World Economics & Finance Bulletin (WEFB)
Available Online at: https://www.scholarexpress.net
Vol. 43,February, 2025
ISSN: 2749-3628,

THE IMPORTANCE AND STATE OF ECONOMETRIC MODELING IN FORECASTING THE DEVELOPMENT DYNAMICS OF THE CHEMICAL INDUSTRY

Sharipova Saodat Ubaydullaevna
TSEU, basic PhD student
saodatsharipova1111@gmail.com

Article history:
Received: 20th November 2024
Accepted: 14th December 2024

Abstract:
The article models the dynamics of the chemical industry development using econometric models. Initially, the article provides a brief description of scientific sources related to the topic. Then, issues such as the regression model, determination, correlation coefficients, the Durbin-Watson test, and empirical probability regarding the Fisher criterion are discussed. In the final section of the article, all the obtained results are summarized.

Keywords: chemical industry, mining industry, development dynamics, forecasting, growth rates, Fisher statistic, regression, correlation, AR model.

INTRODUCTION

The chemical industry plays a crucial role in the economic development of every country today. This sector is significant not only for economic growth but also as one of the leading sources of innovation and technological advancement. Studying the development dynamics of the chemical industry, its interaction with social, economic, and environmental factors, as well as the changes arising from the application of modern technologies, is of great importance in shaping effective strategies for countries and industrial enterprises.

There are many factors that influence the development of the chemical industry. Among them, economic stability, technological innovations, effective management of raw material resources, environmental sustainability, and the optimization of production processes are the most important. At the same time, to fully understand the dynamics of the sector's changes, a complex and systematic approach is required. Through these approaches, key issues such as the growth rate of the chemical industry, the efficiency of resource utilization, the environmental risks of production processes, and their social impacts can be explored.

This article examines the modeling of the chemical industry's development dynamics. Our goal is to investigate various modeling methods to forecast the future development directions of the industry, taking into account its growth rates, technological innovations, and environmental factors. This includes analyzing the possibilities of simulating industrial development based on complex approaches that account for system dynamics, economic models, and ecological sustainability. Additionally, the article aims to contribute to the development of strategic approaches necessary for ensuring the future growth of the chemical industry.

Modeling the development of the chemical industry, its effective management, the optimization of resource use, waste reduction, and increasing social responsibility are among the pressing tasks of today. The article explores these issues and creates opportunities for the industry's changing dynamics and sustainable development.

LITERATURE REVIEW

In this section, we analyze scientific articles related to the topic "Modeling the Dynamics of Chemical Industry Development." The articles listed below discuss various issues related to the development of the chemical industry and modeling methodologies. Each article contributes significantly to analyzing different aspects of the chemical industry and its development dynamics. These articles are valuable for general analysis. The first study we examined is John P. Doyle's "Modeling Chemical Industry Growth Using a System Dynamics Approach"[1].

This article explores modeling the growth of the chemical industry using system dynamics. The system dynamics approach takes into account the interaction of the industry's economic, technological, and ecological systems. The article demonstrates the effective use of system dynamics in modeling factors influencing industrial growth, such as resources, supply and demand, technological innovations, and competition. This method allows for a deep analysis of the changes and evolution of the chemical industry.

World Economics & Finance Bulletin (WEFB)
Available Online at: https://www.scholarexpress.net
Vol. 43,February, 2025
ISSN: 2749-3628,

According to the model results, it is emphasized that ensuring the sustainable growth of the chemical industry requires considering temporary technological innovations, strategic investments, and ecological sustainability. This approach will be useful in studying the long-term development trends of the industry.

The next study we reviewed is X.N. Sabirov's "Modeling the Production Volume of the Food Industry"[2]. In this article, the production volume of the food industry sector was modeled using the Cobb-Douglas function. Initially, the article provides a brief description of scientific sources related to the topic. Then, issues such as the regression model, determination, correlation coefficients, the Durbin-Watson test, and empirical probability regarding the Fisher criterion are discussed. In the final section of the article, all the obtained results are summarized.

Another study we analyzed is Yuan Yo, Kai Lan, Thomas E. Graedel, and Narasimha D. Raol's "Decarbonization Models in the Chemical Industry"[3]. This article proposes various technologies and strategies for decarbonizing the chemical industry. Evaluating the decarbonization, ecological, and economic consequences of these technologies and strategies is crucial for identifying pathways to a more sustainable industrial future. This study explores the latest achievements and integration of system analysis models, including process analysis, material flow analysis, life cycle assessment, techno-economic analysis, and machine learning. These models are classified based on promising decarbonization technologies (e.g., carbon capture, storage, and utilization, biomass feedstocks, and electrification) and circular economy strategies across different scales (micro, meso, and macro).

Incorporating forward-looking, data-driven approaches into existing models enables the optimization of complex industrial systems and assessment of future impacts. While advancements in system-level modeling based on industrial ecology, economics, and planetary boundaries support a more integrated evaluation, further efforts are needed to consider the impact on ecosystems. The effective application of these advanced, integrated models requires interdisciplinary collaboration across chemical engineering, industrial ecology, and economics.

METHODOLOGY

In the previous chapter, we conducted a literature review related to modeling the development dynamics of the chemical industry. In this chapter, we will develop the methodological part of our research and define strategies to address the gaps identified in the literature review.

In this section, we will perform econometric analysis using selected variables, with the assistance of the "Stata" software. By utilizing econometric analyses, the most optimal model will be selected. The findings of our study have shown that the linear model is the most optimal for our research.

$$\overline{Y} = a_0 + a_1 x + a_1 x^2 + \ldots + a_n x^n$$

In this, a_0, a_1 and a_n are considered as the selected parameters.

The main objective of multiple regression is to study the interaction of several factors and their influence on the dependent variable, and to find the optimal model that expresses their combined effect on the outcome variable.

$$r_{yx} = \frac{\overline{yx} - \overline{x} * \overline{y}}{\sigma_x * \sigma_y} \quad [4]$$

The quality of the regression equation or the relationship model is characterized by the coefficient of determination.

$$R^2 = 1 - \frac{D_{residual}}{D(y)} \quad yoki \quad R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

[5]

The coefficient of determination varies between 0 and 1:

0≤R^2≤1

The significance of the research model can be determined using the Fisher criterion and approximation error. According to Fisher's F-test, the statistical significance of the regression equation is tested. The F-value can be expressed in terms of the R^2 coefficient of determination as follows:

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - k - 1}{k} \quad [6]$$

Approximating error is the average relative deviation of the theoretical value y from the actual value Y:

$$\varepsilon = \frac{1}{n} \sum \left| \frac{y_i - \hat{y}}{y_i} \right| * 100\%$$

[7]

here, n is the number of observations,

y_i – represents the actual values of the independent variable,

$\hat{y}$ – represents the fitted (theoretical) values of the independent variable.

If the ε value does not exceed 10-12 percent, the constructed regression equation is considered satisfactory.

World Economics & Finance Bulletin (WEFB)
Available Online at: https://www.scholarexpress.net
Vol. 43,February, 2025
ISSN: 2749-3628,

The Durbin-Watson statistic (denoted as *dw* is used to check for the independence of residuals, i.e., to determine whether autocorrelation exists or not.

$$dw = \frac{\sum_{i=1}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2} \qquad [8]$$

here, the *dw* statistic ranges from 0 to 4:

$$0 \le d \le 4$$

To formulate the methodology section, we will initially gather the necessary data for our research from various databases and public records from government agencies in our country. Specifically, we used the World Development Indicators (WDI), which are indicators established by the World Bank, to build the database. Additionally, we utilized the official data repository from the State Statistics Committee, which provides and analyzes both internal and external statistical data of our country, available at stat.uz.

When using secondary data, our research focuses on economic indicators of our country, utilizing data from the last decade. Furthermore, the database was structured in a panel data format.

ANALYSIS AND RESULTS

In constructing the econometric model, we based our hypothesis on the following:

H_0 - At least one of the seven factors has an impact on the dependent variable (Chemical products production %).

H_1 - None of the seven factors have an impact on the dependent variable (Chemical products production %).

To test the null hypothesis above, we will construct the following theoretical model:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon$$

Here:

Y – Percentage of chemical product production;

x_1 – Agricultural products grown by organizations engaged in agricultural activities;

x_2 – Growth rate of investments in fixed capital % (annual);

x_3 – Volume of foreign trade turnover (by regions) in million USD;

x_4 – Mining industry and open pit mining (per thousand people);

x_5 – Trade services;

ε – Error term;

β – Coefficients;

α – Intercept term.

Table 1

Correlation-Regression Statistical Analysis for Selected Dependent and Independent Variables[1]

years	Y(x)	X_1	X_2	X_3	X_4	X_5
2010	108.5	101.6	104.2	22199.2	121.4	121.5
2011	107.1	105	102.6	26365.9	125.7	115.7
2012	102.9	113.7	110.6	26416.1	126	115.4
2013	103.7	101.5	111.3	28269.6	127.3	113.9
2014	104.8	118.4	109.8	27530.1	128.6	115.7
2015	107.5	103.6	109.4	24924.3	130.3	118.5
2016	132.7	90.8	104.1	24232.2	133.2	120.5
2017	94	110.2	119.4	26566.1	83.5	100.3
2018	97.2	135.5	129.9	33430	78.7	104.9
2019	102.3	159.6	138.1	41751	84.8	107.4
2020	107.6	147.5	95.6	36256.1	92.5	103.8
2021	107	129.1	102.9	42170.5	94	112.3
2022	98.1	123.9	100.2	50500.3	57.5	108.3
2023	96.7	112.8	123.4	63528.6	58.9	110.6

In selecting these factors, we will use the Stata18 software to generate a correlation matrix for correlation analysis. This will help identify which factors are strongly correlated with industrial product volume and determine their participation in the model. In econometric modeling, the correlation-regression analysis method will be used.

	Y	X1	X2	X3	X4	X5

[1] Created by the author.

World Economics & Finance Bulletin (WEFB)
Available Online at: https://www.scholarexpress.net
Vol. 43,February, 2025
ISSN: 2749-3628,

Y	1.0000					
X1	-0.3906	1.0000				
	0.1673					
X2	-0.4357	0.3788	1.0000			
	0.1194	0.1817				
X3	-0.4176	0.4313	0.2462	1.0000		
	0.1374	0.1236	0.3961			
X4	0.6118	-0.5497	-0.3565	-0.8353	1.0000	
	0.0201	0.0417	0.2109	0.0002		
X5	0..6391	-0.6515	-0.3767	-0.4079	0.7444	1.0000
	0.0137	0.0116	0.1843	0.1477	0.0023	

Figure 1: Correlation Matrix[2]

We used Pearson's correlation coefficient, and based on the p-value of all the results, we can observe that only two independent variables are significantly related to Y. Therefore, we will construct a model related to these two independent variables.

Regression analysis will be conducted to derive the regression equation from the results obtained using the software tool.

Source	SS	df	MS			
				Number of obs	=	14
				F(2, 11)	=	4.51
Model	501.26328	2	250.63164	Prob > F	=	0.0371
Residual	611.465926	11	55.5878114	R-squared	=	0.4505
				Adj R-squared	=	0.3506
Total	1112.72921	13	85.5945543	Root MSE	=	7.4557

Y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
X4	.1024925	.1128253	0.91	0.383	-.1458342	.3508192
X5	.5941083	.4810853	1.23	0.243	-.4647534	1.65297
_cons	27.87341	45.95986	0.61	0.557	-73.28356	129.0304

Figure 2: Results of Regression Analysis[3]

As a result of the regression analysis, a linear model was used. Based on this, we exclude one of the exogenous factors from the model according to the reliability coefficient of the factors based on the p-value.

[2] Calculated by the author using the Stata18 software package

[3] Calculated by the author as part of the research study.

World Economics & Finance Bulletin (WEFB)
Available Online at: https://www.scholarexpress.net
Vol. 43,February, 2025
ISSN: 2749-3628,

Source	SS	df	MS			
				Number of obs	=	14
				F(1, 12)	=	7.18
Model	416.488524	1	416.488524	Prob > F	=	0.0201
Residual	696.240682	12	58.0200568	R-squared	=	0.3743
				Adj R-squared	=	0.3222
Total	1112.72921	13	85.5945543	Root MSE	=	7.6171

Y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
X4	.2062123	.0769666	2.68	0.020	.0385166	.3739081
_cons	83.76138	8.1869	10.23	0.000	65.92366	101.5991

3-rasm. Liner model natijalari[4]

In the regression equation with respect to one variable, X4 (Mining industry and open pit mining (per thousand people)) showed better results compared to the other variables. Based on the direct regression analysis, the linear-log model took the following form.

$$\hat{Y} = 83.761 + 0.206 * X4 + \varepsilon \text{ (2)}$$

$$Percentage\ of\ chemical\ p, p, = 83.761 + 0.206 \cdot Mining\ industry,$$

Based on the results of the calculated linear model regression analysis, it can be seen that all the results are significant, and positive test results have been achieved. The determination coefficient, R_2 = 0.3743, indicates that the model explains 37% of the real values. Additionally, the F-statistic value of the model is F_{his} = 7.18 \), which is significant according to the p-value. Thus, the linear model we have constructed is significant, and each parameter is reliable based on the t-Student value. The calculated values through the model and the actual real industrial product volume values are represented in a single graph (Figure 4).

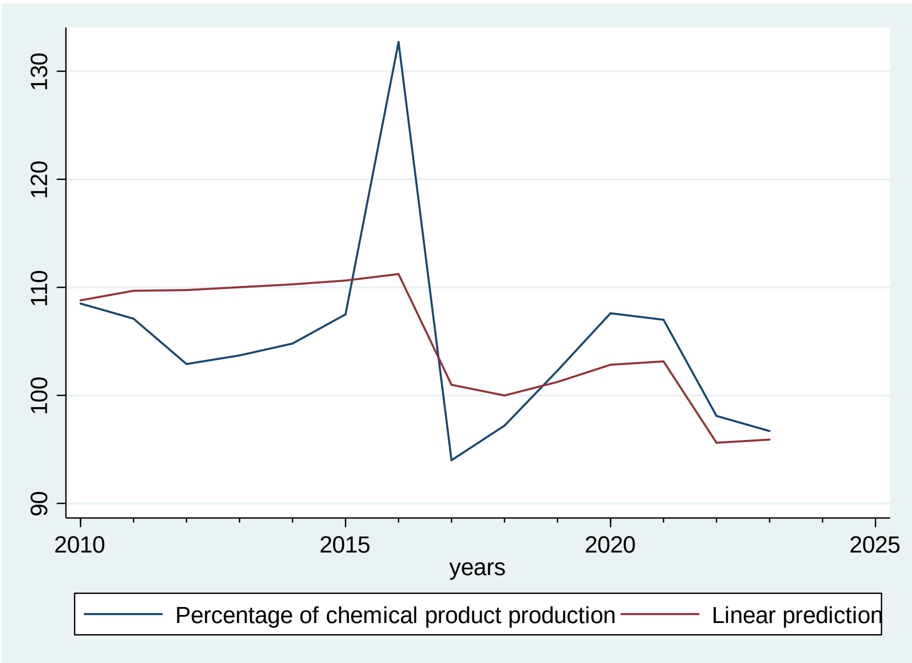


Figure 4: Results obtained through the linear model and actual values[5]

[4] Calculated by the author as part of the research study.
[5] Author's calculations generated using Stata18 software package.

World Economics & Finance Bulletin (WEFB)
Available Online at: https://www.scholarexpress.net
Vol. 43,February, 2025
ISSN: 2749-3628,

Since the model produced positive results, the next step will be to proceed to the forecasting stage. To obtain forecast values, AR models will be used for each exogenous variable, and the results will be derived. Afterward, each result will be applied to the linear model, and the empirical forecast values for chemical product production volumes will be generated.

Source	SS	df	MS			
				Number of obs	=	13
				F(1, 11)	=	18.32
Model	5892.35167	1	5892.35167	Prob > F	=	0.0013
Residual	3538.48549	11	321.680499	R-squared	=	0.6248
				Adj R-squared	=	0.5907
Total	9430.83715	12	785.903096	Root MSE	=	17.935

X4	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
X4						
L1.	.8749403	.2044309	4.28	0.001	.424991	1.32489
_cons	8.501548	22.3176	0.38	0.711	-40.61915	57.62225

Figure 5: AR(1) model results for the quantity of mining industry and open pit mining (per thousand people)[6]

If we focus on the AR(1) model results, all test results are positive, indicating that the model is significant and of good quality. Based on this, the medium-term empirical forecast values for the quantity of mining industry and open-pit mining (per thousand people) (X4) will be derived.

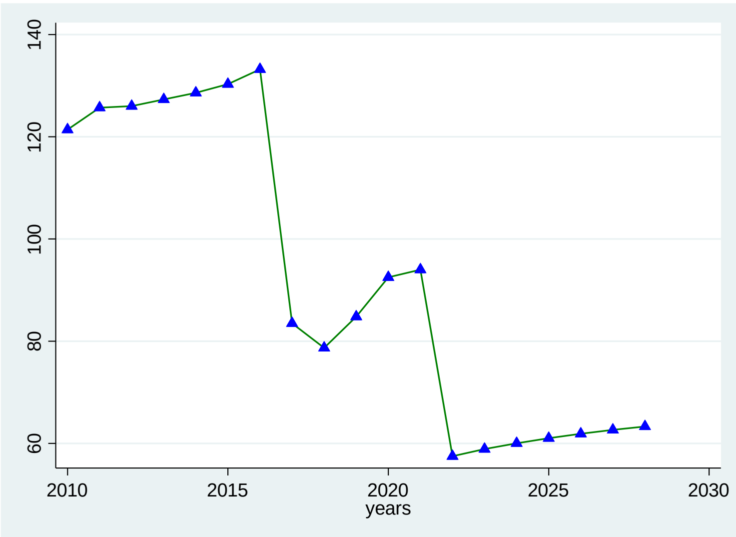
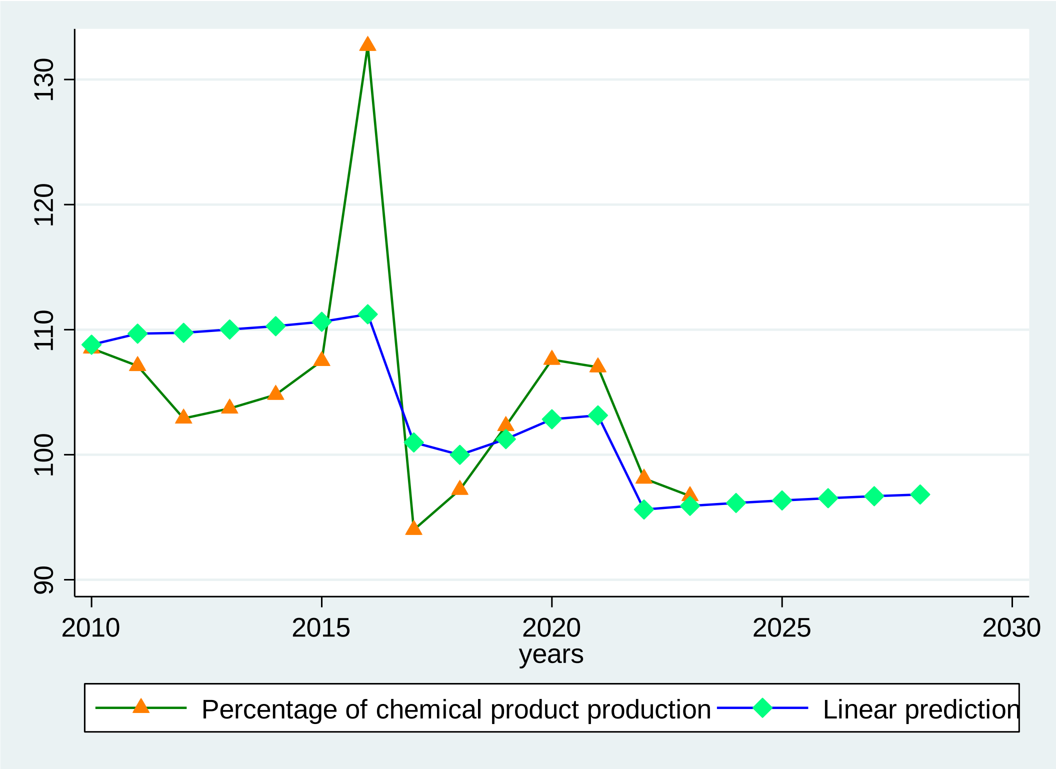


Figure 6: Size of mining industry and open-pit mining (per thousand people) and forecast values[7]

Based on the results of the regression analysis, the generated AR(1) model is significant according to the F-statistic, and the model parameters are reliable based on the t-statistic. The size of the mining industry and open-pit mining (per thousand people) is projected to increase to 63.322 billion soums by 2028, based on empirical data. The forecast values for all exogenous variables obtained through the AR(1) models are then applied to the 2nd formula, i.e., the linear model, to derive the empirical forecast values for the production volume of chemical products.

[6] Calculated by the author as part of the research study.
[7] Author's calculations generated using Stata18 software package.

World Economics & Finance Bulletin (WEFB)
Available Online at: https://www.scholarexpress.net
Vol. 43,February, 2025
ISSN: 2749-3628,



7-figure. The volume of chemical product production (%) from 2010 to 2023 and the empirical forecast value[8]

The forecast values of the exogenous variables mentioned above, based on empirical data, suggest that the size of the mining and open-pit mining sector will reach 60,035 thousand people by 2024, 61,028 thousand people by 2025, 61,897 thousand people by 2026, 62,657 thousand people by 2027, and 63,322 thousand people by 2028. This directly indicates that the production volume of chemical products will continue to grow in a stable manner.

CONCLUSION

The process of econometric modeling of the dynamics of the chemical industry development is of great importance in forecasting changes and developing strategies, considering various economic and ecological factors for this sector. In the study, multiple regression analysis was used to examine the growth rates of the chemical industry and their relationship with various factors.

The analysis results showed that the production volume of chemical products is significantly related to only a few factors, such as the mining industry and open-pit mining. This factor's reliability and significance in the regression analysis are high, confirming the stability and efficiency of the model. The coefficient of determination (R^2 = 0.3743) and Fisher statistics (F = 7.18) validate the significance of the regression equation and indicate the good fit of the constructed model.

Additionally, it was found that good results can be obtained by using the AR model to forecast future growth rates. The obtained forecasts could be useful for optimizing the future development of the chemical industry and efficiently managing resources.

The main conclusion of the study is that the development of the chemical industry can be predicted through complex and systematic approaches, taking into account both economic and ecological factors in the management of this sector. In the future, it is important to improve these models, developing them with the consideration of new technologies and ecological sustainability.

Furthermore, it is necessary to develop strategic approaches to ensure the sustainable development of the industry and implement them in

[8] Author's calculations generated using Stata18 software package.

practice. This, in turn, will help improve the economic and ecological conditions of countries and increase the competitiveness of the chemical industry.

REFERENCES

1. Modeling of Chemical Industry Growth Using System Dynamics Approach" by John P. Doyle (2015)
2. https://www.researchgate.net/publication/261061057_Understanding_US_Biodiesel_Industry_Growth_using_System_Dynamics_Modeling
3. Сабиров, Х. (2020). Моделирование объемов производства пищевой промышленности. Экономика и инновационные технологии, (6), 124–131.
4. https://inlibrary.uz/index.php/economics_and_innovative/article/view/11837
5. Models for Decarbonization in the Chemical Industry Yuan Yao, Kai Lan, Thomas E. Graedel, and Narasimha D. Rao Vol. 15:139-161 (Volume publication date July 2024) https://doi.org/10.1146/annurev-chembioeng-100522-114115
6. Sabirov, K. (2024). KOP OMILLI KORRELYATSION-REGRESSION TAHLIL. *Nordic_Press*, *2*(0002).
7. On the determination coefficient in robust regressionDirk TascheFirst draft: June 28, 2000This revision: November 13, 2000
8. https://www.researchgate.net/publication/2376503_On_the_Determination_Coefficient_in_Robust_Regression
9. When to Use Fisher's Exact Test. Keith M.Bower, M.S
10. https://www.researchgate.net/publication/265026286_When_to_Use_Fisher's_Exact_Test
11. Ding-Xuan Zhou ESTIMATING THE APPROXIMATION ERROR IN LEARNING THEORY November 2011, Analysis and Applications 01(01), DOI:10.1142/S0219530503000089
12. Sabirova, N., & Sadikova, M. (2023). Shortenings based on hydrogenated cotton oils. In E3S Web of Conferences (Vol. 390, p. 02038). EDP Sciences.
13. https://www.researchgate.net/publication/263790633_ESTIMATING_THE_APPROXIMATION_ERROR_IN_LEARNING_THEORY
14. Demonstrating the Durbin-Watson Statistic, December 1998, Journal of the Royal Statistical Society Series D (The Statistician) 47(4):643-644, DOI:10.1111/1467-9884.00161
15. https://www.researchgate.net/publication/243086482_Demonstrating_the_Durbin-Watson_Statistic
16. www.stat.uz
17. www.arxiv.uz
18. www.webofscience.com