For each group, the number of given pre-studied factors is taken into account. The value of the degrees of freedom f is determined using the following formula:

$$f = (n_1 + n_2) - 2 \quad (1.8)$$

After all the calculations, we concluded that the t-criterion's required significance level is appropriate. The calculated t-statistics are sufficiently large, exceeding the critical value. This allows us to conclude that the observed differences are statistically significant (significance level $p<0.05$) [9].

Conclusion. Without the need for extensive experimental testing, mathematical modeling is a potent tool for creating new polymer stabilizers, enabling the prediction of their properties and the optimization of their composition. Molecular dynamics, finite element analysis, and machine learning are examples of contemporary computational techniques that present new possibilities for developing materials with better properties. The efficiency of polymer composition design will be greatly increased by further development of hybrid models and interdisciplinary approaches. After analyzing the t-criterion's calculated values, we came to the conclusion that the t-criterion's necessary degree of significance satisfies the requirements. When compared to a model without independent variables, statistics demonstrate how well the regression model fits the data.

References

1 Smith J., et al. Molecular dynamics simulation of polymer stabilization. *Journal of Computational Chemistry*, 2023.
2 Lee R., et al. Finite element analysis in polymer science. *Polymer Engineering & Science*, 2022.

3 Zhang H., et al. Thermodynamic modeling of polymer degradation. *Materials Science Reports*, 2021.
4 Wang X., et al. Machine learning in polymer material design. *AI in Materials Science*, 2024.
5 A. B. Issa, O. K. Beisenbayev, Zh. K. Artykova, et al., "Polymeric compositions to increase oil recovery," Rasayan J. Chem., vol. 16, no. 4, pp. 876–883, 2023, doi: 10.31788/RJC.2023.1628295.
6 Zh. K. Artykova, O. K. Beisenbayev, A. A. Kadyrov, S. A. Sakibayeva, and B. M. Smailov, "Synthesis and preparation polyacrylonitrile and vinyl sulfonic acid in the presence of gossypol resin for drilling fluids," Rasayan J. Chem., vol. 16, no. 4, pp. 2313–2320, 2023, doi: 10.31788/RJC.2023.1618497.
7 Zh.K. Artykova, O. K. Beisenbayev, et al., "Modification of polymers to synthesize thermo-salt-resistant stabilizers of drilling fluids," Open Eng., vol. 15, no. 1, p. 20240097, 2025, doi: 10.1515/eng-2024-0097.
8 Lazarev, Yu. Modeling Processes and Systems in MATLAB: Textbook. – St. Petersburg, 2005. – 512 p.
9 A.M. Gumerov. Mathematical Modeling of Chemical and Technological Processes: Textbook. 2nd edition, revised. St. Petersburg: Lan., 2014. – 176 p.

Authors:
Zhadyra Artykova
Program: Internship Program 2025–2026
Supervisor: Dr. Anna Tevyashova