Regression Models of Agricultural Production: Evaluation of Data on Russian Regions

Pinkovetskaia Iuliia[1](✉), Aleksandrova Natalya[2], Treskova Tatyana[2], and Berezina Natalia[3]

[1] Ulyanovsk State University, 432000 Ulyanovsk, Russia
pinkovetskaia@gmail.com
[2] Ulyanovsk State Agrarian University named after P.A. Stolypin, 432000 Ulyanovsk, Russia
[3] Chuvash State University, 428015 Cheboksary, Russia

Abstract. The study was based on the development of regression models that characterize the activities of agricultural enterprises in the regions of Russia. Official statistical information on 65 regions of Russia for 2017 and 2018 was used. The conducted research made it possible to identify factors (fixed assets, wages of employees, ratio volume of crop production to livestock production, arable land area) that affect the volume of production in the agricultural sector in the regions of Russia and suggest using four-factor regression models of high quality to describe this influence. The developed regression models are effective management tools that allow assessing the level of use of financial and labor resources. The acquired new knowledge and tools for assessing the activities of agriculture in the Russian regions are of academic and utilitarian value. Results of this study can be used in measuring of the agricultural sector of the economy, monitoring of production volumes, in determining the needs for resources necessary for the development of agriculture, substantiating plans and programs for its development.

Keywords: Agriculture production · Regression model · Volume of production · Regions of Russia

1 Introduction

To date, agriculture in Russia has received significant development. Currently, it is among the top four countries that have the largest areas of arable land. According to some estimates, about 9% of the world's farmland is located in Russia [17]. The solution of managerial tasks in the Russian economy requires an understanding of the factors that affect the volume of agricultural production. In Russia, there is currently an urgent need for accelerated development of agriculture. The efficiency of agricultural production, as one of the directions of increasing the productivity and competitiveness of this branch of the economy, is directly related to the use of resources, with the degree of their involvement in the production process. To a large extent, efficiency depends on the quantitative and qualitative ratio of resources among themselves, on their balance. Determining the cost structure that ensures an increase in output per unit of resource becomes an urgent

B. Alareeni and A. Hamdan (Eds.): ICBT 2022, LNNS 495, pp. 847–855, 2023.
https://doi.org/10.1007/978-3-031-08954-1_71

task of the management system. The justification of the resources mandatory for the successful operating of agriculture can be relying on regression models.

Scientific publications conducted in the twenty-first century have demonstrated the possibility of economic analysis of the activities of farmers who specialize in crop production and animal husbandry using regression models. These mathematical models describe the dependence of agricultural output volumes on factors describing capital and labor costs [4, 15, 16]. Most scientific publications considered data for a number of years (the so-called time series). For example, in the article [8], according to the agricultural sector of India, the efficiency of farms was evaluated using the Cobb-Douglas production function. Evaluation of agricultural production in China was described in the article [1]. While data were analyzed in 12 prefectures during the period from 2009 to 2019. In contrast to the above articles, the study [10] developed models using spatial data for 25 provinces of Cambodia. Four models were built corresponding to the information for each of the four years from 2012 to 2015. Capital and labor costs were used as factors influencing agricultural production volumes in most previously performed scientific studies [3, 14, 19, 20].

The purpose of our research was to develop economic and mathematical models to assess the impact of indicators characterizing the use of resources on agricultural production in the regions of Russia. Our study responds to the calls for taking into account the regional characteristics of agricultural production, formulated in publications [12, 23].

Our article makes a certain contribution to the knowledge about the regional peculiarities of the development of agriculture in Russia. The theoretical contribution is related to the methodology proposed by the authors, which makes it possible to assess the dependence of agricultural production volumes on factors such as fixed assets, wages of employees, ratio volume of crop production to livestock production, arable land area discussed on the development of economic and mathematical models representing regression models. Based on empirical data in the course of the study, new knowledge was obtained about the impact of each factor on the volume of agriculture production. In addition, regions were identified in which high and low values of resource efficiency were noted.

The structure of this article is given below. The following section provides an overview of scientific publications characterizing the production volumes of the agricultural sector in Russia and its regions. The third section presents the methodology and design of the study. The results of empirical data modeling are given in section four. The fifth section is devoted to the discussion of the developed regression models. The sixth section contains conclusions, followed by bibliographic references.

2 Literature Review

Regression models describing the activities of enterprises, organizations and farmers in the agricultural sector of Russia and its regions aroused some interest among researchers. The most interesting of such studies carried out in recent years are presented in Table 1.

Table 1. Characteristics of Russian studies.

Authors	Factor of capital	Factor of labor	Research object
1	2	3	4
Shestakov and Yakovlev, 2020 [21]	Capital expenditures	Labor costs	Agricultural production volumes for 2005–2018 in Russia as a whole
Tolmachev, 2011 [22]	Indices of the physical volume of fixed assets	index of total working time expenditures	Indices of the physical volume of agricultural products in Russia for 1996–2008
Potapov, 2020 [18]	Costs of mechanical engineering products, fuel and energy resources, chemical products	–	Gross agricultural output in Russia in 2011–2015
Naumov, 2017 [13]	Fixed capital	Number of employees	Production volumes in agriculture in the Chelyabinsk region for 2005–2015
Kutenkov, 2020 [11]	Cost of fixed assets	Number of people employed in agriculture per 100 hectares of acreage	three groups of regions of Russia, data for 2017
Zyukin and Zhilin, 2014 [25]	Production funds	Value of labor costs	Volume of production of the agricultural complement of the Kursk region for 2000–2011
Zhilyaskova, 2008 [24]	Fixed assets	Number of employees	Agricultural production volumes in the Rostov region for the period 2004–2006
Germanova and Rudaya, 2017 [7]	Fixed assets	Number of employed workers	Production volumes in agriculture of the Krasnodar territory for 2000–2014

Source: Compiled by the authors.

The data in Table 1 show that in most cases the objects of research are agricultural sectors in specific regions (five cases). The other three publications discuss production functions for Russia as a whole. The initial data in seven studies were time series, only one publication used spatial data for one year. In most studies (seven cases), the number of employees was used as labor costs. In two publications, the values of working time costs were considered. Data on fixed assets of agricultural enterprises were used as capital factors in six publications. In two cases, the costs of production assets were considered, and in one case, the costs of purchasing products from the machine-tool, fuel and chemical sectors. In addition, in one of the articles [24], the area of arable land was used as a factor of the production function. Thus, previous Russian scientific publications did not pay sufficient attention to the comprehensive assessment of regional characteristics of agricultural production in Russia.

3 Methodology and Design

The objects of our research were agricultural complexes and individual farmers who were engaged in crop production and animal husbandry, as well as related activities in each region of Russia. The development of regression models using time series (data for fifteen years or more) does not seem appropriate, since there is a large inflation in Russia. Taking this into account, a methodological approach was used based on the study of data on a large number of regions in one year. Since agriculture in Russia has been widely developed in sixty-five regions, the amount of empirical data was significant and met the requirements for the development of high-quality regression models. It should be noted that the advantage of using spatial data in evaluating such models compared to data for a number of years was demonstrated in the article [2].

As the factors that have the greatest impact on the total output of agricultural products, the following were considered in our study: the total cost of all capital assets in the agricultural sector of each region (factor 1), the total labor costs of agricultural workers in each of the regions (factor 2), the ratio volume of crop production to livestock production in each of the regions (factor 3), arable land area in each of the regions (factor 4). This conclusion followed from the correlation analysis of the influence of these factors on the resulting indicator, that is, the output of agricultural products. At the same time, there was no collinearity between the factors and the resulting indicator. It is essential that the use of these factors and the resulting indicator, as shown in the article [6], provides a good approximation of the initial data, since they all have the same dimension. The empirical data in our study were official statistical data for sixty-five regions of Russia for 2017 and 2018 [5]. In our study, three hypotheses were tested:

- the first hypothesis is that regression models can be used to model the production volumes of the agricultural sector in the regions.
- the second hypothesis is that regression models demonstrate the presence of stable dependence of agricultural production volumes on factors such as fixed assets, wages, ratio volume of crop production to livestock production, arable land area.
- the third hypothesis is that the factors of the total value of fixed assets of agricultural enterprises, as well as the volume of arable land available to them in both regression models affect turnover to a greater extent than the other two factors.

In the course of the study, two regression models were developed, reflecting the dependence of agricultural production volumes on the total cost of all capital assets in the agricultural sector and the total labor costs of agricultural workers in each of the regions, ratio volume of crop production to livestock production, arable land area.

4 Results

Below are the first and second regression models designed on the base of data for 2017 and 2018:

$$y_1(x_1, x_2, x_3, x_4) = 2.056 \times x_1^{0.286} \times x_2^{0.150} \times x_3^{0.035} \times x_4^{0.324} \quad (1)$$

$$y_2(x_5, x_6, x_7, x_8) = 2.238 \times x_5^{0.303} \times x_6^{0.173} \times x_7^{0.023} \times x_8^{0.292} \quad (2)$$

where y_1, y_2 - total output of agricultural products in each of the regions, billion rubles;

x_1, x_5 - total cost of all capital assets in the agricultural sector in each of the regions, billion rubles;

x_2, x_6 - total labor costs of agricultural workers in each of the regions, billion rubles;

x_3, x_7 - ratio volume of crop production to livestock production in each of the regions;

x_4, x_8 - arable land area in each of the regions, thousand hectares.

Table 2 shows the analysis of the model's quality. It presents the calculated values of the correlation and determination coefficients, Fisher-Snedecor and Student's tests (column 2), as well as the significance of the Fisher-Snedecor test and p-values for Student's test (column 3).

The correlation coefficients more 0.9 and close to 1 in both regression models. Regression models are known to be of high quality when determination indexes are more than 0.8. The difference between 1 and this coefficient demonstrates the effect of variables not included in the regressions under consideration is 8.3%. The calculated statistic values (166 and 165) are higher than the table value of the Fisher-Snedecor test, which is 3.98 at a significance level of 0.05. For both regression models, all calculated Student test values for the coefficient and the exponents are in the range from 2.67 to 6.75; in absolute value they exceed the table amount, which is 1.99 at a significance level of 0.05. Results presented in Table 2 allow us to conclude that there is a high quality correlation between the resulting values and the four factors of the regression models (1) and (2). All levels of significance given in column 3 of Table 2 have values less than 0.01. Therefore, the coefficients of the developed regression models and the degree values in these regression models are statistically significant with the precision of 99%.

The data obtained allow us to make a general conclusion that the developed regression models (1) and (2) fully meet the econometric requirements and, therefore, can be used to describe the dependencies of agricultural production volumes in the regions from discussing factors. Consequently, the first hypothesis was confirmed.

Table 2. Values of calculated statistics.

Characteristics	Model (1)	Model (2)	Significance level
Determination coefficient	0.917	0.917	–
Correlation coefficient	0.958	0.957	–
Standard error	0.257	0.258	–
Calculated value of the Fisher-Snedecor test	166.260	164.905	Less than 0.01
Calculated value of the Student's test for y-intersection	2.675	3.041	Less than 0.01
Calculated value of the Student's test for x_1 and x_5	5.088	5.906	Less than 0.01
Calculated value of the Student's test for x_2 and x_6	3.457	2.882	Less than 0.01
Calculated value of the Student's test for x_3 and x_7	5.544	3.638	Less than 0.01
Calculated value of the Student's test for x_4 and x_8	6.748	5.999	Less than 0.01

Source: Calculated by authors.

5 Discussion

The developed regression models (1)–(2) prove the influence of fixed assets, wages, the ratio of the volume of crop production to the production of livestock products, the area of arable land on the volume of production of enterprises and entrepreneurs belonging to the rural sector. The developed regression models show the presence of established stable dependencies of agricultural production volumes in the regions on the factors under consideration for the period from 2017 to 2018. Thus, the second hypothesis was confirmed.

The degree values for the four factors in the developed models are greater than zero. Consequently, an increase in the values of each of the four factors can be used to increase the total volume of agricultural production. In the entire range of changes in the values of factors, the resulting indicators do not reach the maximum values. This indicates the possibility of increasing agricultural production in each of the regions of Russia under consideration. In all regions there are significant reserves for the further development of enterprises in this sector of the economy, including on the basis of the following measures:

- increasing the number of enterprises and the number of employees employed in them;
- increasing the volume of fixed assets;
- increasing the ratio between crop production and animal husbandry;
- expansion of the area of arable land or yield.

Factors of the total value of fixed assets of agricultural enterprises, as well as the volume of arable land available to them in both regression models affect turnover to a greater extent compared to the other two factors. This follows from the comparison of the values of the degrees in the first and second regression models. Thus, the third hypothesis was confirmed. Comparison of total output of agricultural products according to data for 2017 and 2018 shows that the values of this resulting indicator (equal to the sum of the values of degrees in regression models) are almost the same and amount to 0.795 (regression model 1) and 0.790 (regression model 2). This suggests that with the simultaneous increase of four factors, the growth of agricultural production over the years under review was almost the same. The return on scale in agriculture over the years under review was less than 1. This situation is due to the fact that most farmers have a small number of workers. Therefore, the possibilities of specialization of employees are limited, and they are forced to perform a variety of functions. As shown in [9], this leads to a relatively low level of personnel training, a decrease in labor productivity and, as a consequence, low resource efficiency at such enterprises.

To increase agricultural production in the Russian regions, it is advisable to ensure the simultaneous growth of all four factors.

A comparative analysis of the actual values of production volumes and the data predicted on the basis of the regression model (1) showed a high level of resource efficiency in 2017 in the following regions: Krasnodar territory (10.3%), Samara region (12.2%), Orenburg region (12.6%), Rostov region (12.6%), republic of Tatarstan (13.6%), Altai republic (13.8%), Saratov region (13.9%), Volgograd region (14.1%), Trans - Baikal territory (17.5%), Kurgan region (22.0%), republic of Kalmykia (23.7%), republic of Tyva (28.5%). The deviations of the actual values from the predicted values are indicated in parentheses. The low level of use of the considered factors of production was in such regions as Khabarovsk territory (−19.3%), Vologda region (−16.6), Tver region (−12.3%), Kirov region (−12.3%), Primorsky territory (−11.9%), Kostroma region (−11.8%), Vladimir region (−11.5%), Kaluga region (−11.5%), Yaroslavl region (−11.1%).

A comparative analysis of the actual values of production volumes and the data predicted on the basis of the regression model (2) showed a high level of resource efficiency in 2018 in the following regions: Krasnodar territory (10.1%), Orenburg region (10.8%), Rostov region (10.9%), Saratov region (11.1%), Volgograd region (11.7%), republic of Tatarstan (11.8%), Samara region (12.3%), Altai republic (13.2%), republic of Sakha (14.2%), Kurgan region (19.9%), republic of Kalmykia (21.6%), republic of Tyva (32.2%). The low level of use of the considered factors of production was in such regions as Vologda region (−15.8%), Khabarovsk territory (−15.8%), Primorsky territory (−13.9%), Kostroma region (−12.4%), Kirov region (−12.4%), Kaluga region (−11.2%), Yaroslavl region (−10.9%), Vladimir region (−10.7%), Tver region (−10.5%), Smolensk region (−10.1%). The above lists showed that most of the regions in 2018 retained their characteristics shown in 2017.

6 Conclusion

The conducted research has a certain scientific and practical significance. The scientific significance of the study is as follows:

- methodology for the development of regression models demonstrating the dependence of the total volume of agricultural production on four factors - the total cost of all fixed assets in the agricultural sector, the total labor costs of agricultural workers in each of the regions, the ratio of crop production to livestock production, the area of arable land in the agricultural sector. The methodology provided for the use of spatial data by region, characterizing the values of the four factors under consideration according to the results for one year. In our study, these were 2017 and 2018.
- two four-factor regression models were developed during the study. These regression models describe the dependence of production volumes in the agricultural sectors of each region on the factors under consideration.
- an increase in production in the agricultural sector improves the possibility of replacing each of the four factors with another.
- in our study, a ranking was conducted and regions were identified that are characterized by maximum and minimum use of resources.

The proposed regression models allow us to estimate the use of labor costs of agricultural workers, the cost of all fixed assets, the ratio of crop production to livestock production, the area of arable land in the agricultural sector. Therefore, it is advisable to use them when justifying programs and plans for the strategic development of regions. That is, to assess how effectively resources are being used. In addition, regression models allow us to identify an imbalance in the values of factors for each of the regions. Regression models can be used to justify programs to increase each of the four indicators, to form plans and programs for further development of agriculture.

There were limitations in the research process, since 65 regions of Russia were considered in which agricultural production has received significant development. At the same time, data on 17 regions of Russia in which the agricultural sector has not received significant development were not taken into account when constructing production functions. Further research may be related to the development of similar functions in the years following the publication of the relevant official statistics.

References

1. Binghun, W., Zhou, E.: Research of total factor productivity and agricultural management based on Malmquist-DEA modeling. Hindawi. Math. Probl. Eng. **2021**, 2828061 (2021)
2. Charoenrat, T., Harvie, C.: Technical efficiency of thai manufacturing SMEs: a stochastic frontier analysis. Aust. Account. Bus. Financ. J. **7**(1), 97–122 (2013)
3. Czyzewski, B., Majchrzak, A.: Economic size of farms and adjustments of the total factor productivity to the business cycle in polish agriculture. Agric. Econ. **63**, 93–102 (2017)
4. Czyzewski, B., Smedzik-Ambrozy, K.: The regional structure of the CAP subsidies and the factor productivity in agriculture in the EU 28. Agric. Econ. **63**, 149–163 (2017)
5. Federal State Statistics Service. https://rosstat.gov.ru/folder/210/document/13226
6. Felipe, J., McCombie, J.: Problems with regional production functions and estimates of agglomeration economies: a caveat emptor for regional scientists. Levy Economics Institute of Bard College Working Paper №. 725, May 2012. http://www.levyinstitute.org/pubs/wp_725.pdf
7. Germanova, O.E., Rudaya, Y.N.: Dynamics of parameters and type of technological progress in agriculture. Regional Econ. South Russia **3**(17), 158–172 (2017)

8. Ghoshal, P., Goswami, B.: Cobb-Douglas production function for measuring efficiency in Indian agriculture: a region-wise analysis. Econ. Aff. **62**(4), 573–579 (2017)
9. International Labour Conference, 104th Session 2015 Report IV Small and medium-sized enterprises and decent and productive employment creation. International Labour Office (ILO), Geneva (2015)
10. Kea, S., Li, H., Pich, L.: Technical efficiency and its determinants of rice production in Cambodia. Economies **4**(4), 1–17 (2016)
11. Kutenkov, R.P.: Methodology and results of factor forecasting of the dynamics of gross output and labor productivity in agriculture of the regions of the Russian Federation using production functions. Ostrov. Read. **1**, 99–103 (2020)
12. Margono, H., Sharma, S.C.: Technical efficiency and productivity analysis in indonesian provincial economies. Appl. Econ. **43**(6), 663–672 (2011)
13. Naumov, I.V.: Problems of forecasting gross output in the regional socio-economic system. J. Econ. Theory **4**, 68–83 (2017)
14. Nowak, A., Kijek, T., Domacska, K.: Technical efficiency and its determinants in the European Union. Agric. Econ. **61**, 275–283 (2015)
15. Parlinska, M., Dareev, G.: The agricultural production in mathematical models. Probl. World Agric. **11**(26), 1–5 (2011)
16. Petrick, M., Kloss, M.: Identifying factor productivity from micro-data: the case of EU agriculture, Discussion Paper, No. 171, Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Halle (Saale) (2018). http://nbn-resolving.de/urn:nbn:de:gbv:3:2-90157
17. Petrikov, A.V.: The necessity and main features of the new agrarian policy in Russia. APC: Econ. Manage. **12**, 24–34 (2020)
18. Potapov, A.P.: Estimation of agricultural production volumes depending on the structure of resource costs. Econ. Sci. **6**(187), 74–79 (2020)
19. Prager, D.L., Foltz, J.D., Barham, B.L.: Making time for agricultural and life science research: technical change and productivity gains. Am. J. Agric. Econ. **97**(3), 743–761 (2015)
20. Rezitis, A.N., Kalantzi, M.A.: Investigating technical efficiency and its determinants by data envelopment analysis: an application in the Greek food and beverages manufacturing industry. Agribusiness **32**(2), 254–271 (2016)
21. Shestakov, R.B., Yakovlev, N.A.: Analysis of production potential in agriculture based on production function modelling. Bull. Rural Dev. Soc. Policy **3**(27), 9–12 (2020)
22. Tolmachev, M.N.: Problems of building production functions in Russian agriculture. Account. Stat. **4**(24), 88–94 (2011)
23. Zhang, D., Xie, J., Ermanno, A.: An efficiency and productivity analysis of the agricultural sector in Alabama. Int. J. Appl. Econ. **14**(2), 19–36 (2017)
24. Zhilyaskova, N.P.: Production function in agriculture. Econ. Bull. Rostov State Univ. **6**(4), 63–67 (2008)
25. Zyukin, D.A., Zhilin, V.V.: Cobb-Douglas function when assessing the development of agriculture of Kursk area. Curr. Direct. Sci. Res. XXI Cent.: Theory Pract. **4–2**(9–2), 299–302 (2014)