Available online at www.HighTechJournal.org

HighTech and Innovation Journal

ISSN: 2723-9535

Vol. 5, No. 3, September, 2024

Permissible Extrapolation Justification of the Multiplicative Multifactorial Model and Its Application to the White Soot Production Technology

Lyutsiya Karimova [1, 2*], Guldana Makasheva [1, 3], Vitaliy Malyshev [4], Yelena Kharchenko [1, 5], Yerlan Kairalapov [1, 2]

[1] *Metallurgy Laboratory of LLP "Innovation", Karaganda 100024, Kazakhstan.*

[2] *LLP "KazHydroMed", Karaganda 100000, Kazakhstan.*

[3] *Satbayev University, Almaty 050013, Kazakhstan.*

[4] *Chemical-Metallurgical Institute named after Zh. Abisheva, Karaganda 100009, Kazakhstan.*

[5] *Non-Profit Joint Stock Company "Karaganda Industrial University", Temirtau City 101400, Kazakhstan.*

Received 25 May 2024; Revised 23 August 2024; Accepted 27 August 2024; Published 01 September 2024

Abstract

The solution to a specific technological problem is combined with the methodological development of a nonlinear multifactorial relationship to justify the boundaries of its extrapolation beyond the experimental range used. White soot was produced through two-stage carbonization of a silicate solution (composition, g/l: Na_2O = 126.5, SiO_2 = 107.7, Al_2O_3 = 3.1), obtained after processing waste tailings with carbon dioxide in a recirculation system. The influence of the deposition duration, temperature, and final pH value of the pulp on the formation of the specific surface area (S_{sp}, m^2/g) of white soot was studied. The specific surface area was calculated from the average diameter of the white soot particles measured using an electron microscope. A multifactorial experiment was designed, and the experimental results were processed using a probabilistic deterministic method for experiment design (PDED) to obtain a nonlinear multiplicative combined model. A new interpretation of the subordination of the nonlinear multiple correlation coefficient R and the R2 value was given as relating to the structural and adaptive components of complex self-organizing systems. This determines their use for assessing the ratio of the basic R and extrapolated R^2 ranges of variation for each factor and any combinations thereof and multifactorial dependence in general. The results are presented in the form of multifactor tabular nomograms measured by the number of multifactor cells in localized areas of optimal sets that allow isolation by one or another combination of factors. The technological object of extrapolation of the 'white soot' production is linked to the solution of emerging methodological problems and illustrates the accessibility of the engineering application of the proposed method for combining nonlinear and linear approaches to mathematical experimental design.

Keywords: White Soot; Specific Area; Multifactor Model; Correlation; Permissible Extrapolation.

1. Introduction

Recently, mathematical planning of experiments has been developed to improve the accuracy of displaying the relationships of an existing sample using methods for combining linear and nonlinear additive and multiplicative models [1-6]. Moreover, combinatorial enumeration generates an optimal model structure, which is the individual for

* Corresponding author: lyuciya_karimova@bk.ru

http://dx.doi.org/10.28991/HIJ-2024-05-03-08

each sample each time. For example, when using the neural network method, it is directly indicated that an identical generalized mathematical model of an object simply does not exist [1]. In this case, there is no actual planning for the experiment, and it formally belongs to the category of passive experiments, which use a randomly generated set of data without observing any conditions. In the absence of a mathematical model, questions about the possibility of extrapolation, let alone permissible ones, are inappropriate. Meanwhile, the problem of increasing the accuracy and reliability of the results of a multifactorial experiment can be solved by active intervention in the initial data to identify an unknown space of optimal results. In the proposed method, this is achieved by revealing the relationship between the statistical criteria R and R^2 as general systemic criteria of stability, achieved with a ratio equal to the proportion of the golden section. In this case, the intrasystemic limitation of the extrapolation procedure is the adaptive component $R^2 \leq R$.

The production of multifactor linear and nonlinear models of technological processes is well-developed [7-13]; however, the procedure for multifactor extrapolation into unexplored areas of the multifactor technological space and optimization of results remain ununified. Therefore, it is necessary to accumulate such data when solving production problems and to find rational methods for processing them. Over the past decades, there has been intensive growth in the research areas of science and technology based on the use of various forms of silica [14]. Quartz sand and tailings are promising raw materials for the production of amorphous silica, which is used in various industries and is in great demand in the market. Therefore, the study of processes for processing quartz-containing raw materials to obtain pure silicate products (white soot) is an urgent task.

Figure 1 presents a flowchart of the research process.

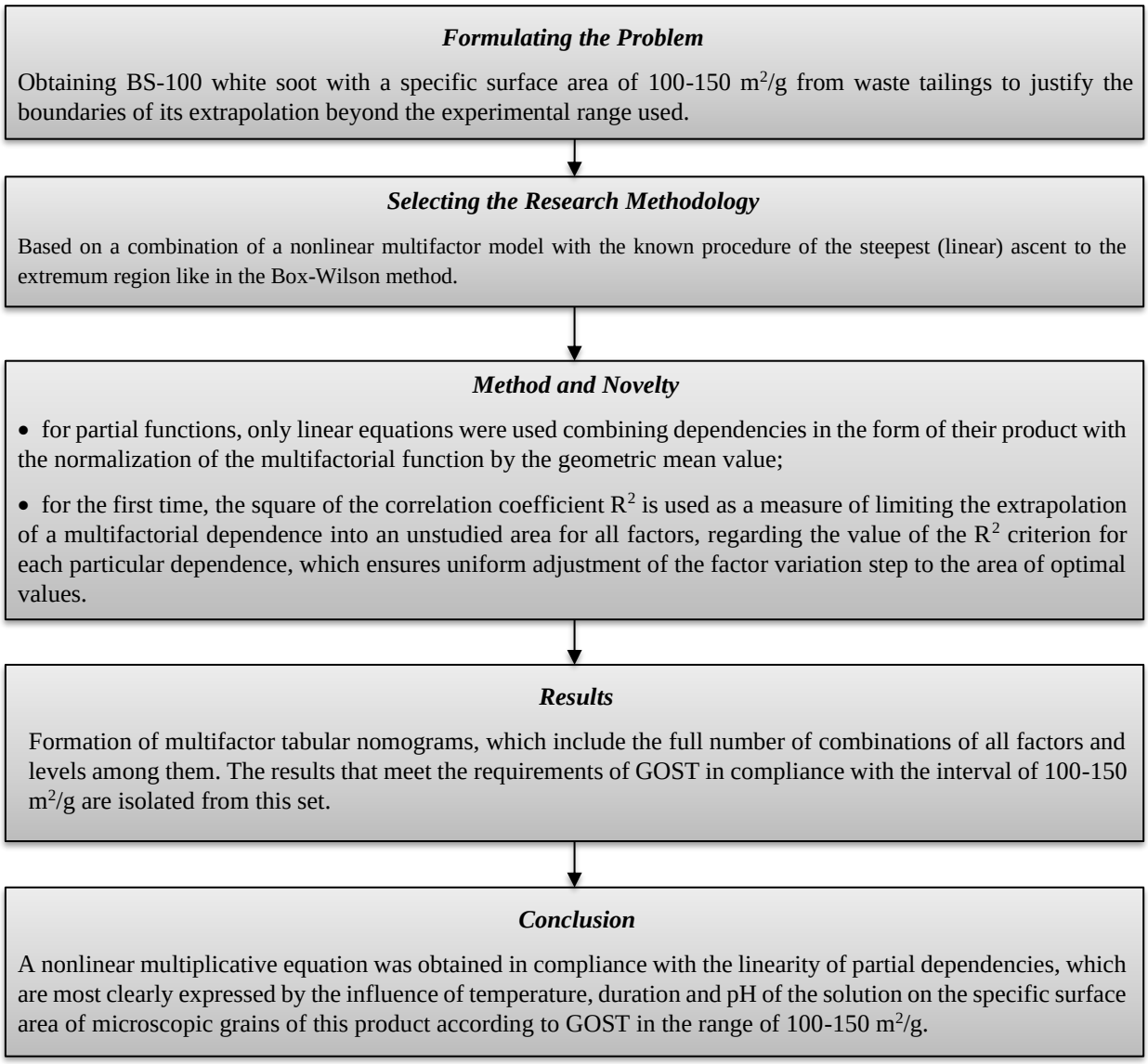


Figure 1. Flowchart of the research process

2. Literature Review

Alkali sintering methods were studied in previous studies [15, 16], and the enrichment of the materials was confirmed. The extraction of silicon from metallurgical waste into a solution can be used for the deposition of white soot. White soot is finely dispersed hydrated silicon oxide containing 85-95% SiO_2 and admixtures of iron, aluminum, magnesium, and sodium oxides. It is in demand as an active mineral filler in tire, rubber, chemical, cosmetic, and other industries. Silicon dioxide is deposited from a solution of sodium silicate (liquid glass) with acid or carbon dioxide, followed by filtration, washing, and drying. The main characteristics of various white soot brands are presented in Table 1 (GOST 18307-78) [17].

Table 1. Main characteristics of white soot

Indicator	BS-30	BS-50	BS-100	BS-120
Mass fraction of silicon dioxide, %, at least	85	76	86	87
Mass fraction of moisture, %, at most	6.5	6.0	6.5	6.5
Weight loss on ignition, %	4.5-7.5	7.0-10.0	5.0-7.0	3.5-7.0
Mass fraction of iron in terms of iron oxide, % at most	Not standardized	0.03	0.15	0.17
Mass fraction of aluminum in terms of aluminum oxide, % at most	Not standardized	0.10	0.15	0.10
pH of aqueous extract				
• For powder white soot	8.0-10.0	9.0-10.5	7.0-8.5	8.0-9.5
• For granular white soot	-	-	-	7.0-8.5
Specific surface area, m^2/g	35±10	45±10	100±20	120±20
By agreement with the consumer, it is allowed to produce white soot BS-100 with a specific surface area of 100-150 m^2/g.				

Depending on the method for producing white soot, the final property of the product is determined by the size and shape of the particles, presence or absence of pores, etc. [18, 19]. The obtained experimental values were processed using the standard methods of probability theory and mathematical statistics. The main approach to solving many problems is the least squares method, which works for linear dependencies, but in practice, nonlinear ones are more common. In this case, approximate nonlinear processing methods were used [20]. The main tasks of regression analysis include establishing the form of dependence, determining the regression function, and estimating unknown values of the dependent variable. To obtain the experimental and statistical functions of objects, mathematical designs of complete factorial experiments (CFE) and fractional factorial experiments (FFE) have been developed [20-23].

Due to the significantly smaller number of experiments compared to CFE designs, FFE designs are widely used in industrial experiments and when it is necessary to study a sufficiently large number of factors with a small number of experiments and determine the factors that have the strongest influence on the property of the factor. If necessary, in an experiment compared to the CFE design, FFE mathematical designs are drawn up depending on the preferred method of statistical analysis of the results. Mathematical experimental design is a constantly improving method; important indicators of the quality of compiled mathematical experimental designs are their orthogonality and optimality [24, 25]. There are different types of compositional designs: three-level Box designs (3k), Box-Wilson design, Box-Hunter design, and Kono design. These designs make it possible to find the regression equation in the following family of polynomials, for example, for coded values of the input factors:

To obtain a quadratic regression equation, it is necessary to have at least three levels for each factor ($ml \geq 3$). Three-level complete factorial experimental designs are known as Box designs. The advantage of Box designs is their high accuracy in determining factor effects. Box designs are not orthogonal and D-optimal. The disadvantages of Box designs also include the large number of experiments, which far exceeds the maximum possible number of coefficients in the quadratic polynomial L (Table 2) [24, 25].

Table 2. Parameters of Box designs

Parameter	Parameter value with the number of input factors *k*				
	2	**3**	**4**	**5**	**6**
Number of experiments $N=3^k$	9	27	81	243	729
L	6	10	15	21	28

Next, the remaining linear regression coefficients are adjusted to simultaneously and uniformly move the expected optimum, for which the existing largest regression coefficient should be reduced and the existing smallest value should be increased. This can be achieved in various ways. Since there is no theoretical justification for limiting the regression

coefficient, further continuation of the Box-Wilson method is provided only in an additional experimental implementation. Special extrapolation methods other than the Box-Wilson method are unknown, including the neural-network-based experimental design method [26]. Therefore, each time it is necessary to adjust the known methods for designing nonlinear experiments, such as the well-known multiplicative [27] and above-mentioned research methods.

In contrast to similar methods, this method is developed only for linear terms of a multifactor regression equation and with the obligatory elimination of regression terms in the form of a product of the first-degree factors; otherwise, the procedure for simultaneously reaching the global extremum region will not be reliable.

This research aims to study the possibility of obtaining a commercial product, BS-100 white soot, in the range with a specific surface area of 100-150 m^2/g using a new method by determining the optimal values of this surface based on combining the known nonlinear multiplicative multifactor model [27] with the known steepest (linear) ascent procedure to the extremum region using the Box-Wilson method.

3. Research Method

To converge with the Box-Wilson method (BWM), by steeply ascending to the region of the global optimum based on a linear regression equation, the following modification of PDED is proposed: for partial functions, only linear equations are used, and their arithmetic mean values are introduced into the multiplicative model to normalize their partial dependencies in dimensionless form. For the first time, the square of the correlation coefficient R^2_i was used as a measure of extrapolation of multifactorial dependence into an unstudied area for all factors simultaneously. This is achieved, as in the BWM, by adjusting the rate of entry into the global extremum region using the value R^2 for each particular dependence rather than by averaging the linear regression coefficients. Extrapolation and optimization processes are detailed in multidimensional tabular nomograms and measured by the number of multifactor cells in localized areas of optimal (acceptable) values. These nomograms can be used to monitor and control the technological process, as illustrated by the example of white-soot production.

Our research is characterized by the possibility of combining linear and nonlinear mappings in multifactorial dependencies, if it is necessary to extrapolate beyond the studied range to obtain separate areas of optimal values of a multidimensional function. The combination method includes the coupling of additive and multiplicative fragments, of which the former are represented by rectilinear partial dependencies normalized by arithmetic mean values, and the latter are represented by the products of partial functions and geometric mean values. The research objective (in contrast to the known ones) is to establish the degree of extrapolation of multifactorial dependence based on the R^2 value, which eliminates the possibility of unjustified extrapolation and unreliable results.

This research was conducted by conducting technological operations with multifactor dependencies within optimal areas and using them, for example, in the event of a random exit from the optimal mode of a technological object and an accelerated return to the optimal mode based on four-factor tabular nomograms. Carbon dioxide was used as a neutralizing agent to separate white soot from a silicate solution. White soot ($mSiO_2 \cdot nH_2O$) was obtained through the two-stage carbonization of a silicate solution (liquid glass) with carbon dioxide in a recirculation system, bringing the pH to 9-10 within 30 min, and then within 60 min, until the residual alkali content in the solution was 90 g/l.

The main reaction for producing white soot with carbon dioxide is as follows:

$$Na_2SiO_3+CO_2 = Na_2CO_3 + SiO_2\downarrow$$

with sedimentation in the form of $mSiO_2 \cdot nH_2O$. At the stage of preliminary desiliconization of the rough concentrate into a solution under arbitrary search conditions, a silicate solution was obtained with the following composition: g/l: Na_2O = 126.5, SiO_2 = 107.7, Al_2O_3 = 3.1. In a series of experiments, carbon dioxide was purged through the solution volume for such a time that the required final pH value (9.5 - 9.8 units) of the pulp was achieved within a technologically acceptable duration.

When conducting experiments according to a special design, the influence of deposition duration (τ, min), temperature (t, °C) and the final pH value of the pulp on the formation of the specific surface area (S_{sp}, m^2/g) of white soot was studied. The resulting sediment was separated by filtration, washed, and dried at 105°C. The specific surface area was calculated from the average diameter of the white soot particles measured using an electron microscope.

The experimental design for the sequential study of the operating factors was implemented using a method that involved setting up a generalizing central experiment with a single experimental point for all factors. This is indicated on the graph of partial functions and is considered when constructing them and determining the correlation coefficient [27, 28]. The resulting partial dependencies regarding the significant functions to describe the set of operating factors were generalized according to Malyshev et al. [27] in the form of their product with normalization by the arithmetic mean experimental value of each function. In all cases, when deriving the equation, to check its adequacy, we used the nonlinear multiple correlation coefficient R and its significance t_R, which are expressed by the following formulas [27, 28]:

$$R = \sqrt{1 - \frac{(n-1)\sum_{i=1}^{n}(y_{e\,i} - y_{d\,i})^2}{(n-k-1)\sum_{i=1}^{n}(y_{e\,i} - \overline{y_{e\,av}})^2}} \quad (1)$$

$$t_R = \frac{R\sqrt{n-k-1}}{1-R^2} > 2 \quad (2)$$

here $y_{e,i}$ is an experimental value; $y_{d,i}$ is a design value; $y_{e,av}$ is an average experimental value; n is the number of independent (non-repeated) experimental data; k is the number of operative factors; $(n\text{-}1)$ is the number of degrees of freedom for reproducibility variance; $(n\text{-}k\text{-}1)$ – the number of degrees of freedom for adequacy variance.

The geometric mean of all experimental values $\overline{y}_{e\,g}$ was introduced into the multifactor equation as a normalizing divisor of the generalized function $y/\overline{y}_{e\,g}$:

$$y = \overline{y}_{e\,g} \prod_{i=1}^{i=n} \frac{y_i}{\overline{y}_{i\,e\,\text{a}}} \quad (3)$$

In its most general form, this equation expresses a nonlinear multiplicative multifactor function and fundamentally differs from a multifactor linear regression equation in the form of a sum of terms. The regression equation has two major drawbacks.

- It is not reset to zero at zero values for any factor, that is, a generalizing function is not identically transformed into a particular function.
- The regression equation does not allow the inversion of variables because of the representation of each variable, not only separately, but also as a product with other factors.

The nonlinear multiplicative multifactor Equation 3 is deprived of these shortcomings because of the obvious algebraic operations involving zero and the transfer of variables on the left or right sides of equality during ordinary algebraic procedures. The mandatory normalization of partial dependencies by their arithmetic mean values is a significant feature of Equation 3. This results in the reduction of the natural dimensions of partial dependencies and their consideration in the form of shares in a unit. Moreover, the generalized multiplicative function y is strictly equal to unity if all the partial dependencies used in the multifactor equation are equal to their arithmetic mean values.

This ensures the possibility of reducing it to unity at the average values of all partial functions $\frac{y}{\overline{y}_{i\,\text{э}}} = \prod_{i=1}^{i=n} \frac{\overline{y_i}_{\text{p}}}{\overline{y}_{i\,\text{э a}}} = 1$. The multiplicative structure of the multifactor Equation 3 considers the influence of each factor according to its deviation from unity: when deviating upward from unity, the influence of this factor increases; when deviating from unity downward, it decreases; if it is equal to unity, the influence of each factor is neutralized.

Determining the optimal conditions for obtaining technological products using multifactor models usually involves searching for extreme values in each partial dependence, with further substitution of the corresponding largest (or smallest) partial optima into the generalizing dependence and obtaining a single multifactor extremum. This practice has taken root since the time when there were no multifactorial mathematical models and the multifactorial process was studied in a sequential transition from the best indicators in the previous partial dependence to the initial conditions in the subsequent one, etc., to the last one.

In fact, it turned out that in this way, it is possible to enter the region of a particular extremum, located arbitrarily far from the global extremum. Special methods have emerged for using multifactor models to determine the shortest paths to reach the region of optimal values of a multifactor function. Most of these methods are based on simplex-lattice progression to an extremum, moving in the opposite direction to the worst conditions (rather than in the continuation of the best). Another method is the so-called "steepest ascent" to the optimum area, based on a multifactorial regression equation with the correlation of the step of progress toward the target result, while simultaneously "stepping" in a general order, implementing the last stage of the "ascent" by staging a step-by-step experimental completion of the process when the indicators begin to deteriorate. In the new experimental design methods, including those based on neural networks, the possibility and conditions for limiting extrapolation were not considered.

The main task of the BWM is to provide simultaneous and uniform step-by-step movements toward the optimum region. This is achieved by adjusting the regression coefficients for each factor after obtaining an incomplete quadratic regression model and removing all the products of the factors as non-linear terms. Thereafter, the remaining linear part of the equation was tested for the significance of each factor by its regression coefficient and the removal of all insignificant factors. Then, the most essential procedure for adjusting the regression coefficient is conducted, as they determine the rate (by an absolute value) and direction of the effects of each factor (increase, plus, decrease, minus) with the code designation of the dimensionality of the variables ($в_i \cdot x_i$=1). Vinarsky and Lurie consider several methods of correlating $в_i \cdot x_i$ to smooth out the impact of strong and weak factors, and among them, normalization of $в_i$ by $в_{max}$, averaged $в_i = \overline{в}$ taking the reciprocal value $1/в_i$, change of sign $в_i \Longrightarrow в_i^2$, and normalization by sum $в_i = в_i / \sum_{i=1}^{i=n} в_i$.

The ambiguity of preparing for the steepest ascent to the region of the global optimum forces to use crushing the adjusted steps to the smallest overall step, especially because the ascent is recommended to be conducted experimentally

before the results begin to deteriorate and repeat the entire procedure, starting with obtaining a new regression equation under the conditions of stopping the ascent to the top and so on until the region of global extremum transforms to a limit that cycles without deterioration.

Thus, it is not easy to implement process control with the linear display of a complex object when a change in linearly acting factors is linearly transmitted to a linearly perceiving multifactorial response function.

Our attempt to apply a similar procedure to nonlinear multifactor objects was based on the complete replacement of the main tool for setting up a uniform and simultaneous ascent to the region of optimal results of regression coefficients with correlation coefficients of particular and generalized dependencies. Compared to regression, correlation seems to be a more general concept, including probabilistic and information aspects and those related to the stability of complex systems. This requires expanding the concept of the optimum and is represented by its multidimensional mapping. At the same time, we tried to adhere to the conditions of subordination of the requirements for linear systems and the basic idea of the steepest ascent – uniform and rectilinear ones (as in Newton's first law), according to the Box-Wilson method.

To combine the steepest (linear) ascent to the optimum region (in this case, for the best S_{sp} values in the GOST interval) with the Box-Wilson method, basic nonlinear model (3) was supplemented with the following restrictions:

- The original partial-point experimental dependencies are approximated only by the equation of straight lines, which are typically used for linear multifactor models in the form of regression equations.
- The basic nonlinear (multiplicative) model is expressed as a product of linear partial functions with normalization by their arithmetic mean values, while the dimensionalities of the partial dependencies for each factor are reduced during normalization and become dimensionless (unit fractions, u.f.).
- To normalize a nonlinear multifactor function using the general average, which is defined as the geometric mean.

These limitations relate to the experimental design to obtain a single multifactor nonlinear equation with fragments of a linear response in combination with nonlinear ones for better adaptability of the combined model to the object of display, in this case, to the technological process for a wider scope of optimal conditions for obtaining a product of a given quality.

In addition to these limitations, an entropy-information justification for the structure of the correlation coefficient (1) and its significance (2) has been added to contain structural *i* and adaptive *h* components normalized by the maximum value of information entropy H_{max} in the form of the law of conservation of the sum of information and entropy [29, 30]:

$$i + h = 1 \tag{4}$$

This supplement to the meaning of nonlinear correlation makes it possible to isolate the additive influence of the structural and adaptive components on the limit of extrapolation and optimization in an expanded multifactor space equal to the proportion of the golden section [31, 32] and in normalized form, not leaving the equality:

$$i + h = 0.618 + 0.382 = 1 \tag{5}$$

This equality is valid for the first level of self-organization of a complex system (n=2) [28].

The values of i=R are obtained from the dependence of the structural component on the degree of coherent (integer) correspondence of the measures of the information (i) and entropy (h) components as a result of the analytical or numerical solution of the equation [28].

$$i^n + i - 1 = 0 \tag{6}$$

When n=2, the superiority of i>h is achieved for the first time, and in all cases $R \geq R^2$, so that neither within the framework of general system laws nor according to purely mathematical properties of the value $0 \leq x \leq 1 \Longrightarrow x^2 < x$,

In this case, this made it possible to more strictly substantiate the need for double consideration of the correlation coefficient, not only to express the real dependencies of the multifactorial empirical connection of data, but also for reliable extrapolation to the same extent in a broader process control mode. In addition, the nonlinear multifactor model acquires additional reliability and degrees of freedom, which can be used for linear models, while simultaneously eliminating the approximate and contradictory interpretation of R^2 in relation to the determinism and functionality of any model.

To verify these equations, it is necessary:

- To use the value R^2 as the permissible fraction σ of the expansion of the studied interval for each factor Δx_i by multiplying it by R^2_i.

$$\sigma_i = \Delta x_i R_i^2,$$

The greater R_i, the greater R^2_i, but it is always within Δx_i. When $t_R \leq 2$, the correlation coefficient was considered insignificant; that is, when R_{min}=0.

$\sigma_{i\,min} = \Delta x_i R^2_{i\,min} = 0,$

- With R_i>2, the share of σ_i is distributed equally between the left and right boundaries of the range of values x by adding or decreasing x by $\Delta x_i \cdot R^2_i/2$, depending on the direction of extrapolation $\sigma_i = \pm\Delta x_i R^2_i/2$

toward increasing or decreasing y_i *in proportion to* x_i, which is a consequence of the primary processing of dependencies on the equation of straight lines $y=ax+b$. For nonlinear functions, their behavior during extrapolation can be greatly distorted.

- Extrapolation for all factors is not conducted immediately over the entire range but with an adjustment for an increase from 50 to 100% of the full share.

To prepare particular dependencies and a multifactor model together for extrapolation procedures, the initial data should be tabulated in the order in which they are sequentially filled out.

The results of extrapolation and optimization of acceptable options for quality indicators (S_{sp} = 100-150 m^2/g of white soot) were in a wider range of combinations of operating parameters of the white soot production process.

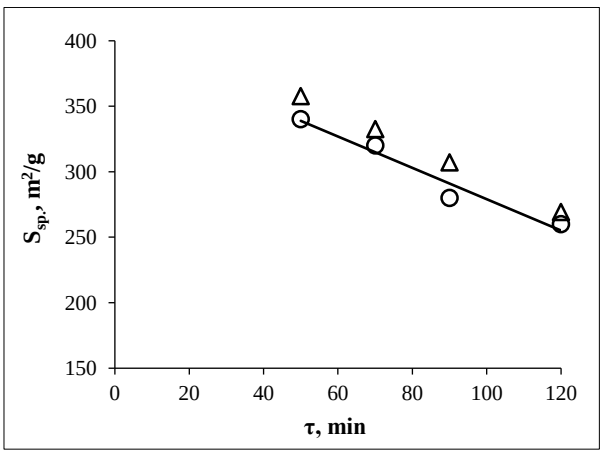
4. Results and Discussions

The results of the experiments using the sequential study of factors after conducting a central experiment are presented in Table 3.

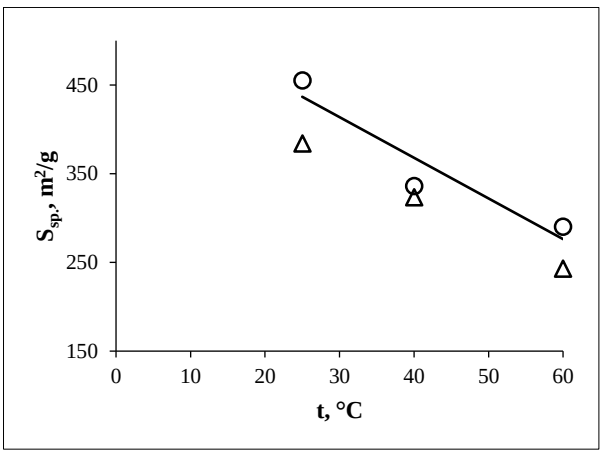
Table 3. Experimental data on the influence of pH, leaching duration and temperature on the specific surface area of white soot

Factor under study Experimental conditions		$S_{sp.}$, m^2/g exp.	$S_{sp.}$, m^2/g for partial functions	$S_{sp.}$, m^2/g by Eq. 7
τ, min. (C_{Na_2O}– 126.5 g/l, t– 40°C)	50	340	338.9	357.7
	70	320	314.9	332.4
	90	280	291.0	307.1
	120	260	255.1	269.3
t, °C (τ–60 min., C_{Na_2O}– 126.5 g/l)	25	455	436.7	383.7
	40	330	368.0	323.3
	60	290	276.3	242.8
pH, units (t–40°C, τ – 60 min., C_{Na_2O}– 126.5 g/l)	9.5	125	168.5	178.6
	9.7	***316.1***	***255.3***	270.6
	10.2	455	472.4	500.7

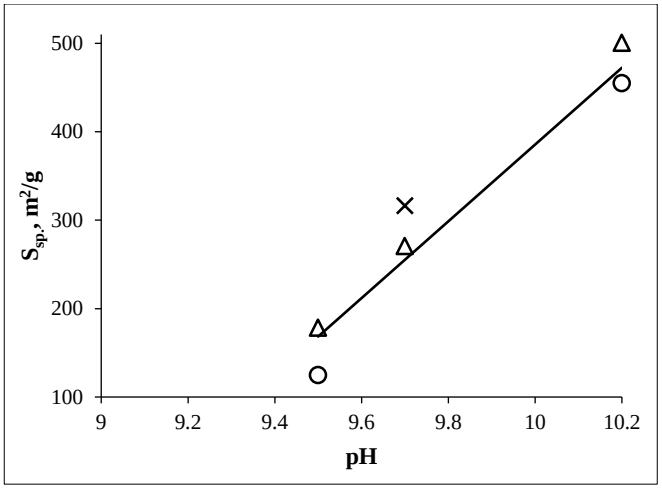
The processing of the experimental results on the equation of a straight line (using the least-squares method) to identify significant linear functions is presented in Figure 2. Crosses represent the coordinates of the central experiment. In all experiments, the alkalinity was adjusted at the final stage to a level of 126.5 g/l. The obtained linear partial dependences of the specific surface area with the determination of the nonlinear multiple correlation coefficient R and its significance t_R are listed in Table 4.



(a)



(b)



(c)

Figure 2. Influence of various carbonization factors on the specific surface area of white soot: (a) duration, min; (b) temperature, °C; (c) solution pH (developed by the authors). *Note*: ○ – experimental data; × – average value for all partial functions; Δ – according to Equation 7.

Table 4. Particular functions of the specific surface area ($S_{sp.}$, m²/g) of white soot for τ, *t* and pH of the solution.

Functions	*R*	t_R
$S_{sp,} = (398{,}7 - 1{,}196 \cdot \tau)$	0.967	21.21
$S_{sp,} = (551{,}3 - 4{,}58 \cdot t)$	0.887	4.166
$S_{sp,} = (434{,}2 \cdot pH - 3956)$	0.958	11.73

Based on experimental data, partial equations were obtained (Table 4), which were used to derive a mathematical model [27] for the specific surface area of white soot. Partial equations were generalized in the form of their product with normalization to the average experimental value (in this case, m²/g: 300.0 for τ =60 min; 360.3 for t=40°C; 298.8 - pH=9.7 units). The generalized equation for the specific surface area is expressed as:

$$S_{sp} = 9.335 \cdot 10^{-6} \cdot (398.7 - 1.196 \cdot \tau) \cdot (551.3 - 4.58 \cdot t) \cdot (434.2 \cdot pH - 3956) \quad (7)$$

The nonlinear multiple correlation coefficient was *R*=0.843, its significance was t_R =6.001>2 (calculated by equations 1 and 2), and based on the geometric mean value $\overline{y}_{e,a}$ *R*=0.847, t_R=6.249> 2. A comparison of the experimental data on the specific surface area of white soot and those obtained from Equation 7 is shown in Figure 2 (triangles).

Generalized Equation 7 makes it possible to identify the joint influence of the existing factors. Moreover, not exceeding the permissible values for the influence of factors can be ensured by many combinations of specified levels rather than by what was used during the central experiment under average conditions of the recorded factors. This can be demonstrated using a multifactor tabular nomogram, which is presented in Tables 5 and 6. Each mini-rectangle (cell) displays a multifactor nonlinear function with a flat multifactor approximation for each of the common coordinate systems. Isolating groups of cells according to the condition of being in the permissible (allowed) limits of the objective function with varying accuracy creates an image of optimal behavior and control of the object. Table 5 shows the initial data for an allowable extrapolation of the multifactor model into the unexplored area of the factor space regarding R^2 for partial functions (Table 4). Moreover, R^2<*R* for the entire range *R* from zero to one.

Table 5. Input data for allowable extrapolation of a multifactor model into unexplored areas of factor space.

Factor	R_i *(Table 4)*	R_i^2	Δx_i	$\Delta x_i \cdot R^2_i/2$	0% σ	for (7)			
						25% σ	50%σ	75% σ	100%σ
τ, min.	0.967	0.935	70	32.72	120	128.2	136.4	144.5	152.7
t, °C	0.887	0.787	35	13.77	60	63.44	66.88	70.33	73.77
pH, un.	0.958	0.918	0.7	0.321	9.5	9.419	9.339	9.259	9.179

Table 6. Dependence of the specific surface area of white soot on various factors (an interval of 100-150 m^2/g is highlighted)

τ, min.	*t, ℃*	*pH*				
		9.259	9.339	9.5	9.7	10.2
50	25	88.2	136.2	232.8	352.8	652.7
60		85.1	131.4	224.5	340.3	629.6
70		82.0	126.6	216.3	327.9	606.6
90		75.8	117.0	199.9	302.9	560.5
120		66.4	102.5	175.2	265.6	491.4
136.4		61.3	94.7	161.8	245.2	453.6
144.5		58.8	90.8	155.1	235.1	434.9
50	40	74.3	114.8	196.1	297.2	549.9
60		71.7	110.7	189.2	286.7	530.5
70		69.1	106.7	182.3	276.2	511.1
90		63.8	98.6	168.4	255.3	472.3
120		55.9	86.4	147.7	223.8	414.0
136.4		51.7	79.7	136.3	206.6	382.2
144.5		49.5	76.5	130.7	198.1	366.5
50	60	55.8	86.2	147.3	223.2	412.9
60		53.8	83.1	142.0	215.3	398.3
70		51.9	80.1	136.9	207.4	383.7
90		47.9	74.0	126.5	191.7	354.6
120		42.0	64.9	110.9	168.0	310.9
136.4		38.8	59.9	102.3	155.1	286.9
144.5		37.2	57.4	98.1	148.7	275.2
50	66.88	49.4	76.3	130.5	197.7	365.8
60		47.7	73.6	125.8	190.7	352.9
70		45.9	70.9	121.2	183.7	339.9
90		42.5	65.5	112.0	169.8	314.1
120		37.2	57.5	98.2	148.8	275.4
136.4		34.4	53.0	90.7	137.4	254.2
144.5		32.9	50.9	86.9	131.7	243.8
50	70.33	46.2	71.4	122.0	184.9	342.1
60		44.6	68.9	117.7	178.4	330.1
70		43.0	66.4	113.4	171.9	318.0
90		39.7	61.3	104.8	158.8	293.8
120		34.8	53.7	91.9	139.2	257.6
136.4		32.1	49.6	84.8	128.5	237.8
144.5		30.8	47.6	81.3	123.2	228.0

The results of using the proposed method for extrapolation of the nonlinear multiplicative model are presented in Table 6 and are characterized by the following features:

- The three-dimensional spaces of the three-factor multiplicative model presented here are projections onto a plane in the form of three-dimensional cells, densely packed, and filling the entire space, including the extrapolated part for each factor.
- The cells are a discrete set defined by the number of levels of each factor, including those in the extrapolation region. The multiplicative combination of the levels of all factors creates an enumerable set of results, which are selected by compliance or non-compliance with any optimality feature, in this case, by being in the interval specified by compliance with the quality of white soot within 100-150 m^2/g. For clarity, the cells corresponding to the optimality conditions are highlighted in color or graphically.

- Cells can be grouped into more or less extensive localizations, while due to the complete distinguishability of cells, which is ensured on the basis of the combination formula, the resulting diversity of system elements turns out to be proportional to its stability and preference for a greater probability of self-organization processes.
- Enumerability of the localization volume by the number of cells in them creates opportunities for primarily assessing the effectiveness of permissible extrapolation, which can be performed by the ratio of optimal cells with and without regard to permissible extrapolation. The same table makes it possible to determine the contribution of each factor to the implementation of permissible extrapolation, limiting it to the projection of the total number of optimal cells in the range of levels used for each factor.

These possibilities are exemplified by the effect of the pH factor (Figure 2), which turned out to be extreme despite the specified linearity of partial functions and is explained by the opposite direction of pH and τ, pH, and t. This creates the realism of the formation of a global extremum; in this case, its formation or non-formation is entirely determined by the ratio of gradients in partial dependencies by their sign and magnitude, and is not predictable in advance. It is simplified by means of nomograms similar to those in Table 6, which record the as-yet undisclosed patterns of location distribution by the number of cells in them, including single ones. Therefore, regarding the possibility of controlling the extremum formation process and returning to it along a specific trajectory, we have doubts that are remotely similar only to a probabilistic solution and submission to the Poincaré recurrent theorem to avoid the homonymous "catastrophe". In this regard, we consider it possible to provide only a preliminary recommendation on the use of Table 6 for moving from one mode to another in a purely graphical manner. This way is determined by the fact that each cell on this surface is surrounded by three other cells with other indicators of specific surface area. To move the process from a non-optimal region to an optimal location, it is necessary to select one of the three cells with an interval closer to the optimal one in the direction of the nearest location. Thus, a transition from the initial position to the next cell in the direction of movement to the nearest location occurs. However, because of the random orientation of the sides of the elementary rectangle relative to the location direction, the transition from one cell to another occurs at a random angle. As a result, the cell will be directed to the location along a broken trajectory, maintaining the aim at the same location. After the process transitions to a new cell, all procedures are repeated until the cell is on the border with the location and becomes part of it.

As shown in Table 6, not exceeding the permissible specific surface area of white soot can be ensured at pH values of 9.339-9.7 by reducing the carbonization duration, at which the rate of formation of solid-phase nuclei decreases with insufficient volume of gas transmission. However, at pH-9.5, the optimal specific surface area was obtained with increasing solution temperature. In any case, this multifactorial dependence can be used to control the process when obtaining a commercial product under real conditions, with the adjustment of existing factors depending on their changes for technological reasons. It is especially important to keep in mind the large areas where the optimal final pH values are exceeded (Figure 3). The number of three-factor cells of the specific surface area of white soot in the zones of optimal results in accordance with the BS-100 brand, under GOST 18307-78, with the best coordinates of pH = 9.5 and G = 17 units.

Figure 4 shows the morphology of the white soot grains with a specific surface area of 100 m^2/g obtained under optimal conditions.

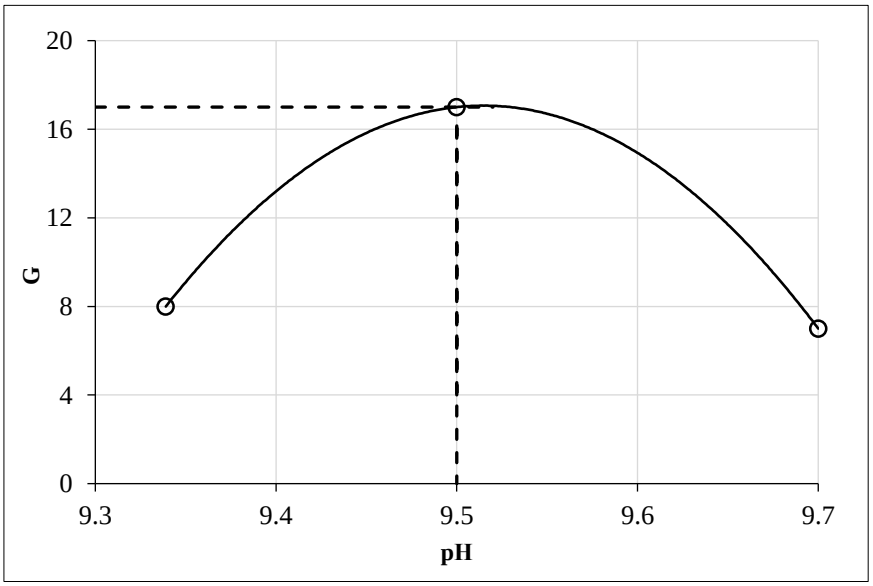


Figure 3. Dependence of the number of cells in zones of optimal areas (G, units) on the pH of the solution

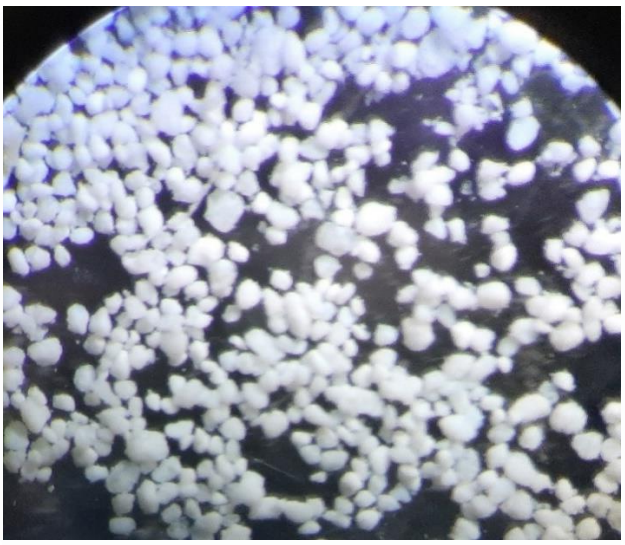


Figure 4. The morphology of white soot grains

In real-world conditions, going beyond the GOST interval implies receiving defective products.

In this case, prompt decisions are required, and such a multifactor nomogram can be used as a flow chart for a gradual exit from the current situation along the most optimal path without a sharp change in the process mode. To do this, a well-known method should be used for moving to the optimum according to the principle of transition along the nearest path with a sequential transition from cell to cell toward the nearest optimum. The shortest path in terms of the number of cells on the broken transition line should be chosen from the various transition paths analyzed.

5. Conclusion

Based on the analysis of the advantages and disadvantages of the Box-Wilson method of the "steepest ascent to the region of a multifactor optimum using an exclusively linear version of the regression equation, a modification of the nonlinear multiplicative Protodyakonov equation (PE) was proposed for uniform and simultaneous extrapolation of a multivariate function regarding the nonlinear multiple correlation coefficient R for each factor and the function as a whole. The PE-based modification of the method consists of representing partial functions only in the form of a straight-line equation with normalization by the arithmetic mean, and their generalization in the form of a product by the geometric mean as part of a multidimensional function. Extrapolation uniformity and simultaneity were ensured by considering the degree of correlation of each partial function in the form of R^2, which limited the admissibility of expanding the initial range of arguments.

The display of multidimensional areas of allowable and non-allowable values of the generalized function is provided by their localization in the plane of visual representation of a maximum of four-factor tabular nomograms, with the possibility of a four-factor review of any n-dimensional space in any sequence.

The proposed modification of the probabilistic deterministic method and experimental design (PDED) makes it possible to most simply determine and unify the procedure for extrapolation and optimization of complex processes, which is exemplified by the technological process for white soot ($mSiO_2 \cdot nH_2O$) production within the permissible content required by GOST. This is of independent importance, because of the demand for this product in several industrial production areas. A multifactor tabular display of a generalizing function with localization of its equivalent values $y \pm \sigma$ appears to be the most visual and easily formalized for computer execution. This finding may be worthy of independent theoretical and practical research.

As applicable to the process of producing white soot of a set quality, a nonlinear multiplicative equation was obtained with straight dependencies on temperature, duration, and pH of the solution of the specific surface area of microscopic grains of this product according to GOST ranging from 100 to 150 m^2/g.

The proposed procedure for extrapolation and optimization of the calculation result using multifactorial dependence made it possible to establish the greatest localization (in terms of the number of multifactorial elementary cells equivalent in product quality). In this case, the choice of the greatest diversity in the implementation of distinguishable microstates (cells) of the system complies with the combinatorial concept of entropy (a measure of the diversity of the system), which tends to reach a maximum spontaneously.

5.1. Strengths and Limitations of the Research

Thus, the main advantage of the proposed method and extrapolation procedure as applicable to a nonlinear multifactorial dependence is the guarantee of a directly proportional transfer of the permissible degree of expansion of the interval of change in the argument to the corresponding partial function. The localization of optimal areas, rather than the single best value of a multifactorial function, is guaranteed at the cost of refusing to consider the nonlinearity of partial functions and the deterioration of the correlation coefficient up to the exclusion of clearly extreme partial dependencies. This may have an independent significance for the study of objects, admitting the existence of equivalent values that determine any of its qualities. Strictly speaking, any value of a multifactor function, including the global extremum, can correspond to an infinite number of factor level combinations with an infinitesimal difference between them. Nevertheless, the developed modification of the PDED method has serious limitations in the form of partial dependencies and can be classified as a disadvantage.

5.2. Recommendations and Further Research

Further research is related to the complex processing of waste copper tailings according to the flowchart of the final product. It is planned to conduct a multifactor analysis of technological operations by identifying the zones of the optimal process modes.

6. Declarations

6.1. Author Contributions

Conceptualization, L.K. and G.M.; methodology, V.M.; software, G.M.; validation, V.M.; formal analysis, Y.Kh.; investigation, Y.K.; resources, Y.K.; data curation, V.M.; writing—original draft preparation, L.K. and G.M.; writing—review and editing, V.M., Y.K., and Y.Kh.; visualization, G.M.; supervision, L.K.; project administration, L.K.; funding acquisition, L.K. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

Data sharing is not applicable to this article.

6.3. Funding

This study was conducted under the grant project AP 19675340, funded by the Committee of Science of the Ministry of Education and Science of the Republic of Kazakhstan.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

7. References

[1] Amin, B., & Golla, R. (2023). Linear and Multiplicative Maps under Spectral Conditions. Functional Analysis and Its Applications, 57(3), 179–191. doi:10.1134/s0016266323030012.

[2] Kolenchukov, O. A., Bukhtoyarov, V. V., Shvetsov, N. S., Sergienko, R. B., & Bashmur, N. A. (2023). Forecasting useful life of pumping equipment elements based on multivariate regression approaches. Science & Technologies Oil and Oil Products Pipeline Transportation, 13(5), 456–465. doi:10.28999/2541-9595-2023-13-5-456-465.

[3] Bazilevskiy, M. P. (2024). Algorithms for constructing non-elementary linear regressions using forward selection method. Proceedings of Voronezh State University. Series: Systems Analysis and Information Technologies, (1), 104-113. doi:10.17308/sait/1995-5499/2024/1/104-113.

[4] Orlov, A. I. (2024). The paradigm shift in mathematical methods of research. Industrial Laboratory Diagnostics of Materials, 90(7), 5–7. doi:10.26896/1028-6861-2024-90-7-5-7.

[5] Aben, E., Yussupova, S., Akhmetkanov, D., Yelzhanov, E., & Sarybayev, N. (2024). Research into Uranium Characteristics and Content in a Pregnant Solution During Leaching with Oxygen Saturation. Civil Engineering Journal, 10(5), 1606–1615. doi:10.28991/cej-2024-010-05-016.

[6] Khomyk, I., & Obelovska, K. (2024). Model of the truck management process using innovative technologies. Information Technology Computer Science Software Engineering and Cyber Security, 2, 142–146. doi:10.32782/it/2024-2-18.

[7] Bavrina, A. P. & Borisov, I. B. (2021). Modern rules for application of correlation analysis. Medical Almanac, 3(68), 70–79.

[8] Grjibovski, A. M., Ivanov, S. V., & Gorbatova, M. A. (2017). Correlation analysis of data using Statistica and SPSS software. Science and Healthcare, 1, 7–36.

[9] Egoshin, V. L., Ivanov, S. V., Savvina, N. V., Kalmakhanov, S. B., Zhamaliyeva, L. M., & Grjibovski, A. M. (2018). Analysis of continuous data using R. Human Ecology, 25(11), 51–64. doi:10.33396/1728-0869-2018-11-51-64.

[10] Bavrina, A. P. (2020). Basic concepts of statistics. Medical Almanac, 3(64), 101–111. (in Russian).

[11] Bavrina, A. P. (2020). Modern rules for the use of descriptive statistics methods in biomedical research. Medical Almanac, 2(63), 95–104.

[12] Gogtay, N. J. & Thatte, U. M. (2017). Principles of correlation analysis. Journal of The Association of Physicians of India, 65, 78-81.

[13] Egoshin, V. L., Ivanov, S. V., Savvina, N. V., Ermolaev, A. R., Mamyrbekova, S. A., Zhamaliyeva, L. M., & Grjibovski, A. M. (2018). Correlation and simple regression analysis using R. Human Ecology, 25(12), 55–64, doi:10.33396/1728-0869-2018-12-55-64.

[14] Nghia, N. H., Zenitova, L., Dien, L. Q., & Truyen, D. N. (2019). The method of obtaining amorphous nanosized silicon dioxide from rice production waste. Ecology and Industry of Russia, 23(4), 30–35. doi:10.18412/1816-0395-2019-4-30-35.

[15] Shoppert, A., Loginova, I., Chaikin, L., & Rogozhnikov, D. (2017). Alkali fusion-leaching method for comprehensive processing of fly ash. KnE Materials Science, 2(2), 89-96. doi:10.18502/kms.v2i2.952.

[16] Shoppert, A. A., Karimova, L., & Zakharyan, D. (2018). Novel method for comprehensive processing of low-grade copper concentrate. Solid State Phenomena, 284, 856–862. doi:10.4028/www.scientific.net/ssp.284.856.

[17] GOST 18307-78. (2024). White soot. Specification. Available online: https://docs.cntd.ru/document/1200017789 (accessed on July 2024).

[18] Bochevskaya, Y. G., Karshigina, Z. B., Sargelova, E. A., & Abisheva, Z. S. (2017). Deposition of amorphous silicon dioxide from silicate solutions obtained after processing of mineral high-silica ore. Bulletin of Science and Education, 36, 18–23. 10.20861/2312-8089-2017-36-006.

[19] Kutishcheva, E. S., Usoltseva, I. O., & Perederin, Yu. V. (2021). Methods for producing highly dispersed silicon dioxide. Polzunovskiy Vestnik, 2, 188–193.

[20] Glukhikh, V. V., Shkuro, A. E., Artemov, A. V., Shyshlov, O. F., & Krivinogov, P. S. (2023). Mathematical Design of Experiments and Analysis of Their Results Using Computer Programs: Textbook. Ural State Forestry University Press, Ekaterinburg, Russia.

[21] Semenov, S. A. (2021). Planning and Processing the Results of an Experiment: A Textbook. RTU MIREA, Moscow, Russia. Available online: https://e.lanbook.com/book/176518 (accessed on July 2024). (in Russian).

[22] Pen, R. Z. & Pen, V. R. (2022). Statistical Methods for Mathematical Modeling, Analysis and Optimization of Technological Processes, St. Petersburg, Russia. (in Russian).

[23] Ligotskiy, D., & Argimbaeva, K. (2024). The sectional formation technology of an anthropogenic deposit with its subsequent mining using a hydraulic pull shovel. Sustainable Development of Mountain Territories, 16(1), 111–121. doi:10.21177/1998-4502-2024-16-1-111-121.

[24] Kokurina, G. N. (2019). Mathematical Methods for Processing Experimental Data: A Textbook. Ivanovo State Chemical Technology University, Ivanovo, Russia. (in Russian).

[25] Makarov, R. I. & Khorosheva, E. R. (2021). Methods for Data Analysis: A Textbook. Vladimir State University named after Alexander and Nikolay Stoletovs (VLSU) Press, Vladimir, Russia, 2021. (in Russian).

[26] Slavutskii, L. A. & Slavutskaya, E. V. (2021). Choosing a neural network structure for signal processing as an experiment planning. Bulletin of the Chuvash University, 123–132, doi:10.47026/1810-1909-2021-3-123-132.

[27] Malyshev, V. P., Makasheva, A. M., Bekturganov, N. S., & Kazhikenova, S. Sh. (2019). Probabilistic-deterministic mapping as a natural scientific basis for understanding, creating and controlling technological processes. Collection of proceedings of the International Scientific Practical Conference "Innovation in the field of natural sciences as the basis for export-oriented industrialization of Kazakhstan." 4-5 April 2019, Almaty, Kazakhstan. (in Russian).

[28] Malyshev, V. P. (2008). Kinetic and technological analysis of generalizing mathematical models of chemical and metallurgical processes. Reports of the National Academy of Sciences of the Republic of Kazakhstan, 2, 13–18. (in Russian).

[29] Soroko, E. M. (1984). Structural Harmony of Systems. Science and Technology, Minsk, Belarus. (in Russian).

[30] Petrukhnova, G. V., Boldyrev, I. R., Gorina, M. O., & Gruzdev, D. A. (2024). Data Collection System. Bulletin of Voronezh State University, 1(20), 20–24. doi:10.36622/1729-6501.2024.20.1.003.

[31] Gamova, N. A. & Sirotkin, V. V. (2021). Manifestation of the golden section. Trends in the Development of Science and Education, 74, 66-70. doi:10.18411/lj-06-2021-310.

[32] Lavrus, V. (2000). Golden Section. Electronic Library: Science and Technology. Available online: https://n-t.ru/tp/iz/zs.htm (accessed on July 2024). (in Russian).