

# Application of Statistical Methods in Assessing the Role of Agriculture in the Formation of Regional Income

Romantseva Yu.N.\* , Kagirova M.V., Galyautdinova D.F.

*Russian State Agrarian University — Moscow Timiryazev Agricultural Academy, Moscow, Russia*

*\*Corresponding author. Email: romantceva@rambler.ru*

## ABSTRACT

In the new conditions of digitalization of the economy, the issue of the operational analysis of economic processes in various spheres of the economy acquires particular relevance. A feature of proposed methodology is taking into account a large number of various factors that affect income in the agricultural sector. The author's approach to improving the analysis is based on the use of econometric modeling to justify the directions of state agrarian policy. Regions of the Russian Federation were considered as an object of the research. As a result, there were distinguished some groups of regions that differ in soil, climatic and economic conditions of management, and in the efficiency of the agricultural sector. So, in regions with unfavorable agricultural conditions, the agricultural sector is ineffective even if additional investments are engaged. Nevertheless, given the low population density, the industry needs state support as the most important area of rural development. For regions producing a significant volume of agricultural products, government support of the industry is required for providing social stability. And for the growth of gross value added, it is necessary to pursue a policy of reducing the disparity in prices between agricultural and industrial products. In the most developed regions from the point of view of agricultural efficiency, the role of the state is to maintain their investment appeal, since they effectively use the available resources.

**Keywords:** *agriculture, agrarian policy, income of agricultural producers, gross value added, econometric modeling*

## 1. INTRODUCTION

The relevance of the study of the income of agricultural producers in modern conditions, as well as their contribution to the formation of regional value added, is due to the need to maintain the financial stability of the agricultural sector and rural development. This in turn contributes to the food independence of the state. A variety of factors influence the formation of value added in Russian agriculture. On the one hand, this is the continuing disparity in prices for agricultural and industrial products, low technical and technological security of the industry, insufficient level of government support, common for all regions; on the other hand, differences in the soil and climatic features of the regions, their specialization, uneven distribution of government support. Taking into account the strong differentiation of the constituent entities of the Russian Federation in terms of soil and climatic conditions and the level of development of agriculture and the economy as a whole, it seems appropriate to improve the analysis methodology, using the methods of econometric modeling more extensively. The identification and ranking the factors in accordance of their significance is of great importance for determining the directions of state policy in relation to the agricultural sector, which is the main aim of the study.

## 2. LITERATURE REVIEW

The need to improve observation, methodology for calculating and analyzing the income of agricultural producers, as well as the contribution of the industry to the development of regions and the country has been noted by many researchers in Russia [Zinchenko, A.P.. 2017, Ushachev, I.G. 2014] and abroad [Efstratios L. et al. 2019]. To assess the role of agriculture in the formation of value added by the economy in world practice, the indicator of the share of GVA of agriculture, hunting and the provision of services in these areas in GDP is widely used [Spirkova et al. 2017]. Its growth has not only a strong impact on economic recovery and GDP growth. [Đokić 2017], but also determines the degree of agricultural activities in the region.

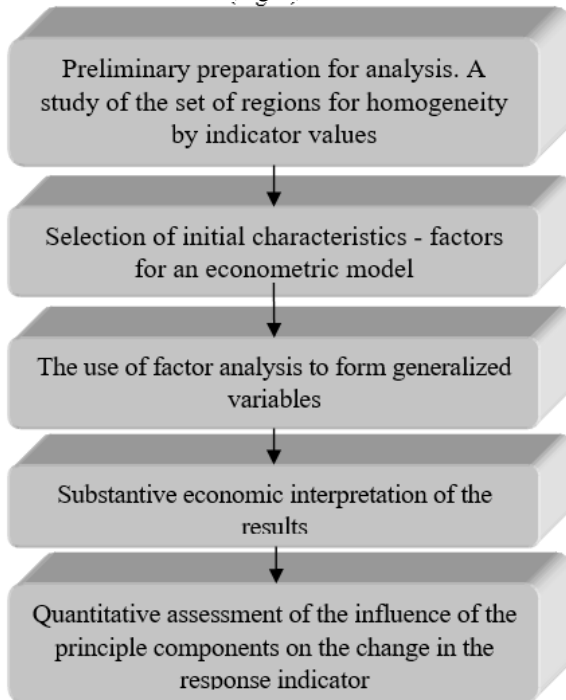
The possibilities of using econometric modeling methods in world practice have significantly expanded thanks to modern software. The most widely used method is Correlation and Regression analysis [Barowski 6]. However, the need to take into account and include in the model a large number of variables that allow a comprehensive assessment of the influence of external and internal factors will lead to multicollinearity, which can be avoided by using factor analysis based on the Principal component analysis (PCA) [Iantovics, L.B. et al. 2019; Dien, J. et al. 2005; Liu, G et al. 2018]. The possibilities of econometric modeling for the detailed

development of more reasonable measures of state regulation are described in the works of many scientists [Bin Li et al. 2017; Lemishko, O.O. 2018], including income of farmers [Herrera, G.P. et al. 2018]. The use of this approach in the analysis of the contribution of agriculture to the development of the region will form the basis for combining similar characteristics-factors into a generalized (integrated) value, which will allow more reasonably to identify and argue the direction of state agricultural policy. Thus, a distinctive feature of this study is the improvement of the methodology for analyzing the income of agricultural producers and their contribution to the formation of regional income based on the inclusion of a large number of factors in the econometric model.

### 3. METHODOLOGY

In order for statistical and economic analysis to meet modern requirements of public administration in speed and scientific justification, it is necessary to use modern research tools, for which STATISTICA Software package was used, which includes the complex of mass data processing methods.

The proposed approach to a comprehensive analysis of the contribution of agriculture to the formation of gross regional product (GRP) includes several combined stages reflected in the scheme (Fig. 1).



**Figure 1** Stages of a comprehensive analysis of the contribution of agriculture in the formation of GRP

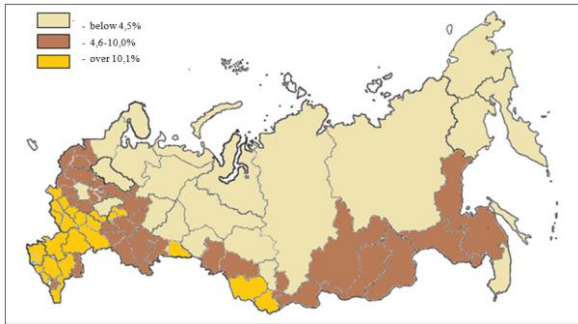
The objects of income research can be enterprises of various forms of managing, as well as regions and municipalities. For them, as multidimensional objects with many characteristics, a system of indicators is formed at

the beginning. Our study used data for 2016, the year of the All-Russian Agricultural Census, the final results of which allowed us to obtain data that are not recorded by current statistics. Based on the developed system of indicators, a correlation matrix is built, its structure is investigated. The inclusion of a complex of factors in the econometric model is based on a theoretical analysis of the value of pair correlation coefficients, which evaluate the tightness of the relationship between the indicators. If there are a large number of factors (as in our case), they are replaced by components (integrated factors), which will subsequently form the basis for the classification of subjects. This, in turn, will allow more reasonably determine the direction of state regulation of income generation in order to achieve the target indicators of agricultural policy.

### 4. RESULTS

At the preliminary stage of econometric modeling, we studied the set of regions for homogeneity by the values of the indicators used. As a result, cities with federal status and the Chukotka Autonomous Okrug were excluded from the study population due to the practical lack of agricultural production. Moreover, the degree of differentiation of regions in terms of the share of GVA of agriculture in GRP, estimated using a system of indicators, is characterized as strong. Thus, the minimum value of the share of GVA of agriculture in the region's GRP is inherent in the Murmansk region (0.4%), and the largest - in the Republic of Kalmykia (31.6%). The ratio of the standard deviation to the average value, expressed as a percentage, is 69.9%. [Romantceva, Yu.N. 2018].

Since it is required to carry out correlation and regression analysis within qualitatively homogeneous groups, a grouping of regions was carried out by the specific gravity of the GVA of agriculture, hunting and forestry in GRP. This indicator characterizes the role of the industry in the formation of value added in the region's economy, which determines the level of agricultural development. The conducted grouping of subjects of the Russian Federation made it possible to distinguish three typical groups of regions: up to 4.5% (a group of regions with a share of GVA below the average for the economy), from 4.6 to 10% and more than 10.1%. The distinguished types of regions differ significantly in number, resource potential, level and conditions of development of agriculture and the economy as a whole. Also from Figure 2 it can be seen that the distribution of regions across the territory of the country is inherent in some regularity associated with the peculiarities of soil and climatic conditions of activity.



**Figure 2** Distribution of constituent entities of the Russian Federation by the share of GVA of agriculture, hunting and the provision of services in these areas in the regional GRP (%)

The first group of regions in which agriculture does not make a significant contribution to the formation of GRP are regions with the least favorable soil and climatic conditions of activity. The population of the group is concentrated in cities (80%) at a low density (3.5 people). The mining industry is well developed here, therefore, GRP per capita is higher than the national average by 1.5 times, and income by 20%. Despite the lowest production growth rates over the past 3 years, the GVA output per 100 ha of farmland is slightly lower than the average for the population (Table 1).

**Table 1** The characteristic of typical groups of regions according to the share of the GVA of agriculture, hunting and the provision of services in these areas in the regional GRP

Indicators	Typical groups			Average
	lowest	middle	highest	
Number of regions	18	34	25	77
The share of GVA of agriculture, hunting and forestry in GRP, %	2.3	7.1	15.4	6.9
GVA per 100 ha of agricultural land, thou. rubles	1544	1166	1877	1507
Actual normal productivity, point	77.2	98.1	102.5	86.9
Agricultural production index for 2014-2016, %	100.0	113.1	130.7	116.4
GRP per capita, thou. rubles	569	329	271	379
Population density, people per sq km	3.5	10.9	31.0	7.7
Cash income per capita, thou. rubles	33.7	25.6	25.1	27.8
The share of the rural population, %	20.0	28.0	38.0	29.0
Investments in fixed assets per 1000 rubles value of fixed assets, rubles	83.0	93.0	96.0	89.0
Subsidies per 1 ruble of GVA, rubles	0.056	0.053	0.048	0.051
The ratio of prices for agricultural and industrial products for 2010-2016, ±	-17.7	-14.8	-2.0	-12.0
The proportion of agricultural land in the total land area, %	2.9	21.7	68.8	13.4
The share of the value of fixed assets in agriculture in the total value of the funds, %	1.2	3.9	8.0	3.4
Agricultural production per capita, thou. rubles	19.8	37.1	70.9	42.8
Growth rates of cattle stock for 2006-2016, %	76.0	80.9	96.9	86.2
Fixed assets per 100 hectares of agricultural land, thou. rubles	125.7	77.6	116.3	100.0

Source: calculated by the authors according to FSSS (<http://www.gks.ru/>)

The most numerous second group with an average GVA of 7.1%, which includes regions mainly of the Central and Volga Federal Districts, as well as the southern regions of the Siberian and Far Eastern Districts, is characterized by high actual normal productivity (98.1 points). While the growth rate of gross output here is higher than in the regions of the first group.

Regions of the third group are represented by 25 regions of the South, North Caucasus and southern regions of the Central Federal District, which affected the population density, which is 4 times higher than the average, GRP per capita (lower by 28.5%), income. Due to favorable soil and climatic features, expressed by indicators of actual normal productivity, which is higher by 18% than the

population average, GVA per 100 ha of farmland exceeds the average Russian by a quarter.

The most favorable conditions for agricultural production are observed in the regions of the third group, in which investment in fixed assets per 1000 rubles of the value of fixed assets is higher by 7.9%, and the indicator for the first group - by 15.7%, but it is still very insignificant to replace massively obsolete funds with new, high-productive. While price disparity was present in the regions of the third group, the difference in growth rates did not exceed 2%, then in the regions of the second, and especially the first group, the parity of relations was greatly violated. Moreover, the situation in the first group can hardly be corrected by a higher level of subsidies per ruble of GVA.

As a result of the operating conditions, the regions of the third group produce most of the gross output and account for more than half of all revenues, which affects the volume of production per capita (70.9 thousand rubles). Regions provide not only for their food needs. Agriculture of the second typical group, playing a significant role in the formation of the GRP, also has the necessary resources for this. So, the share of agricultural land in the total area is 21.7% with a level of plowing 47.7%. The increase in the number of pigs, sheep, and goats is not as significant as in the third group, and there is a significant decrease in cattle (almost 20%). In the first typical group, agriculture is practically undeveloped: the share of agricultural land is less than 3%, the share of fixed assets is insignificant, but the indicator of the value of fixed assets per 100 hectares of agricultural land is, on the contrary, the highest in the population, with their severe deterioration

The obtained typical groups differ significantly, therefore, at the next stage of the study, it is necessary to identify and evaluate the influence of factors on the contribution of agriculture to the formation of the GRP, which will form the basis for substantiating the directions of state policy in relation to each group of regions.

The study of the data of the constructed grouping made it possible to select a set of indicators that affect the income of agricultural producers and characterize the conditions for the functioning of the sector in the regions, while the relationship between the result indicator and factors was significant. For further modeling, a wide system of indicators was considered, and indicators were combined into groups:

- Indicators of economic development;
- Indicators of food demand;
- Indicators of conditions and development level of agriculture;

Indicators of the level of intensification and efficiency of agricultural production;

Indicators of the level of government support.

However, the inclusion in the model of the whole spectrum of existing factors is impossible, since this not only does not reduce the residual variance, but also leads to statistical insignificance of the regression parameters by the t-criterion.

Before constructing a regression model for each typical group, a mathematical and statistical analysis of the relationship between variables was preliminarily carried out on the basis of a matrix of pair correlation coefficients. At this stage, some of the indicators with a low correlation coefficient were excluded from consideration. The analysis of the pair correlation matrix also showed the presence of multicollinearity of factors. The desire to take into account the influence of the maximum possible number of factors on the result indicator for a relatively small set of regions of each group led to the need for factor analysis, the purpose of which is to reduce the number of variables (data reduction) and obtain new generalized (integrated) variables (components) by combining variables, strongly correlated with each other.

To determine the optimal number of components in the model, the Kaiser criterion was used. Thus, in the first group, the use of three components is recommended, which account, respectively, for 44.5, 25.9, and 10.4% of the total variance, which together accounts for 80.9% of the variation.

The methodology for identifying the main components is based on the rotation of the factor matrix, which allows one to obtain factor loads (Table 2), understood as correlation coefficients between variables and main components (factors).

**Table 2** The matrix of rotated components for the first typical group (Varimax rotation method with Kaiser normalization)

Indicators	Components (generalized factors)		
	1	2	3
Average cash income of the population per capita (per month; rubles)	0.113	0.202	<b>0.952</b>
Consumer spending on food, rubles per month	0.371	-0.009	<b>0.806</b>
The share of agriculture in the value of fixed assets, %	-0.514	-0.181	-0.641
The share of investment in fixed assets in agriculture, %	-0.610	0.041	-0.453
The share of agricultural land in the total land area, %	<b>-0.885</b>	0.064	-0.122
Investments in fixed assets per 100 hectares of agricultural land, thou. rubles	0.041	<b>0.839</b>	0.258
Subsidies per 1 ruble of GVA, rubles	0.575	0.301	0.425
Potential agricultural land productivity	<b>-0.907</b>	-0.179	-0.190
Agricultural products per 100 ha of agricultural land, thou. rubles	-0.101	<b>0.940</b>	0.254
Fixed assets per 100 hectares of agricultural land, thou. rubles	0.172	<b>0.906</b>	-0.156
GVA of agriculture per 100 ha of agricultural land, thou. rubles	0.137	<b>0.970</b>	0.055
Total variance	2.782	3.550	2.562
Share of total variance	0.253	0.323	0.233

According to the data obtained, it can be seen that the indicators involved in the formation of the first generalized factor (the principal component) reflect the impact of production factors external to the agricultural producer that characterize the conditions of agricultural production (the share of agricultural land in the total land area (correlation coefficient with the first component is 0.885), reflecting the result of the impact of the soil-climatic conditions of agriculture, and the potential productivity with normal moisture and the most complete use of thermal resources (-0.907)).

The formation of the second principal component is dominated by indicators of the intensity of agricultural production (the value of investments in fixed assets per 100 hectares of agricultural land, the value of fixed assets per 100 hectares of agricultural land), as well as its effectiveness (volume of production and GVA of agriculture per 100 hectares of agricultural land). Therefore, this component should be defined as a factor characterizing the internal level of development of agricultural producers.

The demand factor includes the average cash income of the population per capita and consumer spending on food (the correlation coefficient is 0.952 and 0.806, respectively).

For a quantitative assessment of the influence of the principal components on the level of agricultural development in the region for the first typical group  $\beta$ -coefficients were used. A standardized scale regression equation was constructed:

$t_y = \beta_1 \cdot t_{x_1} + \beta_2 \cdot t_{x_2} + \beta_3 \cdot t_{x_3}$ , where standardized integrated variables for each region are taken as variables:

$$t_y = \frac{y - \bar{y}}{\sigma_y}; t_{x_i} = \frac{x_i - \bar{x}_i}{\sigma_{x_i}}$$

As a result, the following equation was obtained in a standardized form:  $y = -0,44t_1 - 0,32t_2 - 0,52t_3$ .

So, the factors characterizing the demand of the population and the level of its well-being have the greatest impact on the level of development of agriculture in the region (3rd component). As a result of increasing their quantitative value by 1  $\sigma$ , the response variable decreases by 0.52  $\sigma$ , while the level of income of the population has the strongest influence. This is explained by the fact that in conditions of generally unfavorable soil and climatic conditions of agriculture, the share of the sector in the formation of GRP will increase due to the deterioration in the standard of living of the population and the transition to self-sufficiency in food. On the contrary, the presence of effective demand will contribute to the development of inter-regional food exchange and a decrease in agricultural production in this group of regions.

The soil-climatic factor is in the second place by the force of influence on the result indicator, the direction of its influence is also negative. In adverse conditions, agricultural production is rather a necessary measure. Although the population, as well as other forms of small business, produce agricultural products, but the efficiency

of their production is low. In more adverse conditions, the specialization of agriculture is changing in the direction of greenhouse vegetable production and livestock farming on an industrial basis (pig farms, poultry farms, etc.), where the level of efficiency is higher with a small occupied area of agricultural land.

The influence of the second generalized factor determining the level of intensification and production efficiency is not as strong as the components described above, but still significant. The negative direction of the connection between the component and the share of agricultural GVA in the region's GRP is explained by the fact that the specialization of the regions of the typical group is far from agricultural, therefore, even with an increase in efficiency, the share of the industry will remain insignificant due to the insignificant amount of agricultural value added.

The factor analysis carried out according to a similar scheme for the regions of the second typical group also made it possible to distinguish 3 components, which account for 74.6% of the total variance.

The first component characterizes the social conditions of the development of regions and includes the level of employment, the share of the rural population in the total population, the share of the population with incomes below the subsistence level, GVA of agriculture per employed as a characteristic of labor productivity. The second component mainly determines the level of development of the region, since it includes GRP per capita, population density, as well as the share of agricultural land in the total land area. The third component includes the value of the ratio of price index for agricultural and industrial products, which characterizes the relationship between the parity of agribusiness industries.

Multiple correlation and regression analysis carried out using the principal components as explanatory variables showed that the largest share of variance of the response variable is due to the influence of the third generalized indicator - the ratio of the price index for agricultural and industrial products. At the same time, the direction of communication is positive, that is, with the restoration of parity of relations between the agro-industrial sectors, the share of the agrarian sector in the formation of the GRP of region of the Russian Federation increases. The influence of other factors is not so significant.

Three components were also identified for the units of the third typical group. The first integrated factor explaining 32.6% of the total variance characterizes the level of regional development (GRP per capita, population density, cash income per capita), as well as the cadastral value of agricultural land (which is estimated based on the productivity of 1 ha of agricultural land).

The second integrated factor included indicators such as the share of agriculture, hunting and forestry in the value of fixed assets, agricultural production per capita, the growth rate of agricultural production in 2010-2016, the share of agricultural enterprises in production of the sector, investment in fixed assets per 1000 rubles value of fixed assets. All these indicators determine the level of

agricultural development and are characterized by a positive direction of relations. That is, the role of agriculture in the formation of value added increases with increasing growth rates of the industry and its investment appeal. In addition, an important factor determining the development of the industry is the increase in the share of agricultural enterprises in the production of the sector. This is not surprising in view of their more efficient use of resources and high rates of intensification of large-scale production.

The third identified component displays the demand for agricultural products. The negative direction of relation is due to the fact that the regions of the third group are most favorable for agricultural production in terms of soil and climatic conditions. Therefore, personal subsidiary plots, with low material and labor costs, produce a sufficient amount of agricultural products for their own consumption, which reduces solvent demand and the share of food expenditures in total costs.

## 5. CONCLUSIONS

Thus, the conducted econometric modeling, evaluating the influence of the main groups of factors on the share of the GVA of agriculture, hunting and forestry in the GRP, showed that the value of the result indicator for typical groups of regions depends on various factors.

## REFERENCES

- [1] A.P. Zinchenko, Studying the problems of economic reproduction in agriculture of Russia, *Studies on Russian Economic Development*, 28 (2) (2017) 140-146. DOI: 10.1134/S1075700717020137
- [2] I.G. Ushachev, Systemic problems of reproduction in the agricultural sector of Russia, *Economics of the agro-industrial complex*, 7 (2014) 143-149. [http://nbuv.gov.ua/UJRN/E\\_apk\\_2014\\_7\\_22](http://nbuv.gov.ua/UJRN/E_apk_2014_7_22)
- [3] Efstratios Loizou, Christos Karelakis, Konstantinos Galanopoulos, Konstadinos Mattas The role of agriculture as a development tool for a regional economy, *Agricultural Systems*, Volume 173, 2019, - Pages 482-490 DOI: 10.1016/j.agsy.2019.04.002
- [4] D. Spirkova, B. Stehlikova, M. Zubkova, M. Sevela, D. Stiglic, Evaluation of agriculture's economic role in EU countries, *Ekonomicky casopis*, 65(8) (2017) 763-779
- [5] A. Đokić, S. Jović, Evaluation of agriculture and industry effect on economic health by ANFIS approach, *Physica A: Statistical Mechanics and its Applications*, 479(1) (2017) 396-399 DOI: 10.1016/j.physa.2017.03.022

In regions with unfavorable agricultural conditions (the first typical group), the agrarian sector is not developed and exists as an aid to households. At the same time, production is inefficient even if additional investments are made. Nevertheless, given the low population density, the industry needs state support as the most important area of rural development.

For the largest group of regions (the second typical group), producing a significant amount of agricultural products, government support of the sector is required for social stability; and to increase the share of the sector in the formation of economy income, it is necessary to pursue a policy of reducing the disparity in prices between agricultural and industrial products.

The regions of the third group, according to the analysis of variance, are the locomotive of the development of agriculture in Russia, while occupying an important place in the formation of income of the constituent entities of the Russian Federation, as well as ensuring food security. These regions make efficient use of available resources and are attractive to investors. The role of state agrarian policy is to maintain investment appeal.

Since the STATISTICA system is currently one of the most common statistical analysis software products, the proposed methodology, based on the application of the methods laid down in the program, can be used to analyze any set of units and a set of relevant indicators.

- [6] P. Bórawski, M. Grzybowska-Brzezińska, J.W. Dunn, The evaluation of efficiency of Polish agriculture, *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 63 (1) (2015) 175-183.
- [7] L.B. Iantovics, C. Rota, F. Morar, Survey on establishing the optimal number of factors in exploratory factor analysis applied to data mining, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9 (2) (2019)
- [8] J. Dien, D.J. Beal, P. Berg, Optimizing principal components analysis of event-related potentials: Matrix type, factor loading weighting, extraction, and rotations, *Clinical Neurophysiology*, 116 (8) (2005) 1808-1825.
- [9] G. Liu, Y. Liu, C. Zhang, Factor allocation, economic growth and unbalanced regional development in China (2018) 2439-2463.
- [10] Li Bin, Li Tuo, Yu Man, Chen Bin, Can equalization of public services narrow the regional disparities in China? A spatial econometrics approach, *China Economic Review*, 44 (2017) 67-78.

[11] O. Lemishko, Modelling of endogenous factors impacting the efficiency of the aggregate capital in Ukraine's agriculture, *Economic Annals-XXI*, (3-4) (2018) 10-14

[12] G.P. Herrera, R. Lourival, R.B. da Costa, D.R.F. Mendes, T.B.S. Moreira, U.G.P. de Abreu, M. Constantino, Econometric analysis of income, productivity and diversification among smallholders in Brazil, *Land Use Policy*, 76 (2018) 455-459  
DOI: 10.1016/j.landusepol.2018.02.025

[13] Yu. Romantseva, D. Galyautdinova, The role of agriculture research in formation of income in regions *AIC: economics, management*, 9 (2018) 22-31.