The Journal of Social Sciences Research
ISSN(e): 2411-9458, ISSN(p): 2413-6670
Special Issue. 5, pp: 404-410, 2018
URL: https://arpgweb.com/journal/journal/7/special_issue
DOI: https://doi.org/10.32861/jssr.spi5.404.410
Academic Research Publishing Group
Original Research Open Access

Modeling and Forecasting of Socio-Economic Development of the Region

Elena A. Stryabkova*
Belgorod State University, Russia, Pobedy Street, Belgorod, 308000, Russia

Anastasia S. Glotova
Belgorod State University, Russia, Pobedy Street, Belgorod, 308000, Russia

Irina N. Titova
Belgorod State University, Russia, Pobedy Street, Belgorod, 308000, Russia

Julia V. Lyshchikova
Belgorod State University, Russia, Pobedy Street, Belgorod, 308000, Russia

Irina V. Chistnikova
Belgorod State University, Russia, Pobedy Street, Belgorod, 308000, Russia

Abstract
The article is devoted to the modeling and forecasting of socio-economic development of the region. The dependence of GRP per capita of the Belgorod region on the average annual number of employed in the economy, the consolidated budget revenues, the volume of innovative works and services, the consumer price index, the industrial production index, the balanced financial result, exports was established. The analysis of the matrix of pair correlation coefficients of the selected indicators allowed to choose as the most significant explanatory variables the consolidated budget revenues and the average annual number of employees in the economy. The models of socio-economic development of the region were built. The quality of the models was evaluated. It was revealed that the most accurate is the power regression model. The forecast of further changes in GRP per capita was built on the basis of the retrospective analysis data. The method of extrapolation based on the construction of trend models for each explanatory variable was used to carry out the forecast.
Keywords: Regional economy; Modeling; Forecasting; Gross regional product (GRP); Correlation and regression analysis.

1. Introduction
The region is an agglomeration of production forces, which nature is determined by natural-geographical and other features. It acts as a central element of economic development throughout the state. Each individual region contributes to the development of the economy of the whole country. In this regard, the problems of socio-economic development of the regions do not cease to be relevant. The study of the factors, which determine the level and dynamics of indicators of socio-economic development of the region, becomes extremely important, as it allows to determine more qualitatively the possibilities of promoting the development of regional potential.

Among the many indicators that characterize the level of socio-economic development of the region, the most important is the gross regional product. The study of GRP and its components along with the indicator of GRP per capita in relation to other socio-economic indicators of the region allows to understand the mechanism of regional economic growth. The importance of identifying the main parameters that affect the changes in the gross regional product is also due to the fact that this indicator is used by the Ministry of Finance of the Russian Federation for the distribution of the financial support fund of the territories, and is taken into account in forecasting of regional development.

2. Methods and Materials of the Research
The choice of a model for forecasting the socio-economic development of the region is always associated with the search for the most qualitative model and determining the influencing factors. The most scientists commonly used is a fairly simple and repeatedly tested model of the Solow production function. Thus, S. V. Baranov and T. P. Skufina make forecasts based on the use of the classical approach to the modeling of production processes in the region with the help of production functions.

Baranov and Skufina (2007) at the same time, the developers of the model faced the problem of a short information base. To overcome these limitations, the developers of the model proposed to evaluate the parameters of the production function not by time series, but by regional ones. The development of the classical version of the Solow model was obtained in the works of G. V. Bakusheva, who modified the equation and adapted it to predict GRP without taking into account the influence of powerful external factors (Bystrov *et al.*, 2017);

*Corresponding Author

V. S. Mkhitaryan and O. A. Khokhlova in their work use the Box-Jenkins method, which allows to build linear statistical models with a high degree of accuracy describing the behavior of time series by combining the processes of auto regression, integration processes, moving average processes. Models rely only on the information contained in the history of the predicted series, which limits the capabilities of the algorithm.

Very often the forecasting of the region development is based on the use of regression equations. Professor M. D. Mamedsupiyev developed a model for prediction of GRP with a high degree of accuracy of the approximation of accounting data. The model of M. D. Mamedsupiyev is based on the number of factors, quite simple to use and allows you to carry out analytical and predictive calculations quickly upon receipt of reporting information (Mamedsupiyev, 2008).

In research of A. N. Petrova in the construction of linear econometric models, a single-factor regression model of the GRP index is proposed, in which the value of the index of the physical volume of Russia's GDP was used as a predictor (Petrov, 2011).

In researches of R. M. Nizhegorodtsev and A. N. Petukhov regression factor models of forecasting GRP of different subjects of the Russian Federation were developed (Nizhegorodtsev and Petukhov, 2011). Among the main factors which influence the production of GRP, the authors highlighted investments in fixed capital, labor costs and research and development costs.

E. V. Zarova and R. A. Prozhivin propose the concept of a balanced scorecard based on regional statistics to solve the problem (Zarova and Protivin, 2008). This approach requires solving problems of developing and testing methods of statistical analysis and econometric modeling of patterns of socio-economic development of the region. Work on the development of a balanced scorecard is reduced to a multi-level econometric modeling of the region. Forecasting of GRP on the basis of econometric models of is also used in the works of O.

V. Skotarenko, where multi-factor and single-factor models based on GRP are used to determine the strategic development of regions (Skotarenko, 2013).

In the work of G. N. Okhlopkov the necessity of application of time series extrapolation for regional forecasting of GRP is shown, however there is a problem of reliability of forecasts for the long-term period when using this method.

V. A. Borisov and E. I. Tsaregorodtsev suggest using exponential smoothing methods to predict GRP, but they have an error of more than 10% (Borisov and Tsaregorodtsev, 2014).

Finally, it should be noted that the input-output method can be used to predict the volume of gross regional product (Chistova, 2013; Litvinenko, 2017). The main obvious problem of using this method is the lack of regional input-output tables, as well as the limitations of its use in open systems.

In this paper, the modeling of statistical relationships of a number of regional indicators and their subsequent economic and statistical evaluation is performed using the methods of correlation and regression analysis (Kupriyanov *et al.*, 2018).

In this study, based on Collections of Federal State Statistics Service (Rosstat) ("Statistical year-book. Belgorod region. 2017", "The Belgorod region in figures. Release of 2017", "Regions of Russia. Socio-economic indexes. 2017" have been used) the dependence of GRP per capita of the Belgorod region on the average number of employees in the economy, consolidated budget revenues, the volume of innovative works and services, consumer price index, industrial production index, net financial result, exports is established. Mathematical modeling will confirm or disprove this dependence and further predict the level of GRP.

The list of explanatory variables includes the most important macroeconomic indicators, which give a clear picture of the situation, both at the federal and regional levels. On the basis of official statistics among the many factors, in our view, the change of the index of GRP per capita is affected by the following predictors:

X_1 - volume of innovative goods, works, services, mln. rubles;

X_2 - industrial production index to the previous year, %;

X_3 - consolidated budget revenues, mln. rubles;

X_4 - balanced financial result of organizations, mln. rubles;

X_5 - export, mln. doll. USA;

X_6 - the average annual number of employed in the economy, thousand people;

X_7 - consumer price index, %.

To carry out econometric modeling, data for 2000-2016 were used. (Table 1).

Table-1. Socio-economic indicators of development of the Belgorod region for 2000-2016

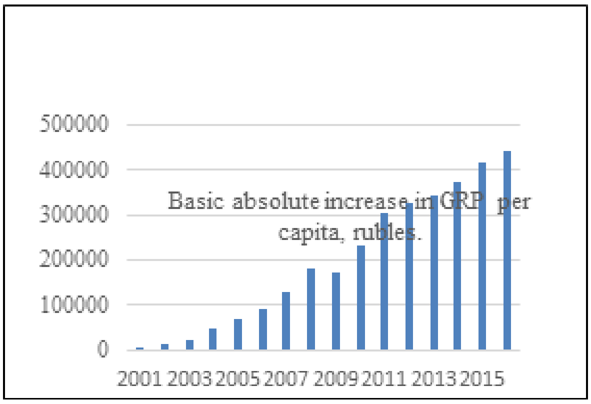
Year	GRP per capita, rubles, Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7
2000	27969.5	341.4	109.1	5396.8	4840	473.2	671.3	120.9
2001	33126.7	3266.7	110.1	7236.3	3687	719.4	677.5	119.2
2002	41327.4	1422.1	116	10173.4	-757	507.8	673.4	116.4
2003	50284	2376.3	106.2	10910	1818	713.6	668.3	112
2004	75650	1897.9	106.3	17269.2	25382	981.4	670.7	111.7
2005	95922	2206.9	112.9	24730.7	39917	1415.1	674.7	112.6
2006	119673.2	2052.7	113.4	29485	39807	1593.3	677.4	109.3
2007	156302.2	13377.9	111.8	41516.5	65587	2316.4	678.7	112.3
2008	208694.7	32978.9	112.6	56852.7	72595	3784.1	679.9	113.6
2009	199229.1	10437.5	104.2	55738.2	27967	1761.9	681	108.9
2010	260015.6	9391.6	113.9	63512.8	77244	2704.1	693.5	109
2011	331010	15457.4	108.7	81394.9	130561	3840.1	698.1	105.5
2012	354570.6	21683.4	105.6	78056.2	112814	3632.8	700.1	106.2
2013	368874.8	21246.5	102.1	77347.2	85118	3412.5	700.2	106.2
2014	400820.8	23098.3	102.5	77580.6	178309	3176	699.1	110.5
2015	443086.2	29348.1	105.5	81081.9	57605	2426.1	754	111.4
2016	470874	56411.5	106.2	82121.6	216056	2184.5	756.8	104.4

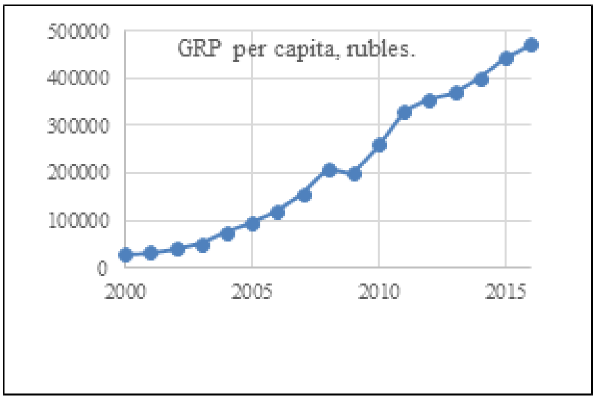
Gross regional product per capita as an indicator, which is an integral part of the sustainable development of the region and reflects its economic condition, depends on the socio-economic stability of the country, which determines the practical significance of the assessment of the selected factors.

3. Results of the Research

The value of the gross regional product per capita in the region has a stable positive trend, as evidenced by its dynamics and absolute increases (Figure 1). For the period 2000-2016, the average growth rate of GRP per capita was 120%, the average annual growth rate of 20%.

Figure-1. Dynamics of gross regional product per capita in the Belgorod region for 2000-2016





To select the most significant explanatory variables X_i, to explain the dependent variable Y, a matrix of pair correlation coefficients is compiled, taking into account the following conditions:

- the corelation between the dependent variable and the explanatory variable should be higher than the interfactor corelation, and the explanatory variables with r_{yx_i}<0.85 are excluded from the model;
- the corelation between the explanatory variables should be no more than 0.85. If the matrix has $r_{x_i x_j}$>0.85, there is multicollinearity in the multiple

regression models;

- - in the presence of multicollinearity, variables with a large correlation coefficient r_{yx_i} are selected.
-

Calculation of pair correlation coefficients was performed using Microsoft Excel and presented in the matrix (Table 2)

Table-2. Matrix of pair correlation coefficients

	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y	1							
X_1	0.83072	1						
X_2	-0.54105	-0.36992	1					
X_3	0.96746	0.76838	-0.47491	1				
X_4	0.84973	0.81029	-0.38504	0.81590	1			
X_5	0.76971	0.60240	-0.28164	0.87945	0.69356	1		
X_6	0.85947	0.82818	-0.39575	0.74029	0.67591	0.41142	1	
X_7	-0.74374	-0.54449	0.40967	-0.78971	-0.69868	-0.66703	-0.51921	1

The analysis of the matrix of pair correlation coefficients of the selected indicators in Table 2, allows us to draw the following conclusions:

1) Variables X_1, X_2, X_5, X_7, are excluded from consideration, since they have an insignificant impact on the effective characteristic, r_{yx_i}<0.85.
2) The value of pair correlation coefficients indicates a very close relationship between per capita GRP and consolidated budget revenues. Also, the average annual number of employed in the economy has a stimulating effect on the growth of the productive feature in the current and future periods. The balanced financial result of organizations is an integrated indicator of the activities of economic units, including both industrial production and services, and agriculture.
3) Note that it is necessary to exclude the balanced financial result from the model, as there is multicollinearity with the consolidated budget revenues. This relationship between these variables is understandable. The formation of the revenue part of the consolidated budget is carried out by deductions of organizations and enterprises of income tax. A balanced financial result represents the amount of profit (loss) of enterprises and organizations from the sale of goods, products (works, services), fixed assets, other property of organizations and income from non-sales operations,

4) reduced by the amount of expenses on these operations. Both of these variables are changing in one direction: the increase in the revenues of the consolidated budget is due to the growth of the balanced financial result. Therefore, to improve the model, it is necessary to exclude the variable whose pair correlation coefficient is less important with the resultant variable.
5) X_3 - consolidated budget revenues and X_6 - the average annual number of employees in the economy should be chosen as the main indicators.

The calculations were performed in Microsoft Excel; the results of calculations are presented in Table 3.

Table-3. The results of the regression analysis

Multiple R			**0.99065383**			
R Square			0.98139501			
Ajusted R Square			0.97873715			
Standard Error			22543.4583			
Observations			17			
	df	*SS*	*MS*	*F*	*Significance F*	
Regression	2	3.75304E+11	1.8765E+11	369.2430601	7.71626E-13	
Residual	14	7114905198	508207514			
Total	16	3.82419E+11				
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-1242816.8	209477.18	-5.932946	3.65363E-05	-1692100.655	-793532.92
X Variable 1	3.75911637	0.278162809	13.5140869	2.00653E-09	3.162516485	4.35571626
X Variable 2	1850.87705	316.6027145	5.84605552	4.24799E-05	1171.831759	2529.92233

We've got a two-factor model $Y = -1242816.6 + 3.8X_3 + 1850.9X_6$

Calculations of the coefficient of determination, which is $R^2 = 0.981$. This coefficient shows the proportion of GRP dispersion per capita due to the influence of the factors included in the equation. Accordingly, 98.1% of changes in GRP per capita are determined by changes in the explanatory variables, i.e. the average annual number of employees, consolidated budget revenues.

We check the quality of our model and the coefficients of the linear regression equation.

The significance of the regression equation is verified on the basis of Fisher's F-test. The value of the Fisher's F-test is presented in Table 3.

The table value of the F-test for the confidence probability $\alpha = 0.95$ and the number of degrees of freedom equal to $\nu = n - k = 17 - 2 = 15$ is 2.13, and it's less than calculated. Therefore, the regression equation should be considered as significant, that is, it can be used for further analysis.

Evaluation of the significance of the coefficients of the model based on the results of the Excel report is carried out in three ways:

1. Observed values of t-statistics for all coefficients of the equation (Table 3) more than the critical (tabular) value of t-statistics 2.11 Thus, the obtained values of correlation coefficients are significant.

2. Let's check the P-values for each coefficient of the regression equation:

$$P - value\ (\alpha 1) = 3{,}7 * 10^{-5} < 0.05$$
$$P - value\ (\alpha 2) = 2 * 10^{-9} < 0.05$$
$$P - value\ (\alpha 1) = 4{,}2 * 10^{-5} < 0.05$$

Therefore, all coefficients are statistically significant.

3. The lower and upper 95% of the boundaries of the confidence interval have the same signs (Table 3), which once again proves the importance of the coefficients of the regression equation.

The test of independence (lack of autocorrelation) will carry out using the criterion Durbin-Watson (Table 4) according to the formula:

$$DW = \frac{\sum_{t=2}^{n}[\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1}]^2}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2} = \frac{9041143131}{7114905198} = 1.27$$

Table-4. Independence (no autocorrelation) results based on Durbin-Watson test

№	$\hat{\varepsilon}_t$	$\hat{\varepsilon}_t^2$	$(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2$
1	8005.327702	64085271.62	
2	-5227.804554	27329940.45	175115789.3
3	-479.409366	229833.3403	22547256.86
4	15147.69845	229452768.3	244206498.6
5	12166.62069	148026659.1	8886824.585
6	-3013.534315	9081389.07	230437106.1
7	-2131.669316	4544014.074	777685.8765
8	-13136.61813	172570735.9	121108898.4
9	-20615.73112	425008369.5	55937131.09
10	-27927.76067	779959816	53465776.17
11	-19502.84991	380361154.5	70979121.36
12	-24243.37923	587741436.4	22472618.25
13	8166.028517	66684021.74	1050369710
14	24950.35432	622520180.8	281713592.7
15	58054.94131	3370376211	1095913680
16	-14454.60268	208935538.8	5257633970
17	4242.388293	17997858.43	349577471.6
Total		7114905198	9041143131

Since the calculated criterion falls into the uncertainty zone (d_1=1.02<DW<d_2=1.54), when there is no reason to either accept or disprove the hypothesis of the existence of autocorrelation, another criterion should be applied.

Let's use the Breusch-Godfrey criterion.The Breusch-Godfrey autocorrelation test is a procedure used in econometrics to check autocorrelation of an arbitrary order in random errors of regression models. If the value of the statistics exceeds the critical value, autocorrelation is considered significant, otherwise it is insignificant.

Since the observed value (T=1.49) is less than
critical (T=1.73), it is possible to conclude about the absence of autocorrelation of the first order.

Let's characterize the accuracy of the model using the average relative approximation error by the formula $A = \frac{1}{\bar{y}}\sqrt{\frac{\sum(y-\hat{y})^2}{n}} \cdot 100\%$

Acceptable accuracy of the model in solving practical problems, as a rule, is determined on the basis of considerations of economic feasibility, taking into account the specific situation. The criterion according to which the accuracy is considered satisfactory if the average relative error is less than 15% is widely applied. If A is less than 5%, it is said that the model has a high accuracy; the approximation error within 5-7% indicates a good selection of the model to the original data; 8-10% (up to 15% is allowed) means that the model is satisfactory and can be used for forecasting. It is not recommended to use for analysis and forecasting models with unsatisfactory accuracy when A is greater than 15%.

The calculated value of the average relative approximation error for our model A=9.6%, gives us the opportunity to conclude that this regression equation is satisfactory, does not exceed 15%. However, in order to improve the accuracy of the model, it is necessary to search for another nonlinear regression formula, and to build a power function using Microsoft Excel.

The new model looks like this $Y = 3.7 \cdot 10^{-11} \cdot X_3^{0.9} \cdot X_6^{4.1}$.

The results of the comparative analysis of the quality of the developed models of the Belgorod region are presented in Table 5.

Table-5. The comparative analysis of the quality of the developed models

Model	Fisher's F-test	R2	The average relative error of the approximation
$Y = -1242816.6 + 3.8X_3 + 1850.9X_6$	369.2	0.981	9.6
$Y = 3.7 \cdot 10^{-11} \cdot X_3^{0.9} \cdot X_6^{4.1}$	958.8	0.993	6.7

The data of Table 5 allow us to recommend a model of power function, which indicates a good selection of the model to the original data. According to the chosen model, the greatest impact on the growth of the dependent variable - GRP per capita has an average annual number of employed in the economy, and to achieve the growth rate of the effective indicator of 1%, it is necessary to increase the number of employed in the economy by 4.1%, and the consolidated budget revenues by 0.9% per year.

1) Convinced of the adequacy of the model, we proceed to the implementation of the forecast for the next two years. For this purpose, each factor in the model
2) was studied and a trend was selected, on the basis of which separate forecast values were obtained. $x_{t2} = 0.05t^3 - 0.76t^2 + 4.21t + 667.03$ $(R^2 = 0.903)$

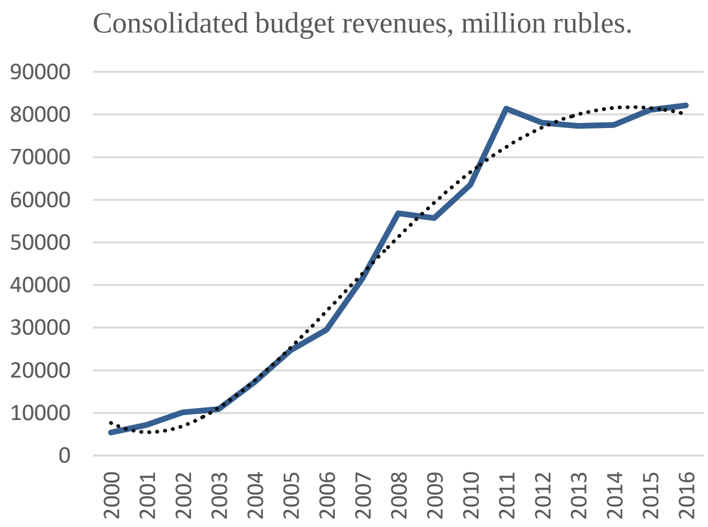
The criterion for choosing a trend is the coefficient of determination (R^2): the closer this indicator is to one, the more preferable is the choice of the corresponding trend.

Thus, based on the data for 2000-2016, the following results were obtained:

3) consolidated budget revenues (X1) have a trend line of the fourth degree polynomial (Figure 2):

(Figure 2):$x_{t1} = 2.85t^4 - 162.97t^3 + 2715.70t^2 - 9212.6t + 14339$ $(R^2 = 0.986)$

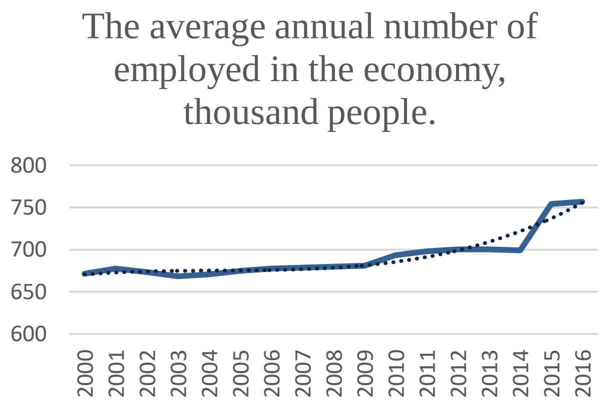
Figure-2. Dynamics of revenues of the consolidated budget of the Belgorod region for 2000-2016



the trend line of the second factor - the average annual number of employed in the economy (X2) is represented by a polynomial of the third degree (Figure 3):

$$x_{t2} = 0.05t^3 - 0.76t^2 + 4.21t + 667.03 \ (R^2 = 0.903)$$

Figure-3. Dynamics of the average annual number of employed in the economy of the Belgorod region for 2000-2016.



Using the trend equations, the predicted values of the model factors and GRP per capita for the next two years are determined (Table 6).

Table-6. The predicted values of the factors and GDP per capita

Year	x_{t1}	x_{t2}	Y_t
2017	77307.5216	777.3628	506321.942
2018	73479.4336	802.8819	551453.473

4. Discussion

As we can see from Table 6, it is planned to reduce consolidated budget revenues by 5.9% in 2017 and by 10.5% in 2018 compared to 2016. It can be

explained by the decrease in gratuitous revenues, namely inter-budgetary subsidies, as well as the negative dynamics show revenues from the sale of tangible and intangible assets, and with them the regional property tax of organizations is presented in the consolidated budget of the Belgorod region with a negative increase.

Despite this fact, in 2017 and 2018, an increase in GRP per capita is projected, which indicates the sustainable development of the Belgorod region. In terms of gross regional product per capita, the region is on the 3rd place in the Central Federal district, second only to Moscow and the Moscow region, and on the 19th place among the regions of the Russian Federation.

On the basis of the report "The results of The Department of Economic Development of the Belgorod region for 2017 and the tasks for 2018" GRP per capita in 2017 is 498900 rubles, which distinguishes it from the projected value of 1.49%, and this indicates a good quality of the model, as the approximation error is less than 5%.

5. Conclusion

Thus, in the course of the study, models of socio-economic development of the region were built, the quality of the models was evaluated and it was revealed that the most accurate is the power regression model. Using the data of the retrospective analysis, the forecast of further changes in GRP per capita was built. The method of extrapolation based on the construction of trend models for each explanatory variable was used to carry out the forecast.

Summing up, it should be noted that the management of socio-economic development of regional systems requires an understanding of the processes, phenomena and relationships, in-depth analysis of the current situation, trends in the various parameters and the consequences of decisions. The use of modern methods and technologies of economic and mathematical modeling is an effective way to improve approaches to strategic planning in the management of socio-economic development of the region, because it allows you to take into account and assess the impact of the impact of multiple intersecting impacts.

References

Baranov, S. V. and Skufina, T. P. (2007). Modeling of gross regional product production in the North and non-Northern part of Russia, Questions of statistics. 2: 57-62.

Borisov, V. A. and Tsaregorodtsev, E. I. (2014). The accuracy of forecasting the gross regional product of the Republic of Mari El, UEKS, 8. Available: http://uecs.ru/uecs68-682014/item/3017-2014-08-19-06-54-02

Bystrov, V. V., Malygina, S. N. and Khaliullina, D. N. (2017). Simulation in tasks of regional security management. *Belgorod State University Scientific bulletin, Economics. Computer Science,* 23(272): 92-104.

Chistova, M. V. (2013). Forecasting the volume of gross regional product by the method of, Input-output, Domestic and foreign experience. *Bulletin of Adyghe state University. Series 5, Economy,* 3(127): 148-57.

Kupriyanov, S. V., Stryabkova, E. A. and Gerasimova, N. A. (2018). Features of economic development of border regions. *Belgorod State University Scientific Bulletin, Economics. Computer Science,* 1(45): 5-14.

Litvinenko, I. L. (2017). To analyze the organizational-economic mechanism of management of regional innovative system. *Belgorod State University Scientific Bulletin, Economics. Computer Science,* 16(265): 40-49.

Mamedsupiyev, M. D., 2008. "Multifactor assessment model gross regionalnorth East Kazakhstan region, Our common home - Altai." In *International partnership, Science, Economy, Education, Culture, Tourism, International. science.- practice. Conf. Ust-Kamenogorsk: publishing house of wkro,.*

Nizhegorodtsev, R. M. and Petukhov, A. N. (2011). *The forecasting of gross regional product on the basis of three-factor regression model in increments.* Vestnik yurgtu NPI,: Volga Federal district. 3: 90-100.

Petrov, A. N. (2011). Set of econometric models to predict the gross regional product index of the Ivanovo region, Proceedings of higher educational institutions. *Economics, Finance and Production Management,* 1: 93-101.

Skotarenko, O. V. (2013). Forecasting of GRP in regions on the basis of econometric models. *International Research Journal,* 8(15): 64-68.

Zarova, E. V. and Protivin, R. A. (2008). Balanced system of indicators of regional development: the rationale and economic modelling. *Problems of Statistics,* 8: 59-66.