IEEE

PROCEEDINGS
2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies

DESSERT

2020 Theme:
IoT, Big Data and AI for a Safe & Secure World and Industry 4.0

ORGANIZED BY

National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine
Banking University, Kyiv, Ukraine
Leeds Beckett University, Leeds, United Kingdom
IEEE United Kingdom and Ireland Section
IEEE Ukraine Section

Ukraine, Kyiv
May 14-18
2020

Table of Contents

Linear Regression Model for Substantiation of Sustainable State Policy in a Digital Economy

Oleksandr Dorofyeyev
Department of Public Management and Administration of Poltava State Agrarian Academy
Poltava, Ukraine
dorofiejew@ukr.net

Tamara Lozinska
Department of Public Management and Administration of Poltava State Agrarian Academy
Poltava, Ukraine
tnloz@rambler.ru

Yuriy Ponochovnyi
Department of Information Systems and Technologies of Poltava State Agrarian Academy
Poltava, Ukraine
yuriy.ponch@gmail.com

Tetiana Vlasenko
Department of production, business and management organization of Petro Vasylenko National Technical University of Agriculture
Kharkiv, Ukraine
tatyanavlasenko2011@gmail.com

***Abstract*—The article discusses the use of digital mathematical models in justifying public policy for sustainable economic development. The wider expediency uses of statistical analysis methods to substantiate public administration decisions in the digital economy is explained. The technique of forming and preparation of experimental data for the construction of a mathematical regression model is described. The hypothesis of normal distribution of the input data was tested according to the statistical criteria of Kolmogorov-Smirnov and Lilliefors and Shapiro-Wilk. Box-Cox transformation was used to normalize the data. The linear multiple regression coefficients are determined by the least squares method. The quality of the resulting model was evaluated using the Student's and Fisher's criteria. The results of the forecast of financial stability of agricultural enterprises for 2020-2022 and the factor indicators indicate the negative dynamics of the effective indicator of their financial condition. The proposed model can be used to continuously monitor the financial condition of agricultural enterprises by linking it to the State Statistics Service of Ukraine.**

Keywords—digital economy; state policy of sustainable development; agricultural enterprise; financial sustainability; mathematical simulation; linear multiple regression; normal distribution.

I. Introduction

The linear regression analysis method is quite often used in agrarian economics to make a link between a set of independent and dependent variables. However, in many cases, the detection of certain dependencies does not find application in public administration practice. The technological processes of the digital economy and the need to ensure the sustainability of economic development of agricultural enterprises actualize the issue of monitoring the financial performance of their activities with the aim of timely transformation of agricultural policy.

The results of this analysis can be used for these purposes.

In the scientific literature of recent years, attempts to have been made to apply these methods to identify the relationship between factors and results. Malyarets E. et al. [1] are concerned with the search for a model for assessing the financial condition of enterprises; Voloshin O. and Kulik V. also consider production activity as an object of system analysis [2]; quantitative parameters for assessing the stability of the agricultural market are explored by Pasemko G. et al. [3]. The development of original modeling methods to support managerial decisions are devoted to the work of Russian researchers Nasonov A. [4], Timofeeva G. and Antamoshkina E. [5]. Authors such as Ahmet M. and Feng M. [6] focus on identifying modeling capabilities under deterministic chaos, and Pelivan N and Sahin A. on using regression methods to establish a link between clear factors and fuzzy results [7]. As we can see, there is no recent attempt to use regression analysis to justify agrarian policy.

The purpose of this work is to predict the dynamics of changes in the financial result of agricultural enterprises using an abovementioned model, taking into account the indicators of financial stability. To do this, you need to analyze the input data and bring them to a normal distribution law. The research question is to substantiate the feasibility of reviewing the state policy to support the sustainable development of the agrarian economy.

The article is structured as follows. The following section describes the construction of a linear regression model. The set of input parameters was checked for normality and normalized through Box-Cox transformations. The calculation of the values of the seven factor indicators of the model. The third section is devoted to compilation of the profitability forecast of agricultural enterprises using the model. The fourth section references related work. Section 5 summarizes the conclusions and perspectives of future research

II. Determination Of Parameters Of The Linear Regression Model

A. Description of the input validation process for model building

The initial data for model construction are taken from [8,9]. For convenience, we used the unified designation of the input parameters of the model as Xi: resultant indicator - Y. Their initial values are presented in Table I. The nomenclature of economic development indicators that were analyzed is as follows:

X1 = Kef – coefficient of efficiency of use of own funds as the ratio of profit to equity;

X2 = Kfi – coefficient of financial independence as the ratio of equity to the value of all capital;

X3 = CA/TA ratio of current assets to total assets;

X4 = NP/Rev – the ratio of net profit to revenue;

978-1-7281-9957-3/20/$31.00 © 2020 IEEE

TABLE I. INPUT DATA FOR THE MODEL BUILDING

Para-meter	Value (in years)									
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
X1	0,10	0,47	0,55	0,11	-0,61	1,10	1,16	0,05	0,07	0,05
X2	0,11	0,50	0,55	0,12	-0,53	0,98	1,11	0,06	0,08	0,06
X3	0,19	0,53	0,59	0,17	-0,42	0,87	1,11	0,10	0,16	0,10
X4	0,21	0,54	0,61	0,20	-0,44	0,83	1,13	0,11	0,18	0,11
X5	0,18	0,54	0,63	0,16	-0,41	0,87	1,18	0,10	0,16	0,10
X6	0,10	0,50	0,62	0,09	-0,49	1,00	1,24	0,05	0,08	0,05
X7	0,13	0,42	0,68	0,10	-0,45	1,38	1,63	0,06	0,09	0,06
X8	0,37	0,40	0,75	0,28	-0,47	1,49	1,87	0,15	0,30	0,15
X9	0,25	0,24	0,86	0,22	-0,74	3,16	3,57	0,06	0,25	0,06
Y	0,16	0,48	0,70	0,15	-0,50	1,09	1,46	0,08	0,16	0,08

X5 = (Pr - Ac) / Ac – ratio of difference between revenue and assets to assets;

X6 = Kaeq – ratio of attracted and equity, as a proportion of the distribution of the amount of borrowed funds in the amount of equity;

X7 = Kman – the coefficient of maneuverability of own funds as a ratio of own working capital to equity;

X8 = Kfr – ratio of financial resources (total assets) as a ratio of profit to total assets;

X9 = NP/Cos – ratio of net profit to the cost of sales (goods, works, services);

Y – Financial result from ordinary activities before tax / Balance sheet at the end of the reporting period.

The input data presented is a random sample of independent observations from 2008 to 2017 [8,9].

A prerequisite for applying the least squares linear regression coefficient finding procedure and further statistical justification of the obtained model using the Student's and Fisher's statistical criteria is the assumption of a normal distribution of the original data [8, 9].

Table II also presents the results of testing the source data for compliance with the normal distribution law. The MS Excel spreadsheet was used for the tests. The tests were performed by Kolmogorov-Smirnov and Liliefors (K-S&L) and Shapiro-Wilk (Sh-W) [10].

TABLE II. RESULTS OF THE NORMAL DISTRIBUTION HYPOTHESIS TEST

Parameter	X1	X2	X3	X4	X5	X6	X7	X8	X9	Y
Norm. test	**Yes**	**No**	**Yes**	**Yes**	**No**	**No**	**No**	**No**	**Yes**	**No**
K-S&L	Yes	No	Yes	Yes	No	No	No	No	Yes	No
Sh-W	Yes	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes

Testing the hypothesis allowed us to establish a statistically significant difference in the distribution of input data from the norm.

B. *Description of the process of modifying the input data to build the model*

To provide a further procedure for constructing the regression model, in order to normalize the data, the Box-Cox transformation (1) and the corresponding inverse (2) [11] were applied:

$$\tilde{x}_i = \begin{cases} \dfrac{(x_i+\alpha)^{\lambda}-1}{\lambda}, \lambda \neq 0; \\ \ln(x_i+\alpha), \lambda = 0. \end{cases} \tag{1}$$

$$x_i = \begin{cases} (\lambda \tilde{x}_i + 1)^{\frac{1}{\lambda}} - \alpha, \lambda \neq 0; \\ e^{\tilde{x}_i} - \alpha, \lambda = 0. \end{cases} \tag{2}$$

The optimal transformation parameters (1) for each set of output data (Table I) are determined from the condition of maximizing the logarithm of the likelihood function [12]:

$$f(x,\lambda,\alpha) = -\frac{N}{2}\ln\left[\sum_{i=0}^{N-1}\frac{\left(x_i(\lambda,\alpha)-\bar{x}(\lambda,\alpha)\right)^2}{N}\right] + \\ +(\lambda-1)\sum_{i=0}^{N-1}\ln\left(x_i(\lambda,\alpha)\right) \tag{3}$$

The values of λ and α were calculated using Statistica software. The results of application of transformation (1), taking into account (3), as well as the results of tests for normality of data were obtained using MS Excel and are presented in Table III:

TABLE III. RESULTS OF THE TRANSFORMATION OF THE INPUT DATA AND THE CHECK OF THEMSELVES FOR THE NORMAL DISTRIBUTION

	X1	**X2**	**X3**	**X4**	**X5**	**X6**	**X7**	**X8**	**X9**	**Y**
λ	-0,44	5,75	-4,56	-5,00	5,24	-2,47	-2,77	-0,87	0,16	-0,87
α	0	0	0,45	0,907	1	0,17	0	0	0	0
2008	-4,10	-0,17	0,00	0,02	-0,19	0,18	0,12	-15,79	-2,13	-15,79
2009	-3,66	-0,17	0,01	0,02	-0,19	0,12	0,09	-12,83	-2,05	-12,83
2010	-2,43	-0,17	0,04	0,06	-0,18	0,05	0,09	-7,28	-1,57	-7,28
2011	-2,28	-0,17	0,06	0,08	-0,18	0,00	0,10	-6,57	-1,50	-6,57
2012	-2,53	-0,17	0,07	0,06	-0,18	0,04	0,13	-7,56	-1,61	-7,56
2013	-4,08	-0,17	0,06	0,00	-0,18	0,13	0,16	-14,88	-2,08	-14,88
2014	-3,26	-0,17	0,10	0,01	-0,18	0,27	0,27	-13,07	-2,01	-13,07
2015	-1,23	-0,17	0,13	0,12	-0,18	0,29	0,30	-4,83	-1,11	-4,83
2016	-1,93	-0,17	0,16	0,09	-0,19	0,38	0,35	-12,28	-1,25	-12,28
2017	-2,83	-0,17	0,11	0,05	-0,19	0,18	0,24	-9,66	-1,59	-9,66
Norm. test	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
K-S&L	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sh-W	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

III. CONSTRUCTION OF A LINEAR REGRESSION MODEL FOR JUSTIFICATION OF STATE SUSTAINABLE DEVELOPMENT POLICY

A. *Description of the model construction process*

Dependence of Y (Financial result from ordinary activities before tax / Balance sheet at the end of the reporting period) on factor indicators X1-X9 (Table I) is represented by a linear regression model:

$$\begin{aligned} Y &= b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \\ &+ b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + \\ &+ b_8X_8 + b_9X_9 = \sum_{i=0}^{9} b_iX_i \end{aligned} \tag{4}$$

The coefficients bi, i = [0..9] of model (4) indicate the magnitude by which the resultant index Y changes if the values of the factor exponent Xi are increased by one (at fixed values of other factors).

Table IV summarizes the results of the multiple regression analysis performed by the stepwise regression method in the MS Excel spreadsheet. The following notation was used:

b* - standardized model coefficients;

Std.Err. b * is the standard error of estimation b*;

b - model coefficients;

Std.Err. b - standard error of estimation of model coefficients;

t(3) is the value of the Student's t test to evaluate the significance of the model coefficients;

p-value - the level of significance of regression coefficients.

TABLE IV. RESULTS OF MULTI-REGRESSION ANALYSIS VOR INPUT VALUES

Intercept	b*	Std.Err. b*	b	Std.Err. b	t(3)	p-value
			0,116	0,019	6,262	0,008
X1	0,011	0,091	0,004	0,037	-2,373	0,049
X2	-0,179	0,066	-0,067	0,025	-2,463	0,045
X3	0,024	0,022	0,009	0,008	-2,749	0,035
X4	0,971	0,084	0,543	0,047	11,498	0,001
X5	0,391	0,031	0,133	0,010	12,800	0,001
X7	-0,451	0,052	-0,020	0,002	-8,738	0,003

For the coefficients of model X1… X7, the p-value is <0.05, so the specified regression coefficients are significant.

Thus, according to equation (4), we obtained the following mathematical model:

$$\begin{aligned} Y &= 0{,}116013 + 0{,}004402 \cdot X_1 - \\ &-0{,}067131 \cdot X_2 + 0{,}008652 \cdot X_3 + \\ &+0{,}542929 \cdot X_4 + 0{,}133140 \cdot X_5 - \\ &-0{,}020398 \cdot X_7 \end{aligned} \quad (5)$$

Note that the factor scores X6, X8, X9 were not included in the model because the corresponding regression coefficients were not significant.

B. *Description of the process of checking the quality of the model*

The results were calculated using MS Excel; Statistica software was used to test statistical hypotheses.

TABLE V. RESULTS OF MULTI-REGRESSION ANALYSIS FOR PREDICTED VALUES

Years	Observed	Predicted	Residual
2008	0,045398	0,045221	0,000177
2009	0,056610	0,056571	0,000039
2010	0,101240	0,101794	-0,000555
2011	0,111980	0,111979	0,000000
2012	0,097569	0,096826	0,000743
2013	0,048379	0,048706	-0,000327
2014	0,055497	0,055354	0,000143
2015	0,150380	0,150393	-0,000013
2016	0,059265	0,059268	-0,000003
2017	0,076067	0,076272	-0,000205
Minimum	0,045398	0,045221	-0,000555
Maximum	0,150380	0,150393	0,000743
Mean	0,080239	0,080239	0,000000
Median	0,067666	0,067770	-0,000001

In order to evaluate the quality of the constructed model, the value of the coefficient of determination R2 = 0,99989905 was calculated. This shows that the constructed model explains 99.9% of the variation of the trait. Testing the quality of the model by analysis of variance showed that the model is statistically significant (F(6,3) = 4952,604; p <0.000004). In this case, the standard error of estimating the values of the resultant feature, made according to this model, was equal to 0.00060, which does not exceed 5% of the average value (0.004012).

Table V presents Observed (Table I, column Y) and predicted values. Their comparison indicates the normal distribution of the residual regression (Residual) and the absence of correlation between the regression residues and theoretical values (Fig. 1).

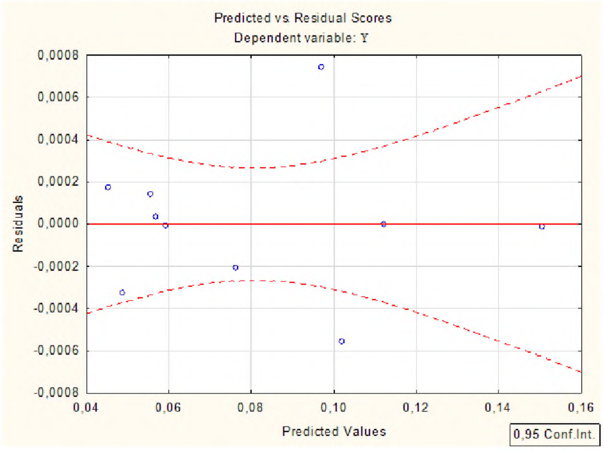

Fig. 1. Scatter plot of regression residuals relative to the theoretical values of the resultant indicator..

From the calculated empirical value of the Darbin-Watson statistics (Durbin-Watson d = 2.506539), it follows that the remains of model (5) are uncorrelated [13]. This demonstrates the independence of estimation errors in the regression model and confirms its adequacy. Therefore, the constructed model (5) is adequate.

IV. CASE STUDY – USING A LINEAR REGRESSIVE MODEL FOR FORECASTING FACTOR AND RESULTING INDICATORS

Table VI presents the results of the 2019-2022 forecast of the factor indicators included in the model (5) and the corresponding values of the effective indicator Y. The polynomial trend of the 6th order will be used to predict the values of the factor indicators X1… X7. This minimized the approximation error. The calculations were performed in MS Excel.

TABLE VI. FORECAST OF FACTORS AND RESULTS OF MULTIPLE REGRESSION ANALYSIS

Years	X1	X2	X3	X4	X5	X7	Y
2019	0,1787	0,4187	0,6455	0,174	-0,5675	1,8069	0,076333
2020	0,1833	0,3948	0,6543	0,1813	-0,6088	1,9807	0,072953
2021	0,1879	0,3681	0,6631	0,1892	-0,6615	2,1735	0,068182
2022	0,1925	0,3386	0,6719	0,1977	-0,7268	2,3853	0,061859

The combined results of the approximation of factor indicators and their forecast are presented in Fig. 2. Since the use of a linear regression model involves the use of normalized factors, the results are presented in Fig. 2 without displaying the initial data in Table I.

Changes in the normalized factor indicators in 2008-2022 are characterized by the following trends: X7 (coefficient of maneuverability of own funds as a ratio of own working capital to equity) - shows a marked monotonous growth; X3 (ratio of current assets to total assets), X2 (coefficient of financial independence as the ratio of equity to the value of all capital), X1 (coefficient of efficiency of use of own funds as the ratio of profit to equity), X4 (ratio of net profit to revenue) - remain virtually constant; X5 (ratio of difference between revenue and assets to assets) - shows a slight decline.

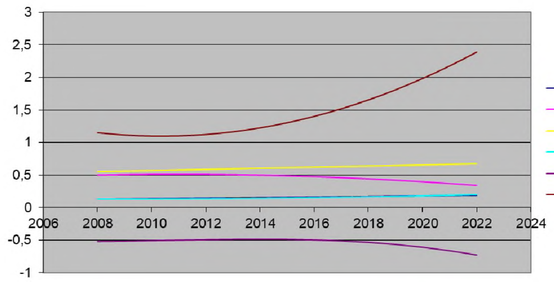

Fig. 2. Dynamics of change and forecast of factor indicators.

Dynamics of these indicators are accordingly reflected in the forecast of the dynamics of the effective indicator Y (Financial result from ordinary activities before tax / Balance sheet at the end of the reporting period) (Fig. 3). The performance metric was calculated in MS Excel based on (5).

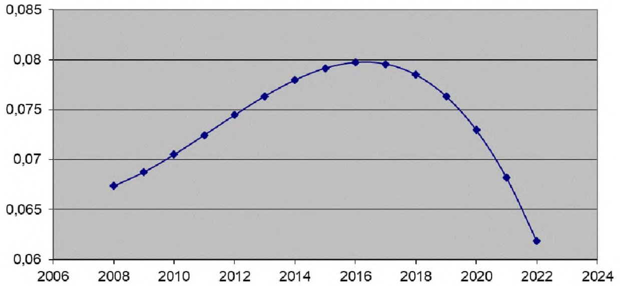

Fig. 3. Dynamics and forecast of Y in 2008-2022.

As the indicator of the financial condition of the agricultural enterprises of Ukraine Y is the main one for determining the state policy of sustainable development, Fig. 4 presents a comparison of its actual values in 2008-2018 (Table I) and calculated by (5) values in 2008-2022.

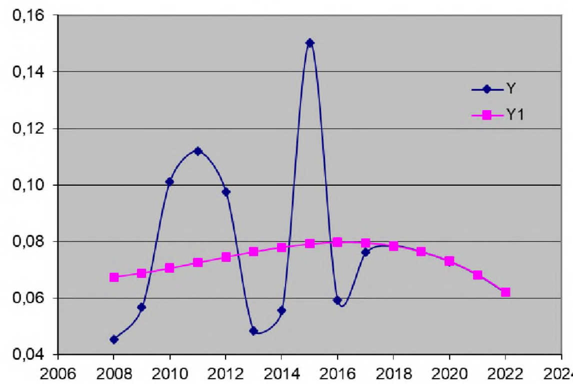

Fig. 4. Comparison of actual values of Y in 2008-2018 and Y1 of model (5) for 2008-2022.

The results of comparison in Fig. 5 showed that the trend of the resultant index Y1 of model (5) relative to the actual values in Table I occupies an intermediate "averaged" position. The best match for model values and actual data in 2018, with a maximum deviation of 0.07 in 2015.

To detail the forecast of factor and result indicators, the period 2019-2022 is presented on separate graphs (Fig. 5, Fig. 6).

The results of the model factor forecast for 2019-2022 indicate a linear increase of X7 (coefficient of maneuverability of own funds as a ratio of own working capital to equity) of 0.5784 (32%) and a decline of X5 (ratio of difference between revenue and assets to assets) of 0.16 (28%).

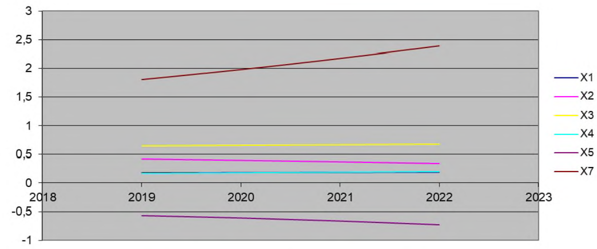

Fig. 5. Forecast of dynamics of factor indicators of the model for 2019-2022.

However, more research is needed to determine the causes of these changes and the ways in which they may be affected.

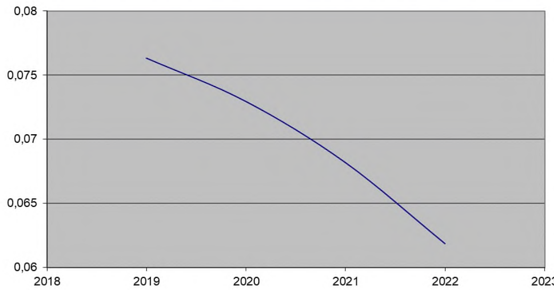

Fig. 6. Forecast of the dynamics of the performance indicator for 2019-2022

According to the model (5) forecast, in 2019-2002, the resultant indicator Y (Financial result from ordinary activities before tax / Balance sheet at the end of the reporting period) has a clear downward trend. The absolute value of the reduction is 0.014, the relative value is 18.9%.

V. Related Works

The importance of using regression analysis methods in public administration is emphasized by many researchers, in particular [14]. In [15,16] the impossibility of timely decision-making without timely processing of Big Data. Special Data Mining methods are used to process large amounts of data. At the same time, the tasks of data finalization, which also include regression analysis, remain relevant. In the Table VII shows examples of using information packages to perform statistical and regression analysis tasks.

TABLE VII. Related Works Studying Software For Multiple Regression Analysis

Paper	Area / Industry	Software
P. van Schaik et al. [17]	public health	R package & fingertipsR
Amani Ayad, et al. [18]	Software project management	java + apache common mathematical library
K. T. Ejlerskov et al. [19]	Supermarket policies	Stata/SE v14.2 & R v3.3.1
Kumari K, Yadav S. [20]	General questions	SPSS statistical software & MS Excel
M. Valsamis et al. [21]	Healthcare	MATLAB & MedCalc software suite

Analysis of the use of software for statistical surveys showed that the choice of tools depends more on the preferences of the (preferred) authors. In this paper, the choice of MS Excel and Statistica tools is justified not only by the presence of the necessary functions in the programs, but also by the need for additional visual teaching of the material to students who master this tool.

VI. Conclusion and Future Work

The mathematical model of linear regression for forecasting the financial performance of agricultural enterprises is constructed. The expediency of wider use of methods of statistical analysis in the decision-making of government in digital economy is substantiated. The technique of forming and preparation of experimental data for the construction of a mathematical regression model is described. The hypothesis of normal distribution of the input data was tested according to the statistical criteria of Kolmogorov-Smirnov and Lilliefors and Shapiro-Wilk. It showed deviations from the normal distribution of input parameters X2, X5… X8. Box-Cox transformation was used to normalize the data.

The linear multiple regression coefficients are determined by the least squares method. For the coefficients of model X1… X5, X7, the significance level of the regression coefficients was p-value <0.05. Factors X6, X8, X9 were not included in the model because the corresponding regression coefficients were not significant.

The quality of the resulting model was evaluated using the Student's and Fisher's criteria. From the calculated empirical value of the Darbin-Watson statistics (Durbin-Watson d = 2.506539), it follows that the residuals of the constructed regression model are uncorrelated, which confirms its adequacy.

The results of the forecast of financial stability of agricultural enterprises for 2020-2022 and the factor indicator X5 indicate a negative dynamics. The factor X5 is projected to decline by 0.16 (28%). The forecast of the absolute value of the decrease of the effective indicator of the financial condition of the agricultural enterprises of Ukraine amounted to 0,014, relative - 18,9%. Based on the obtained values, it is necessary to insist on the expediency of transformation of state agricultural policy in the direction of changing the structure of instruments of regulatory influence, in particular strengthening measures to support the sustainability of development of agricultural enterprises: increase of budget financing of insurance of economic losses from changes in market conditions; the use of public procurement of socially significant types of agricultural raw materials; introduction of indicative prices for these products and so on. When updating the data in the statistical bulletin, it is necessary to repeat all the model calculation procedures to correct the coefficient values and confirm / refute the hypothesis about their normal distribution.

The paper highlights the results, the novelty of which is the use of linear regression analysis for an economic realm of the Ukrainian economy - the agricultural sector. The methods used in this work are limited by the requirement to bring the values of the input parameters to normal form. The direction of further research is to use the proposed model to constantly monitor the financial condition of agricultural enterprises by linking it to the State Statistics Service of Ukraine.

References

[1] L. Malyarets, Y. Misyura and V. Koybichuk, Mathematical Methods and Models in the Management of Economic Processes. Kharkiv: KhNEU im.S. Kuznetsya, 2016.

[2] O. Voloshin and V. Kulik, "Production activity as an object of system analysis and simulation modeling", in Mathematical And Simulated Modeling Systems "MODS 2018", Chernigiv: ChNTU, 2018, pp. 165-167.

[3] O. Dorofyeyev, T. Lozinska, G. Pasemko, "Application of the principle of harmony in strategic management by economic entities for agricultural market stability", Actual Problems Of Economiics, vol. 11, no. 185, pp. 168-177, 2016.

[4] A. Nosonov, "Modern Methods For Spatial Modeling Of The Development Of Agriculture", Bulletin of the Moscow State Regional University (Natural Sciences), no. 3, pp. 62-74, 2018. DOI: 10.18384/2310-7189-2018-3-62-74

[5] G. Timofeeva and E. Antamoshkina, "Efficiency of agri-food policy: models and assessment methods", Public servic, vol. 5, pp. 10-16, 2015.

[6] M. Akhmet and M. Fen, "Replication of Chaos in Neural Networks, Economics and Physics", Nonlinear Physical Science, 2016. DOI: 10.1007/978-3-662-47500-3

[7] N. Yapıcı and P. Aynur Şahin, "An Integrated approach for fuzzy logistic regression", Journal of Statisticians: Statistics and Actuarial Sciences, vol.1, pp. 42-54, 2018. Available: https://dergipark.org.tr/tr/download/article-file/514721

[8] "Activities of economic entities: statistical summary for 2010", Ukrstat.gov.ua, 2011. [Online]. Available: http://www.ukrstat.gov.ua/druk/publicat/kat_u/2018/zb/11/zd_2018.pdf. [Accessed: 07- Mar- 2020].

[9] "Activities of economic entities: statistical summary for 2018", Ukrstat.gov.ua, 2018. [Online]. Available: http://www.ukrstat.gov.ua/druk/publicat/kat_u/2018/zb/11/zd_2018.pdf. [Accessed: 07- Mar- 2020].

[10] M. Wei, Y. Gan and S. Tang, "Bifurcations of Travelling Wave Solutions for the B(m,n) Equation", American Journal of Computational Mathematics, vol. 04, no. 02, pp. 104-118, 2014. DOI: 10.4236/ajcm.2014.42010.

[11] M. Hossain, "The Use of Box-Cox Transformation Technique in Economic and Statistical Analyses", Journal of Emerging Trends in Economics and Management Sciences (JETEMS), vol. 2, no. 1, pp. 32-39, 2011.

[12] A. Porunov, "Box-Cox transformation and the illusion of "normality" of the macroeconomic series", Business Informatics, vol. 2, no. 12, pp. 3-10, 2010.

[13] "Durbin-Watson Table | Real Statistics Using Excel", Real-statistics.com, 2020. [Online]. Available: http://www.real-statistics.com/statistics-tables/durbin-watson-table/. [Accessed: 07-Mar- 2020].

[14] P. Maravelakis, "The use of statistics in social sciences", Journal of Humanities and Applied Social Sciences, vol. 1, no. 2, pp. 87-97, 2019. DOI: 10.1108/jhass-08-2019-0038.

[15] J. Höchtl, P. Parycek and R. Schöllhammer, "Big data in the policy cycle: Policy decision making in the digital era", Journal of Organizational Computing and Electronic Commerce, vol. 26, no. 1-2, pp. 147-169, 2015. DOI: 10.1080/10919392.2015.1125187.

[16] V. Kharchenko and O. Illiashenko, "Concepts of Green IT Engineering: Taxonomy, Principles and Implementation", Green IT Engineering: Concepts, Models, Complex Systems Architectures, pp. 3-19, 2016. DOI: 10.1007/978-3-319-44162-7_1.

[17] P. van Schaik, Y. Peng, A. Ojelabi and J. Ling, "Explainable statistical learning in public health for policy development: the case of real-world suicide data", BMC Medical Research Methodology, vol. 19, no. 1, 2019. DOI: 10.1186/s12874-019-0796-7.

[18] A. Ayad, I. Marsit, N. Mohamed Omri, J. Loh and A. Mili, "Using Semantic Metrics to Predict Mutation Equivalence", Communications in Computer and Information Science, pp. 3-27, 2019. DOI: 10.1007/978-3-030-29157-0_1.

[19] K. Ejlerskov, S. Sharp, M. Stead, A. Adamson, M. White and J. Adams, "Supermarket policies on less-healthy food at checkouts: Natural experimental evaluation using interrupted time series analyses of purchases", PLOS Medicine, vol. 15, no. 12, p. e1002712, 2018. DOI: 10.1371/journal.pmed.1002712.

[20] K. Kumari and S. Yadav, "Linear regression analysis study", Journal of the Practice of Cardiovascular Sciences, vol. 4, no. 1, p. 33, 2018. DOI: 10.4103/jpcs.jpcs_8_18.

[21] E. Valsamis, D. Ricketts, H. Husband and B. Rogers, "Segmented Linear Regression Models for Assessing Change in Retrospective Studies in Healthcare", Computational and Mathematical Methods in Medicine, vol. 2019, pp. 1-9, 2019. DOI: 10.1155/2019/9810675.