

# **Eurasian Journal of Soil Science**



Journal homepage : http://fesss.org/eurasian\_journal\_of\_soil\_science.asp

## Mathematical models of fertility for the soils of Azerbaijan

## Amin Ismayilov <sup>a,\*</sup>, Fariz Mikailsoy <sup>b</sup>

<sup>a</sup> Azerbaijan National Academy of Sciences, Institute of Soil Science and Agrochemistry, Baku, Azerbaijaan <sup>b</sup> Iğdır University, Faculty of Agriculture, Department of Soil Science and Plant Nutrition, Iğdır, Turkey

#### Abstract

Article Info

Received : 15.07.2014 Accepted : 19.01.2015 The experience shows that when constructing soil fertility models, many researchers prefer single-valued regression analysis. This is primarily due to the fact that regression analyses require simpler statistical calculations, and on the other hand, regression equations enable a physical explanation of the process under study. The research goal is to determine the effect of soil fertility indices and mineral fertilizers on the yields of crops (cereals) grown in the Karabakh Steppe on gray-brown irrigated soils.

Keywords: mathematical modeling, soil fertility, regression analysis, adequacy of the model

© 2015 Federation of Eurasian Soil Science Societies. All rights reserved

### Introduction

The pace of modern life demands from science quick integration of previously accumulated knowledge about fertility and identification and filling the existing gaps. One of the ways of this generalization is fertility models which enable to arrange and structure the considerable body of factual data. The requirement for a radical increase in the productive capacity of soils and soil cover also calls for scientific interpretation and integration of the available experience in the form of mathematical models of fertility.

The issues of soil fertility study and the practical results of crop yield analysis have always been the focus of researchers. The analysis of the factors affecting crop yielding capacity and forecasting the impact of those factors on crop yields are still topical issues. Soil fertility model is a combination of experimentally determined fertility indices which are in close correlation with crop yield value. Fertility model is developed for specific soil, climatic and technological conditions of crop cultivation. The main factors of soil fertility are as following: humus content, phytosanitary condition of the soil (the occurrence of weeds, pests and crop pathogens), arable layer thickness, particle size composition, structure, labile nutrient content, and soil solution reaction.

A fertility model may be a tool for programming crop yields for the specific conditions of each field of a crop rotation on a farm. It is known that soil fertility simulation uses the results of long-term studies. The models reflect the soil properties and soil regime parameters, agro-climatic indices, the series of land reclamation measures and different fertility levels (low, medium and high) including all changes in soil regimes and properties according to those levels. Our analysis shows that regardless of the differences in the taxons, natural and economic regions, soil-cover structure, etc., when developing fertility models in agro-ecosystems, all models are mainly designed to address the following objectives:

a) determining the indices of predicted high soil fertility level in an agro-ecosystem;

Tel.: +994502229889

ISSN: 2147-4249

<sup>\*</sup> Corresponding author.

Azerbaijan National Academy of Sciences, Institute of Soil Science and Agrochemistry, Baku, Azerbaijaan

- b) determining the current state of soil fertility in an agro-ecosystem;
- c) determining quantitative characteristics, internal and external relations of fertility model parameters in the specific conditions of an agro-ecosystem. It should be noted that in model development, researchers widely use regression analysis methods (Frid, 1990).

The research goal is to determine the effect of soil fertility indices and mineral fertilizers on the yields of crops (cereals) grown in the Karabakh Steppe on gray-brown irrigated soils.

### **Material and Methods**

In fertility model development, the crop being grown and its yielding capacity are mainly taken into account. At the same time, it enables using efficient techniques for determining the parameters of the mathematical model. There are many known soil fertility models developed for different soil and climatic conditions and based on the existing theoretical and practical knowledge about fertility indices. Some researchers propose fertility models consisting of various block-diagrams which are used to develop the measures of fertility improvement and management. When developing soil fertility models, for a complete description of the object investigated, it is important to determine the minimum parameters which reflect the main properties of the object being modeled. At the same time, the adequacy of the model being developed is directly dependent on the number of parameters being determined and the accuracy of those parameters.

The experience shows that an excessively large number of parameters (more than 15-20) is not always useful for more accurate description of the dynamical processes in the simulated objects. The determination of one or several interrelated parameters which reflect the object's properties enables describing the other indices.

Along with the above, it should be noted that soil data have some specific features and it is somewhat difficult to take them into account in the modeling process. The examples of such features may be horizontal and vertical heterogeneity, varying boundaries, dynamic nature of soil characteristics, etc. Failure to take into account those features in simulation results in the complexity of the model or obtaining erroneous results.

Thus we are led to a restricted model which covers a specific soil type and is applicable for a specific region only. In this case, several models are developed for one region and they differ in quality and quantity of soil parameters.

It should be noted that in model development researchers widely use regression analysis methods. On the one hand, this is due to the simplicity of regression analysis application, and on the other hand, regression equations directly describe the result. It should be noted however that without a preliminary analysis of the factors affecting the studied object of an agro-ecosystem the use of regression analysis methods creates some difficulties, some of which are listed below:

- a regression model, as other mathematical models, while reflecting the basic properties of the studied agro-ecological process, is not able to fully reproduce its behavior. The fact is that in this case the model being developed gives incomplete information about the physical trajectories of the process change, and consequently the model is not able to evaluate the effects of all factors, alone and in combination, on the resulting character (crop yield);
- data support is critical in the study of soil fertility models;
- a comparative analysis of informative indices may reduce or optimize the number of those indices thus easing the mathematical description of the fertility model;
- the application of fertility models developed for specific crops on large land areas results in significant deviation of the actual result from the predicted one;
- the lack of a unified concept of fertility model development in terms of mathematical analysis and a large number of variables taken into account restrict wide application of the models;
- soil indices are determined by laboratory tests of soil samples and field measurements. Therefore, mathematical processing of the obtained data is of a relative nature. Inevitable measurement errors, either systematic or random, lead to the deviations from the expected results. This is indicative of the imperfection of the regression model being used.

All the above proves the topicality of the issue and the need for using the minimum (optimum) number of unified soil parameters when developing an optimum and adequate mathematical model of soil fertility.

In recent years a number of soil fertility models for different soil and climatic zones of the Republic of Azerbaijan has been developed (Nuriyeva, 1994). It should be emphasized that the lack of a unified concept of soil fertility model development, the variety and multitude of factors taken into account, and, finally, different approaches of the researchers to this issue, create some difficulties for wide application of the models. Due to the lack of a common concept in soil fertility modeling, each author with the individual approach independently determined the range and number of factors to be taken into account. Sometimes fertility models with different factors were proposed for the same region (Poluektov, 1991).

Thus, the lack of common fertility modeling methodology and informative (descriptive) nature of the existing models proves the need for new scientific approaches. Along with the disadvantages, it should be stated that the fertility modeling studies have accumulated a sufficient amount of data for further scientific data integration and the identification of possible regularities by means of mathematical analysis. It is obvious that the development of optimum, adequate and simple models using the minimum number of the same factors is a topical issue of scientific and practical importance (Semenov, 1982).

The description of natural objects' dynamics is based on the idea of their systemic organization. Mathematical modeling is one of the basic tools of system analysis enabling in some cases to avoid timeconsuming and expensive field experiments. Modeling is a recognized means of knowing reality. The process consists of two major stages: model development and the analysis of the developed model.

Statistical models are most commonly used in mathematical soil fertility modeling. Statistical models are constructed on the assumption that the process being studied is a random process and may be investigated by means of statistical methods of system analysis (Khomyakov, 1996). These include: empirical and dynamic statistical models, correlation and factor analysis, and time series analysis.

Empirical models of soil fertility are mainly represented by the so-called production functions which are the regression equations relating the final result (crop yield and its quality indices) with the acting factors. The production functions should meet a number of requirements: the model should take into account the main factors affecting the crop yield, cover a wide range of factor values, and the approximating function should maximally correspond to the real biological regularities.

The construction of a multiple regression model (or multifactor empirical model) is in finding the relation among and several indices y and  $x_1, x_2, x_3, ...,$  etc., i.e. it is defined how the variation of the indices

 $x_i (i = 1, 2, ..., m)$  will affect the value of y.

**The construction of a multiple regression equation begins with a decision on model specification.** With regard to a multiple regression, prior to defining the model type, the factors should be selected. The factors included in the model should be as following:

1) they should be quantifiable.

2) the factors should not be intercorrelated and all the more be in perfect relationship. If the factors are highly correlated, it is impossible to determine their individual effect on the resulting index and the regressors are not interpretable.

In a general way all empirical models may be written symbolically (Mikayilov, 2014):

$$\mathscr{Y} = f(x_1, x_2, ..., x_m; a_1, a_2, ..., a_m)$$
(1)

where  $\tilde{y}$  is the studied property of the environment (dependent variable),  $x_i$  is the environmental factors (independent variable),  $a_j$  is the coefficients of empirical models (i.e., those of regression), and m is the total number of the analyzed factors.

The numerical values of  $a_j$  parameters in (1) are chosen with the best possible fit condition of the theoretical (calculated according to the formula (1)) and the experimental data. In this case the more observations have been conducted, the more redundant information is available and the more accurate smoothing is (Methods of Mathematical Biology, 1982).

To uniquely determine the regressor values of (1), the sample size *n* should not be less than the number of regressors, i.e.  $n \ge m+1$ . Otherwise, the regressor values cannot be determined uniquely.

Among the advantages of empirical models there are reasonably good formal computer ways of identifying (enumeration of equations) different model structures; those models are very easy for calculations.

The disadvantage of these models is the impossibility of taking into account the cause-effect relations between the variables, and environmental hypotheses. It is common for empirical models to have a small number of input values ( $x_i$ ) which reflect the action of the environmental factors; this implies low accuracy of these models. Another and the most important disadvantage is that empirical models do not reveal the mechanisms of the phenomenon being studied and therefore they cannot be used under the conditions different from those they were constructed under (Pachepskiy, 1992).

The construction of a multiple regression equation begins with a decision on model specification. The core of the problem involves two arrays of issues: the selection of factors and the type of regression equation. Due to the clear interpretation of the values, linear, parabolic, exponential, power, exponential-and-power, irrational types, etc., are most widely used.

Linear multiple function is most commonly used to analyze the relationship:

$$\tilde{y} = a_0 + \sum_{i=1}^m a_i x_i \tag{2}$$

In the linear multiple regression the values  $a_i$  at  $x_i$  are referred to as net regression coefficients. The value of each regression coefficient  $a_i$  equals to the mean change of y as  $x_i$  increases by one unit provided only that all other factors are unchanged.

As a rule, the model (2) does not reflect the actual regularities. The actual relations are much more complex than linear relations, therefore to construct adequate models of fertility multiple nonlinear correlations have to be used most often.

After the construction of a linear regression equation, the significance of both the equation as a whole (model adequacy) and its individual regressors is examined. The significance of regression equation is examined by variance analysis.

To verify the significance of the regression equation (or regression model adequacy) means to determine whether the mathematical model, which expresses the relationship between the variables, corresponds to the experimental data and whether the explanatory variables added to the equation are sufficient to describe the dependent variable. The adequacy of the regression equation is thought of as the correspondence of this equation with the experimental data. The adequacy of the regression models, i.e. their correspondence with the factual statistical data, is of great importance for their practical application.

The significance of regression equation as a whole is examined by using Fisher's exact test (F-test) preceded by variance analysis. In mathematical statistics, variance analysis is considered as an independent statistical analysis tool. In econometrics it is used as an auxiliary tool to study regression model quality. Mathematical model adequacy is checked by using F-test which is determined from the formula

$$\mathbf{F} = \frac{S_{fakt}^2}{S_{ost}^2} = \frac{\sum \left(\tilde{y}_i - \overline{y}\right)^2}{\sum \left(y_i - \tilde{y}_i\right)^2} \cdot \frac{n - m - 1}{m} = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m}$$
(3)

Here,  $y_i$  is the result of experimental measurements,  $\tilde{y}_i$  – the values of the dependent variable calculated from the model (1),  $\bar{y}$  – the mean of  $y_i$  values, n – the number of observations, m – the number of significant coefficients in the model (1);  $S_{fakt}^2$  – the variance of the factual values of the observable variable y;  $S_{ost}^2$  – residual variance;  $R^2$  – multiple determination coefficient.

The factual value of F-test is compared with the tabulated point  $\mathbf{F}_{tabl}(\alpha, k_1, k_2)$  whose value should be determined by a special table based on the significance level  $\alpha$  and the degrees of freedom  $k_1 = m$  and  $k_2 = n - m - 1$ . Here,

- if  $\mathbf{F} > \mathbf{F}_{tabl}(\alpha, k_1, k_2)$ , then the model is adequate and the statistical significance of the equation as a whole is recognized, and the calculated value of F-ratio is considered to be significant. At this point the regression analysis ends;

- if  $\mathbf{F} < \mathbf{F}_{tabl}(\alpha, k_1, k_2)$ , then the model is inadequate. Then the starting model should be changed and all the calculations be repeated.

While analyzing the regression equation (model) adequacy to the investigated process, the following may outcome:

1. The model constructed on the basis of F-test is generally adequate and all regression coefficients are significant. This model can be used for decision-making and forecasting.

2. The model based on F-test is adequate, but some coefficients are not significant. The model may be used for some decisions, but not for forecasts.

3. The model based on F-test is adequate, but all the regression coefficients are not significant. The model is considered to be completely inadequate. The model cannot be used for decision-making and forecasting.

To have a general idea of the model quality, the mean approximation error is determined for each observation based on relative divergence.

The adequacy of the regression equation (model) is checked by the mean approximation error whose value should not exceed 10-12% (advisable):

$$\overline{\varepsilon} = 100\% \cdot \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - \tilde{y}_i|}{y_i}$$
(4)

For practical purposes it is often required to compare the effect rendered on the dependent variable  $(\tilde{y})$  by different variables  $(x_i)$  when the latter are expressed by different units of measurement. In this case, standardized regression coefficients  $\beta_i$ , elasticities  $\beta_i$  and the mean elasticites are used:

$$\beta_i = a_i \frac{s_{x_i}}{s_y} \quad \mathbf{M} \quad \mathcal{P}_i = b_i \frac{\overline{x}_i}{\overline{y}} \quad (i = 1, 2, ..., m)$$
(5)

where  $\sigma_y$  and  $\sigma_{x_i}$  – are mean square (standard) deviations of y and  $x_i$  characteristics;  $\partial \tilde{y} / \partial x_i$  – is a

partial differential coefficient of the function (2) by factor variable  $x_i$ .

Elasticity's, along with the indices of correlation and determination for non-linear forms of relationship, are used to characterize the relationship between the resultant variable and factor variables. The amount of dependence between x and y variables may be estimated by using the elasticities.

The elasticity shows the percentage change in the value of the resultant variable y if the value of the factor variable changes by 1%.

Elasticity's may be calculated as the average and point coefficients.

The average elasticity characterizes the percentage change in the resultant variable *y* relative to its average level  $\overline{y}$  if the input variable  $x_i$  changes by 1% relative to its average level  $\overline{x}$ .

The average elasticity's are calculated from individual formulas for each type of non-linear functions.

The standardized regression coefficient  $\beta_i$  shows the value  $s_y$  mean change in the dependent variable y as only *i* explanatory variable increases by  $s_{x_i}$ .

The elasticity  $\mathcal{P}_i$  shows the average percentage change (relative to the mean) of y with an increase by 1% only.

Standardized regression coefficients  $\beta_i$  are comparable. Their intercomparison may rank the factors in terms of their impact on the result, i.e. on the dependent variable  $\tilde{y}$  .

That is the main advantage of standardized regression coefficients as opposed to net regression coefficients which are incomparable.

After calculating  $\beta_i$  we can write the regression equations in a standardized scale:

$$\hat{t}_{y} = \beta_{1} t_{x_{1}} + \beta_{2} t_{x_{2}} + \ldots + \beta_{m} t_{x_{m}}$$
(6)

The intercept term in the equation (12) is missing since all standardized variables are of zero mean value. The analyzed meaning of the standardized regression coefficients enables using them in factor selection – the factors with the least value  $\beta_i$  are eliminated from the model.

The application software for the construction of multiple regression equations depending on the solution algorithm being used enables obtaining either the regression equation for the initial data only, or the regression equation in a standardized scale.

The practical significance of a multiple regression equation is estimated by multiple correlation index (correlation ratio  $\eta$ ) and its square – multiple determination coefficient  $R^2$  whose computational formulas are of the following form:

$$\eta = \sqrt{1 - \sum_{i=1}^{n} (y_i - \tilde{y}_i)^2 / \sum_{i=1}^{n} (y_i - \overline{y})^2}, \quad R^2 = \eta^2$$
(7)

The multiple correlation index  $\eta$  characterizes the closeness of the relationship between the considered range of factors  $x_i$  and the investigated characteristic y, or otherwise, estimates the closeness of the factors' joint effect on the result.

Multiple determination coefficient is a way to express the variation proportion of the dependent variable y under the effect of the studied factors  $x_i$ . The determination coefficient estimates the dispersion *ratio* (*variance*) y which is explained by  $x_i$  in a regression model.  $\eta$  and  $R^2$  may take on a value from 0 to 1. The closer those coefficients are to one, the greater the regression equation explains the behavior of y.

If the number of parameters (m+1) approaches n, the numerator and denominator in (7) are corrected for the number of degrees of freedom of residual and total variance respectively.

The formula of the *adjusted multiple determination coefficients* are of the following form:

$$\overline{R}^{2} = 1 - \left[ \left( n - m - 1 \right) \sum_{i=1}^{n} \left( y_{i} - \widetilde{y}_{i} \right)^{2} \right] / \left[ \left( n - 1 \right) \sum_{i=1}^{n} \left( y_{i} - \overline{y} \right)^{2} \right]$$
(8)

#### **Results and Discussion**

The experiments were conducted on gray-brown irrigated soils under cereal crops in the Karabakh Plain. The experimental results are presented in the following table:

No.	x1	x2	x3	x4	x5	x6	x7	У	<b>y</b> *	8
	%		mg/kg	mg/kg	mg/kg	%		dt/ha	dt/ha	%
1	0.80	0.00	15.0	20.0	350	30.0	7.5	20.0	20.0	0.00
2	0.97	9.83	15.5	20.5	355	30.2	7.6	21.0	21.46	2.21
3	1.14	9.66	16.0	21.0	360	30.4	7.7	23.0	22.45	2.40
4	1.31	9.49	16.5	21.5	365	30.6	7.8	23.5	23.43	0.29
5	1.48	9.32	17.0	22.0	370	30.8	7.9	25.0	24.41	2.34
6	1.65	9.15	17.5	22.5	375	31.0	8.0	24.6	25.40	3.24
7	1.82	8.98	18.0	23.0	380	31.2	8.1	25.5	26.38	3.46
8	1.99	8.81	18.5	23.5	385	31.4	8.2	29.8	27.37	8.17
9	2.16	8.64	19.0	24.0	390	31.6	8.3	26.5	28.35	6.98
10	2.33	8.47	19.5	24.5	395	31.8	8.4	29.0	29.33	1.15
11	2.50	8.30	20.0	25.0	400	32.0	8.5	31.0	30.32	2.21
Total	18.15	90.65	192.5	247.5	4125	341.0	88.0	278.9		
Avg. value	1.65	8.241	17.5	22.5	375.0	31.0	8.0	25.355		

where **x1**- humus content; **x2** - CN=C:N; **x3** - N/NO<sub>3</sub>+N/NH<sub>3</sub>; **x4** - P<sub>2</sub>O<sub>5</sub>; **x5** - K<sub>2</sub>O; **x6** - total absorbed bases; **x7** - pH; **y** – factual yield of cereal crops; **y\*** - the calculation results from the model (1);  $\epsilon$  - mean relative approximation error.

When employing STATISTIKA-5-0 software package to reveal quantitative interdependence between crop yield and other soil indices, it has been found that the following multiple linear equation is the best mathematical model:

$$\tilde{y} = 0.435727 + 0.277681 \cdot x_1 + 0.048 \cdot x_2 + 6.016173 \cdot x_3 + 7.983546 \cdot x_4 - -1.753666 \cdot x_5 + 11.982141 \cdot x_6 + 3.166334 \cdot x_7$$
(9)

To compare the effect of the factors  $x_i$  on the result, we calculate the values of average elasticity's by using the formula (5). Using the data of the Table for the parameter  $\overline{\mathcal{P}}_i$ , we find the following values:

$$\mathcal{P}_{1} = 0.018074; \ \mathcal{P}_{2} = 0.015604; \ \mathcal{P}_{3} = 4.153176; \ \mathcal{P}_{4} = 7.085987; \ \mathcal{P}_{5} = -25.9418; \mathcal{P}_{6} = 14.65272; \ \mathcal{P}_{7} = 0.999238$$
(10)

That is, an increase of total absorbed bases alone by 1% increases crop yield by 14.65 % on an average. Thus, a greater effect of the factor  $x_6$  than that of other factors on crop yield y is proved.

Using the data of the Table, we calculate the standardized regression coefficients  $\beta_i$  from the formula (5). Then the standardized regression equation is of the following form:

$$\hat{t}_{y} = 0.007144 \cdot t_{x_{1}} + 0.029945 \cdot t_{x_{2}} + 1.338904 \cdot t_{x_{3}} + 1.776745 \cdot t_{x_{4}} - 39.027985 \cdot t_{x_{5}} + 0.426662 \cdot t_{x_{6}} + 0.028187 \cdot t_{x_{7}}$$
(11)

It is seen from the equation (11) that the third and fourth factors render a great positive effect on the result (as  $\beta_3 = 1.338904$ ,  $\beta_4 = 1.776745$ ) than the rest. At the same time, the fifth factor renders a negative effect as  $\beta_5 = -39.027985$ .

Since the least value  $\beta_1 = 0.007144$  was for  $\beta_i$ , the factor  $x_1$  should be excluded from the model (9). Then finally the fertility model takes on the following form:

$$\tilde{y} = 0.435727 + 0.048 \cdot x_2 + 6.016173 \cdot x_3 + 7.983546 \cdot x_4 - 1.753666 \cdot x_5 + 11.982141 \cdot x_6 + 3.166334 \cdot x_7$$
(12)

Since  $\sup_{\forall i} {\beta_i} = \beta_4 = 1.776745$  and on the ground of comparison of the equations (9) and (12), it may be

concluded that greater effect on production is rendered by the fourth factor rather than the sixth factor as it appears from the regression equation (9) in the natural scale.

Consequently, an increase of  $P_2O_5$  alone by 1% increases crop yield by 7.09% on an average. Thus, a greater effect of the factor  $x_4$  than that of other factors on crop yield y is proved.

To ensure further application of the proposed model in the research, we tested its adequacy. First, we calculate the correlation ratio. We calculate  $\eta = 0.9493$  from the formula (1), and it shows high closeness of the relationship between the input  $(x_1, x_2, \dots, x_7)$  and the resultant  $\tilde{y}$  features of soil fertility models.

Using formula (7) for multiple determination coefficients we have  $R^2 = 0.9012$ .

The unadjusted multiple determination coefficient  $R^2 = 0.9012$  estimates the dispersion ratio of the results due to the factors in the total variation of the result represented in the equation. Here, this ratio makes 90.12% and points to a very high degree of result variation dependence on factor variation, in other words – to very close relationship between the factors and the result.

The adjusted multiple determination coefficients:

$$\overline{R}^2 = 1 - \left(1 - R^2\right) \left(\frac{n - 1}{n - m - 1}\right) = 1 - \left(1 - 0.9012\right) \frac{11 - 1}{11 - 7 - 1} = 1 - 0.0988 \cdot \frac{10}{3} = 1 - 0.3293 = 0.6707$$

defines the closeness of the relationship taking into account the degrees of freedom of total and residual variances. It gives such estimation of the closeness of relationship that does not depend on the number of factors and, therefore, may be compared in different models with different number of factors.

Both coefficients point to quite high (more than 90% and 67.07 %) determinacy of the result y in the model by the factors  $x_1, x_2, \dots, x_7$ .

As the Table suggests, the maximum ratio error is 8.17% and the mean relative approximation error makes  $\overline{\varepsilon} = 2,95\%$  which is an acceptable level for crop yield models.

The quality of the model, based on the relative deviations for each observation, is recognized to be good since the mean approximation error does not exceed 10%.

The statistical significance of the regression equation as a whole will be estimated by using the Fisher's exact test. The factual value of F-test from the formula (3) will be

$$\mathbf{F} = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m} = \frac{0.9493^2}{1 - 0.9493^2} \cdot \frac{11 - 7 - 1}{7} = \frac{0.9012}{1 - 0.9012} \cdot \frac{3}{7} = 3.9092$$

The tabular value of the F-test at 5% significance level ( $\alpha = 0.05$ ) and the degrees of freedom  $k_1 = m = 7$ 

and  $k_2 = n - m - 1 = 11 - 7 - 1 = 3$  is  $\mathbf{F}_{tabl}(0.05; 7; 3) = 8,89$ .

Since  $\mathbf{F} = 3.9097 < \mathbf{F}_{tabl} (0.05;7;3) = 8.89$ , the regression equation is recognized to be statistically non-significant.

We also have the same when  $\alpha = 0.01$ , i.e. as  $\mathbf{F} = 3.9097 < \mathbf{F}_{tabl} (0.01; 7; 3) = 27.67$ .

In this situation, the construction of a regression model based on the variables under study is considered to be unreasonable.

### Conclusion

- A technique of constructing a yield model based on experimental design is proposed. The technique was applied for gray-brown irrigated soils.
- Soil fertility models for gray-brown irrigated soils under cereal crops constructed on the basis of experimental design were analyzed. The statistical analysis shows that the relative error of the proposed model is not more than 8.17, which is acceptable for practical use.

#### References

- Aparin, B.F., 1998. Soil fertility parameters and management systems. *Bulletin of Dokuchaev Soil Institute* 48: 12-15. [In Russian].
- Methods of Mathematical Biology, 1982. Book 4. Methods of identification of mathematical models of biological systems. Kiev, Ukraine. 192 p. [In Russian].
- Mikayilov, F.D., 2014. Modeling of some soil processes. *Bulletin of Altai State Agricultural University* 7(117): 59-64. [In Russian].
- Neuymin, Ya.G., 1984. Models in science and technology. History, theory, practice. Leningrad: Russia. Nauka, 192 pp. [In Russian].
- Nuriyeva, K.G., 1994. Agromelioration evaluation and development of optimal soil fertility model of the Karabakh Plain. Thesis Abstract. Cand. Agr. Sci. – Baku, Azerbaijan. 20 pp. [In Russian].
- Pachepskiy, Ya. A., 1992. Mathematical models of the processes in reclaimed soils. Moscow: Russia. MGU. 85 pp. [In Russian].
- Poluektov, Z.A., 1991. Dynamic models of agro-ecosystems. Leningrad: Russia. Gidrometeoizdat, 312 pp. [In Russian].
- Semenov, V.A., 1982. On soil fertility modeling. In: Soil fertility models and the methods of their development. Scientific Paperes of Dokuchaev Soil Institute. Moscow, Russia. p.36-44. [In Russian].
- France, J., Thornley, J.H.M., 1987. Mathematical models in agriculture; translated from English. Moscow, Russia. Agropromizdat, 400 pp. [In Russian].
- Frid, A.S., 1990. The system of soil fertility models: development and use. Thesis Abstract. Dr. Agr. Sci. Moscow, Russia. 36 pp. [In Russian].
- Shisho, L.L. Karmanov, I.I., Durmanov, D.N., 1987. Criteria and models of soil fertility. Moscow: Russia. VO Agropromizdat, 184 pp. [In Russian].
- Khomyakov, D.M. Khomyakov, P.M., 1996. Fundamentals of systems analysis. Moscow: Mechanics and Mathematics Faculty of Lomonosov Moscow State University, 108 pp. [In Russian].
- Ismayilov, A.I., 2013. Optimization problems of soil fertility. *Soil-Water Journal* 2(1): 1085-1090.