

On the Methodology of Checking Integral Estimates of Socio-Economic Objects*



Dmitrii A. ALFER'EV

Vologda Research Center, Russian Academy of Sciences
Vologda, Russian Federation
Peter the Great St. Petersburg Polytechnic University
Saint Petersburg, Russian Federation
e-mail: alferov_1991@mail.ru
ORCID: 0000-0003-3511-7228; ResearcherID: I-8333-2016



Aleksandr E. KREMIN

Vologda Research Center, Russian Academy of Sciences
Vologda Institute of Law and Economics, Federal Penitentiary Service of Russia
Vologda, Russian Federation
e-mail: akremin@inbox.ru
ORCID: 0000-0001-9405-2255; ResearcherID: I-8188-2016



Dmitrii G. RODIONOV

Peter the Great St. Petersburg Polytechnic University
Saint Petersburg, Russian Federation
e-mail: drodionov@spbstu.ru
ORCID: 0000-0002-1254-0464; ResearcherID: A-9693-2017



Darya S. VELICHENKOVA

Peter the Great St. Petersburg Polytechnic University
Saint Petersburg, Russian Federation
e-mail: velichenkova_ds@spbstu.ru
ResearcherID: ABG-8363-2020

* The article was financially supported by the RF President grant (MK-2249.2020.6).

For citation: Alfer'ev D.A., Kremin A.E., Rodionov D.G., Velichenkova D.S. On the methodology of checking integral estimates of socio-economic objects. *Economic and Social Changes: Facts, Trends, Forecast*, 2021, vol. 14, no. 6, pp. 86–106. DOI: 10.15838/esc.2021.6.78.5

Abstract. A reliable and high-quality assessment of scientific, technological, and innovative development of territories helps to define socio-economic conditions and forecast economic growth dynamics of a given subject. The usage of integral indicators is among the most popular approaches toward assessing science, innovative activity, and other socio-economic objects. However, since a collective synthetic category is estimated, accuracy of this metric's characterization of an intangible subject is uncertain. In this regard, issues related to the development of methodology for checking aforementioned provisions are relevant. The purpose of the study is to define the reliability of artificially derived integral complex estimates that in turn describe various socio-economic processes and phenomena. Scientific novelty of the research is to develop an approach to determining the reliability of integral metrics based on mathematical statistical tools. We attempted to determine the quality of artificially derived integrated estimates that, according to their creators, characterize various manifestations of science, innovation, and technology. We applied corresponding methods (variance, correlation, and regression analysis) using the Innovation Development Index of RF constituent entities and assessment of territories' scientific and technological potential. The results obtained are of practical importance in relation to the evaluation practices of the scientific and innovative sphere carried out in the Russian economy. The theoretical significance of the study is characterized by the development of an approach that can be applied to other socio-economic objects. We conclude that integral assessments become an extremely subjective tool when applied to humanitarian areas. They can be used correctly if there is a confirmed connection with the indicators: integral assessments should influence them or have a strong response from them.

Key words: scientific and technological development, integral estimate, analysis of variance, correlation analysis, regression analysis, logit function, economic growth, innovation activity.

Introduction

Prediction of the economic growth, development, and social consequences is an important task for government agencies. Using this tool, they can anticipate the consequences of planned actions and correctly adjust their activities with policies to avoid serious socio-economic shocks or accelerate the onset of any positive events.

To assess a territory's economic and social well-being, the GDP indicator has been used for many years (the equivalent for lower-level economic entities is the GRP value) [1; 2]. It is intended to characterize the economic growth dynamics. When assessing business entities, the most developed territories are determined by this metric.

In accordance, the search for "levers" of regulation is an important scientific problem. In other words, the scientific community needs to identify the factors by changing which it is possible

to increase or decrease the value of GDP (GRP), and therefore influence economic growth or development.

Such key factors include scientific and technological (in older sources, "scientific and technical") progress and associated innovative activities of enterprises and organizations [3; 4; 5]. According to several authors [6–11], science and technology are the engine of socio-economic development. Therefore, the state should implement a policy of supporting science and the research sector to increase its competitiveness in the international arena, as well as to improve citizens' living conditions in a particular territory [12; 13; 14].

In this regard, various economic and mathematical models are being developed [15; 16] (R.M. Solow [8], S. Rebelo [9], K. Arrow [17], P.M. Romer [18], R. Lucas [19], D. Grossman and

E. Helpman [20], K. Freeman and B.-A. Lundvall [21], C. Griliches [22]). Depending on the indicators chosen and justified by scientists and researchers, they describe functional relationships between economic growth and any economic object (costs for science; amount of research and development works carried out; human capital; enterprises that carry out and implement innovative activities; educational institutions; dynamics of innovative ideas). These models have both a theoretical justification and some empirical implementation that, to a certain extent, is their approbation and confirmation of the formulated ideas.

Scientists also attempt to conduct a comprehensive assessment of scientific and technological changes in the economic environment. For this purpose, various kinds of “integral” estimates are developed and calculated. An example is the index of the scientific and technical potential of a region, the calculation method of which is published in the works of K.A. Zadumkin and I.A. Kondakov [23], or the assessment of the scientific and technological potential of the territories, presented in an article by a team of authors led by K.A. Gulin [24]. A prominent and fundamental work is the draft “Russian Regional Innovation Ranking”, developed and published from 2012 to 2019 by a team of authors of the Institute for Statistical Studies and Economics of Knowledge, part of the HSE¹. This methodology includes an assessment of 37 indicators and has a significant calculation period (7 years).

The mentioned works attempted to assess the cumulative impact of scientific, technological, and innovative development factors on the territories’ socio-economic level (Russia as a separate state, and districts or regions that are part of it, representing similar units of smaller size). As a result of the complexity of the estimates obtained,

¹ Russian Regional Innovation Ranking. Available at: <https://issek.hse.ru/rir>

their interpretation is a kind of conditional unit that characterizes a generalized process or phenomenon (in our case, scientific and technological potential, innovative development, etc.), but does not have an explicit quantitative interpretation that can be obtained during measurement procedures. As a result, the question about the adequacy and reliability of the processes, described by such integral assessments, arises. Their reliability can be confirmed only sometime later, which significantly reduces the significance and practical applicability of such models.

A person often encounters the need to evaluate objects that are characterized by heterogeneous parameters. Most often, such an assessment is carried out intuitively and, as a result, there is a negative result. The use of integral estimates is also associated with several problems, which can be characterized as follows² [25; 26; 27]:

- need to consider the weight and significance of each of the parameters included in the overall assessment;
- need to specify a way of translating qualitative assessments into quantitative ones;
- distribution of the assessed objects into the corresponding groups, characterized by the magnitude of the levels identified in the study;
- possibility of comparing estimates with the ones that will be obtained in the future (socio-economic indicators are often non-permanent and may become irrelevant over time, unlike physical quantities that are measured by objects of the natural sciences).

The problematic aspects of integral assessments described above are also indicated in several domestic scientific papers, which attempt to systematize advantages and disadvantages of such approach regarding the interpretation of socio-

² Minaker B.E., Bychovskyi M.V. Problems of integral assessments of technical systems. **TRIZ-Summit – 2006**. Saint Petersburg, October 13–14, 2006. Available at: <https://www.metodolog.ru/00821/00821.html>

economic objects. Thus, in the E.N. Volkova's article, a methodology for characterizing the socio-economic development of the region is formed and described on the basis of an integral assessment [28]. In the work of E.V. Klyushnikova and E.M. Shitova, the features of constructing integral estimates in accordance with the main stages of modeling are outlined: normalization, aggregation, weighing [29]. Similar studies are being conducted abroad. One of these works is the publication of a team of scientists led by M.-S. Saib [30], in which the authors use an integral indicator to assess the inequality of the population of territories in terms of conditions and factors affecting health.

Among the modern works, special attention should be paid to the publication of A.A. Sidorov [31]. He meticulously systematizes and describes the mathematical nature of the integral approach, thus continuing the work of the famous Russian econometrician S.A. Ayvazyan [32; 33; 34]. In the foreign environment, there are also studies related to the mathematical construction of an integral indicator. Thus, a team of authors led by P. Zhou [35] proposes a variant of aggregating an integral estimate based on the product of adjusted partial values of the included indicators.

Despite the wide variety of works on the use of the integral approach in relation to socio-economic objects, one point is omitted in them – do integral indicators correctly characterize what they are intended to describe? Although nearly all the studies reviewed stipulate that such assessments carry a fair share of indefinite subjectivism.

In this regard, we aim to determine the reliability of artificially derived integral complex estimates by means of mathematical and statistical methods, which, in turn, describe various socio-economic processes and phenomena.

“Reliability” of the estimates, indicated in the work, means their ability to explicitly describe the processes and phenomena that they should characterize.

To achieve this goal, we have solved several tasks in this work:

- defined and described the methods of mathematical statistics, with the help of which the search for the relationship between economic growth and the assessment of scientific and technological development is carried out;
- formed a sample of data for the calculations indicated in the work;
- carried out and described the results of calculations, based on which the relevant conclusions and recommendations were formulated.

Research methods

One-way analysis of variance

In mathematical statistics, analysis of variance is used to investigate the presence of the influence of qualitative factors on the values of a quantitative indicator. In our case, the resulting indicator will be $y:GRP$, and for the factor $x: x \in (I_1 || I_2), (I_1 || I_2) \supset S_1 \& S_2 \& S_3 \& S_4$, where I_1 is the Russian Regional Innovation Index, I_2 is comprehensive assessment of scientific and technological potential, and $\{S_n\}_{n=1}^4$ is their constituent sub-indexes.

Using the one-way analysis of variance, we attempt to define whether the difference y in the Russian Federation subjects, observed on $k = 4$ levels (socio-economic conditions, scientific and technical potential, innovation activity, quality of innovation policy in one case and research and development, personnel, technology, and innovation in the other) is statistically significant.

Algorithm of sample formation for the one-way analysis of variance

The algorithm for selecting elements of four different samples in accordance with levels I_1 will be formed based on the ratings of the subjects of the Russian Federation according to the sub-indexes of the Russian Regional Innovation Index. The distributions obtained by scientists and researchers of the ISSEK include four groups that can be characterized as follows (the groups are modeled

by the authors of the questioned study, but there is no explicit justification for the boundary values in them; a variant of a possible interpretation is presented below):

I group – cities of federal significance (Moscow, St. Petersburg) that have the best indicators for most statistical metrics characterizing socio-economic development. By their nature, their estimates are several times different from other similar objects (regional territorial units), which is why they are clearly out of the overall distribution picture and look like outliers. In this regard, while conducting the one-way analysis of variance, they will be excluded from our study, which will allow evaluating truly equivalent objects among themselves;

II group – regional territorial entities that have the best values for the assessed characteristics (except for objects included in group I), often exceeding the average national estimates. Such subjects can be regarded the ones that have the studied feature;

III group – the RF regions, which often have values according to S_n estimates that are smaller than the average Russian characteristics; these territories only approach the qualitative level of the studied features I , and therefore they cannot be considered the ones possessing the studied characteristics. It means they should be excluded from further calculations;

IV group – territorial units with the lowest values according to the considered characteristics; they can be characterized by a strong spread of estimates in the context of dynamics, instability of growth rates, and the absence of pronounced trends; in this regard, such objects will not be included in the studied sample.

The algorithm for selecting elements of four different samples in accordance with levels I_2 will be formed based on the ratings of the subjects of the Russian Federation by sub-indexes of the assessment of the territories' scientific and technological potential. The distributions, obtained by scientists

and researchers of the VolRC RAS, include five groups (levels: high, above average, average, below average, low), which can be characterized as equally distributed and scaled to a whole dozen ($I_2 \in [0; 10]$).

In accordance with the above-mentioned rule of inclusion of observations in the analysis of variance (by analogy with the rating of the subjects of the Russian Federation), we are interested in the “average” and “above average” level groups. In them, as in the previous version, we imply the presence of the considered integrated assessment, therefore, the observation data should be reflected in the change in the dynamics of GRP. The territories included in the high-level sample are unique single objects, and they represent outliers of some kind. The “below average” and “low” level groups include the bulk of observations and, in fact, are identical objects without any prominent features (in 2015, their shares in the sample groups were 93.75; 95; 95, and 98.75%, respectively, in accordance with the selected subindex).

This is the end of the sampling algorithm, and then the algorithm of the one-way analysis of variance continues.

After sampling the regions of the Russian Federation, we determine the number of objects, included in each of the k levels ($k = 4$) as the sum of the elements included in the considered set, and denote it by m_n , where n is the ordinal number of the level.

Then we determine the total number (m) of objects included in all the samples (*formula 1*):

$$m = \sum_{n=1}^k m_n . \quad (1)$$

Next, we calculate the average GRP value for each of the formed k groups (*formula 2*):

$$\bar{y}_n = \frac{1}{m_n} \sum_{i=1}^{m_n} y_{ni} , \quad (2)$$

where y_{ni} is the GRP value, corresponding to i -region in the sample n .

Next, we will find the average value of the resulting variable y for all available values included in all samples n (formula 3):

$$\bar{y} = \frac{1}{m} \sum_{n=1}^k \sum_{i=1}^{m_n} y_{ni} \leftrightarrow \frac{1}{m} \sum_{n=1}^k \bar{y}_n m_n. \quad (3)$$

The next stage is the search for the sum of squared deviations of the resulting estimates (y_{ni}) by samples from the common average (\bar{y}) (formula 4):

$$S^2 = \sum_{n=1}^k \sum_{i=1}^{m_n} (y_{ni} - \bar{y})^2. \quad (4)$$

Then, sum of squared deviations of averages by groups (\bar{y}_n) from the total average (\bar{y}) (formula 5):

$$S_F^2 = \sum_{n=1}^k (\bar{y}_n - \bar{y})^2 m_n. \quad (5)$$

Next, we calculate the residual sum of the squares of the deviations as the sum of the squares of the difference of the resulting values (y_{ni}) from the average values (\bar{y}), included in the same sample (formula 6):

$$S_{res}^2 = \sum_{n=1}^k \sum_{i=1}^{m_n} (y_{ni} - \bar{y}_n)^2. \quad (6)$$

To check S^2 (formula 4), the following equality can be used (formula 7):

$$S_F^2 + S_{res}^2 = S^2. \quad (7)$$

Let us calculate the factor variance (formula 8):

$$\sigma_F^2 = \frac{S_F^2}{k - 1}. \quad (8)$$

Calculate the residual variance (formula 9):

$$\sigma_{res}^2 = \frac{S_{res}^2}{m - k}. \quad (9)$$

Next, we find the value F_f according to the formula of private factor variance (σ_F^2) to residual (σ_{res}^2) (formula 10):

$$F_f = \frac{\sigma_F^2}{\sigma_{res}^2}. \quad (10)$$

Using the Fisher–Snedecor distribution, a given level of significance (α), and two degrees of freedom (formula 11 and 12), we can define the metric value F_k :

$$df1 = k - 1; \quad (11)$$

$$df2 = m - k. \quad (12)$$

To find F_k , classical Fisher–Snedecor distribution tables, which appear in reference books on mathematical statistics, can be used, but a scientist may encounter a several problems using them. These include the absence of necessary numerical values, which the study is based on. A similar situation associated with the choice of the level of reliability (in the reference literature, there are often only $\alpha: \alpha = 0.1 || 0.05 || 0.01$). Such problem is currently easy to solve using the capabilities of computational computer programs (for example, function “ F_k ” Python “Scipy” libraries).

As a result, it is necessary to compare the values obtained F_f & F_k (formula 13):

$$F_f > F_k \rightarrow A, \quad (13)$$

where A is the statement that the investigated qualitative features really have an impact on the value of the resulting indicator.

Correlation analysis algorithm

Using the correlation analysis, the presence of a statistical relationship between several random variables is determined. One of these tools is the Pearson correlation coefficient (r), which indicates the presence of a linear relationship. This indicator takes values from -1 to 1 inclusive and, in case of a negative value, indicates a decrease in the resulting estimate relative to an increase in the factorial one. In case of a positive value of the metric r , the situation is drastically opposite. The closer the absolute value r to one, the greater the linear relationship between the indicators is. It is worth noting that, at $r = 0$, there is no linear relationship, but the possibility of describing it through a

nonlinear function or the functional exist. The correlation coefficient between indicators, factors, or the resulting feature (r_{xy}) can be found as the ratio of covariance between them (cov_{xy}) and the product of the corresponding mean square deviations (σ_x & σ_y) (formula 14):

$$r_{xy} = \frac{cov_{xy}}{\sigma_x \sigma_y}, \tag{14}$$

where $cov_{xy} = M((x_n - \mu(x_n))(y_n - \mu(y_n))) \leftrightarrow M(x_n y_n) - \mu(y_n)\mu(x_n)$;

$M()$ is the unbiased estimation of the mathematical expectation of the sample;

$\mu()$ is the average value of the studied observations.

Algorithm for constructing the multiple linear regression model

The multiple linear regression model is a tool used in multivariate statistical analysis to describe the relationship of signs (causes) with any result or consequence. Its general form is represented by the analytical formula (formula 15):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \dots, \tag{15}$$

where $\{\beta_n\}_{n=1}^k$ is regression coefficients showing the degree of influence of the factor on the resulting feature;

β_0 is a free parameter of the model that allows the curve to be optimally positioned in space in such a way that the sum of squared deviations (OLS) is the smallest.

To calculate coefficients β , it is convenient to use the matrix search method (formula 16):

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_z \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1n} \\ 1 & x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{z1} & x_{z2} & \dots & x_{zn} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \rightarrow \tag{16}$$

$$\rightarrow \hat{\beta} = (X^T X)^{-1} X^T Y$$

After building the model itself, it is necessary to evaluate its accuracy and significance. The accuracy of the multiple linear regression model is determined using the coefficient of determination (R^2), which can be found by the following formula (formula 17):

$$R^2 = \frac{SS_f}{SS}, \tag{17}$$

where $SS_f = \sum_{j=1}^z (\hat{y}_i - \bar{y})^2$ is the explained sum of regression squares;

$SS = \sum_{j=1}^z (y_i - \bar{y})^2$ is the total sum of regression squares;

j is the ordinal number of the observation included in the generated model.

The closer this metric to one, the better the constructed model approximates the available empirical data.

To compare the accuracy of identical models that differ in the number of regressors, a different metric is used. It is because the coefficient of determination calculated by the method according to formula 17 will always be better with a larger number of parameters. To compare the quality of such regressions, we should use the adjusted R^2 (formula 18):

$$R_{adj}^2 = 1 - \frac{(1 - R^2) \cdot (z - 1)}{z - n - 1}. \tag{18}$$

At the final stage, the significance of the constructed model is evaluated. It is determined using Fisher's F -test. To do this, the required level of significance (α) is needed, and we calculate the following characteristics (formula 19):

$$SS_o = \sum_{j=1}^z (y_i - \hat{y}_i)^2 \leftrightarrow SS - SS_f - \text{is the residual sum of regression squares;}$$

$$(df1_p = n) \ \& \ (df2_p = z - (n + 1)) \quad \text{is (19)}$$

a number of degrees of freedom;

$$\left(MS_f = \frac{SS_f}{df1_p} \ \& \ MS_o = \frac{SS_o}{df2_p} \right) \rightarrow (F_k)_p = \frac{MS_f}{MS_o}.$$

Determination of $(F_k)_p$ happens similarly to the procedure of searching for (F_k) , however, in this case, instead of $df1$ (formula 11) and $df2$ (formula 12), we take accordingly $df1_p$ & $df2_p$ (formula 19). Then, similarly to formula 13, the new corresponding values obtained are compared.

Algorithm for determining the significance of multiple linear regression coefficients

In order to determine the significance of the coefficients of multiple linear regression, the Student's t -test is used. It allows answering the question: "Can the coefficients obtained in the model be interpreted?"

The observed values, obtained from the constructed statistical model, are calculated using the standard error of the parameters $(l_{\beta_0} \& l_{\beta_n}) \beta_0 \& \beta_n$ accordingly (formula 20):

$$t_{\beta_0} = \frac{\beta_0}{l_{\beta_0}} \& t_{\beta_n} = \frac{\beta_n}{l_{\beta_n}}; t = t_{\beta_0} || t_{\beta_n}. \quad (20)$$

Standard errors can be found from formula 16 using the matrix A (formula 21):

$$A = (X^T X)^{-1}. \quad (21)$$

This matrix is square and is determined by the size $(k + 1) \times (k + 1)$. Therefore, its diagonal element can be denoted as a_{nn} (formula 22):

$$l_{\beta_0}^2 = MS_0 a_{00}, \quad (22)$$

$$l_{\beta_n}^2 = MS_0 a_{nn}, n = \overline{1, k}, k = 4.$$

The values obtained in formula 21 are compared with the two-way critical point of the Student's distribution – $t_k(\alpha; z - n - 1)$. With $|t| > t_k$, the corresponding parameter of multiple linear regression is considered statistically significant, the null hypothesis of the form $H_0: \beta_0 || \beta_n = 0$ is rejected.

Algorithm for constructing logistic regression

Using this method, the potential probability of the occurrence of an event can be determined. In our case, the probability of an increase in the GRP dynamics depending on the cumulative annual change of k -factors

In general, the model looks like this (formula 23):

$$\hat{P} = p_0 + p_1 \cdot \Delta x_1 + p_2 \cdot \Delta x_2 + \dots + p_n \cdot \Delta x_n + \dots, \quad (23)$$

where \hat{P} is preliminary assessment of the probability of occurrence of a certain event;

$\{p_n\}_{n=0}^k$ is regression coefficients similar to those that characterize the model (formula 15);

$\{\Delta x_n\}_{n=1}^k$ is the values of the annual change of the corresponding factor.

The final transformation (\hat{P}') can be performed using the *sigmoid* function, and it looks like this (formula 24):

$$\hat{P}' = sigmoid(\hat{P}) = \frac{1}{(1 + e^{-\hat{P}})}. \quad (24)$$

Algorithm for converting annual GRP dynamics into Boolean function values

If y is a current GRP value, and y^{-1} is a previous one, then Δy is favorable dynamics of y by absolute deviation from y^{-1} . Then P can be obtained (formula 25):

$$\Delta y \rightarrow P = \begin{cases} 1, & \text{if } \Delta y > 0 \\ 0, & \text{if } \Delta y \leq 0 \end{cases}. \quad (25)$$

In accordance with the discussed methods, necessary statistical forms are presented below. They were developed by the authors to conduct appropriate statistical assessments, as well as the calculated results and the resulting comments and conclusions.

Table 1. Part of posteriori data set for the one-way analysis of variance

1. Russian Regional Innovation Index of HSE					
Regional territorial unit	Gross regional product, mil. rub. (y)	Socio-economic conditions of innovation activity (n = 1)	Scientific and technical potential (n = 2)	Innovative activity (n = 3)	Quality of innovation policy (n = 4)
Altai Krai	487903.3			0.4003	0.4948
Astrakhan Oblast	322303.0	0.4431			
Vladimir Oblast	368489.2			0.3688	
Vologograd Oblast	740458.0				0.4828
Vologda Oblast	478893.0				0.4986
Ivanovo Oblast	180517.5		0.3968		
Irkutsk Oblast	1001717.6		0.3821		
Kaliningrad Oblast	349818.6	0.4675			
Kaluga Oblast	339760.8	0.4829	0.4385		
Kamchatka Oblast	175404.8				0.4444
.....
Number of sample elements	m = 82*	m₁ = 28	m₂ = 16	m₃ = 15	m₄ = 23
2. Assessment of scientific and technological potential of VoIRC RAS					
		Research and development (n = 1)	Personnel (n = 2)	Technology (n = 3)	Innovation (n = 4)
Number of sample elements	m = 11*	m₁ = 4	m₂ = 3	m₃ = 3	m₄ = 1

* Total number of objects included in all k samples.
 Compiled according to: Rating of innovative development of the subjects of the Russian Federation. Issue 5. Moscow: NRU HSE, 2017. 260 p. Available at: <https://issek.hse.ru/news/206633461.html>; K.A. Gulin et al. Scientific and technological potential of a territory and its comparative appraisal. *Problems of Territory's Development*, 2017, no. 1 (87), pp. 7–26. Available at: <http://pdt.vsc.ac.ru/article/2102>. Federal State Statistics Service. Gross Regional Product (in current basic prices – total (1998–2017). Ed. March 1, 2019. Available at: http://www.gks.ru/free_doc/new_site/vvp/vrp98-17.xlsx (accessed: April 20, 2021).

Table 2. Results obtained during the analysis of variance

Content of the indicator	Unit of measurement	Symbolic designation	Calculated value	
			I_1	I_2
Average GRP value for the group S_1	mil. rub.	\bar{y}_1	1 183 444	1 204 667
Average GRP value for the group S_2	mil. rub.	\bar{y}_2	1 338 762	1 577 121
Average GRP value for the group S_3	mil. rub.	\bar{y}_3	1 810 813	4 743 038
Average GRP value for the group S_4	mil. rub.	\bar{y}_4	1 181 169	837 495
Average GRP value for all sample objects	mil. rub.	\bar{y}	1 327 875	2 237 876
Value of Fisher's calculated statistics	cond. un.	F_f	0.4422	0.4837
Value of Fisher's critical statistics	cond. un.	F_k	2.7218*	4.3468*

* With $\alpha = 0.05$.
Complied according to Table 1.

Results

Analysis of variance

To carry out the analysis of variance according to the algorithm described in the first paragraph of the “Research methods”, an array of data was formed (*Tab. 1*). It presents the values of key sub-indexes (socio-economic conditions, scientific and technical potential, innovation activity, quality of innovation policy) included in the overall assessment of the “Russian Regional Innovation Ranking”, developed by the ISSEK, and similar indicators (research and development, personnel, technology, innovation) to assess the scientific and technological potential of VolRC RAS. Data are taken in accordance with the latest current calculations carried out at the time of the study described in the 2015 work.

Each of the four groups included those GRP values of the regions. For them, the condition of getting the corresponding index in the significance group was met – group II for the HSE methodology and groups of the “average” and “above average” level for the methodology of VolRC RAS.

In accordance with formulas 1–13 and a posteriori data set, statistical metrics were obtained from table 1 (*Tab. 2*). They can be used to interpret the analysis of variance carried out in the work to determine the significant impact of the sub-indexes of the studied methods on the level of territories' economic growth, expressed in the values of the GRP indicator.

During the variance analysis, we tested the hypothesis about the impact and influence of artificially derived integral assessments of the rating of the Russian Regional Innovation Index and estimates of scientific and technological potential on the value of the territories' gross regional product (dynamics of economic growth). While selecting territories where the corresponding ratings were recorded for four sub-indexes for each method, eight samples were formed respectively, which are characterized by a high manifestation of the processes inherent in one of the four key groups of the studied ratings.

The calculated value of the F-statistics in both cases ($F_{f1} = 0.4422$, $F_{f2} = 0.4837$) turned out to be significantly less than the critical value ($F_{k1} = 2.7118$, $F_{k2} = 4.3468$) with a given reliability level of 5% and the corresponding degrees of freedom obtained ($df1_{1||2} = 3$ & ($df2_1 = 78$ || $df2_2 = 7$)) (*Tab. 2*). We may conclude that the differences between the groups of regions, included in the samples by large values (not the maximum possible ones) of the key indicators of the studied methods, are statistically insignificant in relation to the GRP differences. This may indicate two things: either these indicators do not have a significant impact on the regions' economic growth changes, or the compared comprehensive estimates are mutually dependent.

Table 3. Part of posteriori data set for conducting correlation analysis, building a model of multiple linear and logistic regression

Territory	GRP, 2014 (mil. rub.)		GRP, 2015 (mil. rub.)	Growth dynamics	Soc. and econom. cond. of innov. act., 2014 (n = 1)	Soc. and econom. cond. of innov. act., 2015 (n = 1)	Δn_1 by absolute value
	in current prices	in 2015 prices*					
1	2	3	4	5	6	7	8
Altai Krai	446023.8	479475.6	487903.3	↗ 1	0.3042	0.3038	-0.0207
Amur Oblast	232053.0	249457.0	277380.4	↗ 1	0.2893	0.3151	-0.0667
Arkhangelsk Oblast	542695.3	583397.4	627698.1	↗ 1	0.3317	0.2977	-0.0243
Astrakhan Oblast	296319.3	318543.2	322303.0	↗ 1	0.3997	0.4431	-0.0509
Belgorod Oblast	619677.7	666153.5	693379.4	↗ 1	0.4049	0.3946	-0.0191
Bryansk Oblast	242722.4	260926.6	271782.5	↗ 1	0.3245	0.3446	0.0456
Vladimir Oblast	328064.2	352669.0	368489.2	↗ 1	0.3755	0.3651	-0.0101
Vologograd Oblast	715409.6	769065.3	740458.0	0 ↘	0.3255	0.3435	-0.0947
Vologda Oblast	387211.7	416252.6	478893.0	↗ 1	0.3110	0.3191	-0.0215
Voronezh Oblast	717667.2	771492.2	805969.6	↗ 1	0.4253	0.3993	-0.0391
.....

* Deflator Index – 1.075 cond.un.
 Compiled according to: Rating of innovative development of the subjects of the Russian Federation. Issue 5. Moscow: NRU HSE, 2017. 260 p. Available at: <https://issek.hse.ru/news/206633461.html>;
 Gulin K.A. et al. Scientific and technological potential of a territory and its comparative appraisal. *Problems of Territory's Development*, 2017, no. 1 (87), pp. 7–26. Available at: <http://pdt.vssc.ac.ru/article/2102>.
 Federal State Statistics Service. Gross Regional Product (in current basic prices – total (1998–2017)). Ed. March 1, 2019. Available at: http://www.gks.ru/free_doc/new_site/vvp/vrp98-17.xlsx (accessed: April 20, 2021).

Data panel for performing correlation analysis, building multiple linear regression and a logit model

To carry out the selected analyses and modeling options, a statistical panel was formed. It compares the values of changes in the dynamics of the sub-indexes of the Russian Regional Innovation Ranking and the assessment of the RF territories' scientific and technological potential (Δn – column 8, etc., Tab. 3) with the fact of the GRP dynamics' growth (column 5, Tab. 3). The generated data were used in the correlation analysis, where information on 2015 sub-indexes was used (column 7, etc., Tab. 3) and the corresponding data characterizing this year's GRP values (column 4, Tab. 3). The same data set was used in the construction of the multiple linear regression model. To create a logit model, it was necessary to establish the fact of the GRP growth (column 5, Tab. 3) being a positive deviation from the previous year with an absolute deviation of the factors (column 8, etc., Tab. 3), potentially influencing this process.

Correlation analysis

The matrix from *Table 4* characterizes the linear relationship between the resulting indicator of the region's economic growth (GRP) and indicators that are the key sub-indexes of the Russian Regional Innovation Ranking. The data analysis was carried out based on statistics for 2015.

In accordance with the key indicators of the Russian Regional Innovation Ranking and the GRP value, inherent in the studied territories, we have made several relevant conclusions.

Gross regional product

There is a linear relationship between socio-economic conditions of work. A response from the values of the indicator characterizing the scientific and technical potential was revealed. Indicators describing the level of innovation activity and the quality of innovation policy have almost no linear effect on the growth or decline of economic growth in the regions of the Russian Federation.

Socio-economic conditions of innovation activity

This indicator strongly correlates with the values of the indicator characterizing the scientific and technical potential of the RF regions. This may be caused by high correlation between GRP and scientific and technical potential. There is a connection with the indicators of innovation activity, but it is insignificant with the quality of innovation policy.

Scientific and technical potential

The values of this indicator correlate with the values of indicators characterizing the innovative activity of the Russian regions. There is a stronger relationship with the quality of innovation policy than in previous indicators.

Scientific and technological potential

These indicators are more correlated with each other than the rest.

General conclusion

There is a pronounced interdependence between the studied metrics. It confirms the conclusions obtained during the analysis of variance, which showed the absence of a statistically significant

Table 4. The results of the correlation analysis carried out using the ISSEK methodology

	GRP	n_1	n_2	n_3	n_4
GRP	1	0.6669	0.4540	0.2652	0.1571
n_1	0.6669	1	0.5905	0.3935	0.2289
n_2	0.4540	0.5905	1	0.4007	0.2897
n_3	0.2652	0.3935	0.4007	1	0.4892
n_4	0.1571	0.2289	0.2897	0.4892	1
Complied according to Table 3.					

Table 5. Results of the correlation analysis held according to the VoIRC RAS methodology

	GRP	n_1	n_2	n_3	n_4
GRP	1	0.6674	0.6862	0.4558	0.3051
n_1	0.6674	1	0.8034	0.8205	0.3481
n_2	0.6862	0.8034	1	0.6430	0.2631
n_3	0.4558	0.8205	0.6430	1	0.3783
n_4	0.3051	0.3481	0.2631	0.3783	1

Complied according to Table 3.

difference between the GRP value of objects belonging to different categorical groups according to the levels of regions' innovative development.

The matrix from *Table 5* characterizes the linear relationship between GRP and the key sub-indexes of the rating, which assesses the territories' scientific and technological potential (2015).

1) *Gross regional product*

In accordance with the practice of econometric modeling, the relationship of GRP with indicators, which are chosen as predictors, is weak and cannot be properly used to predict or interpret the movement of the dynamics of the resulting estimate as a linear response from them.

2) *Research and developments*

The group of indicators that has the highest correlating values for all the studied sets of sub-indexes of the VoIRC RAS methodology. Its significant level is observed for the indicators of the "Technology" (0.8) and "Innovation" (0.82) groups. At the same time, these groups correlate with each other much less (0.64). Perhaps, it can be successfully used to build a model of multiple linear regression, where "Research and development" will act as the resulting estimate.

3) *Technologies*

Weak linear relationship with the group of indicators "Innovation" and "Personnel".

4) *Innovation*

There is practically no linear relationship with the last group of indicators characterizing the abstract systematization of scientific personnel.

General conclusion

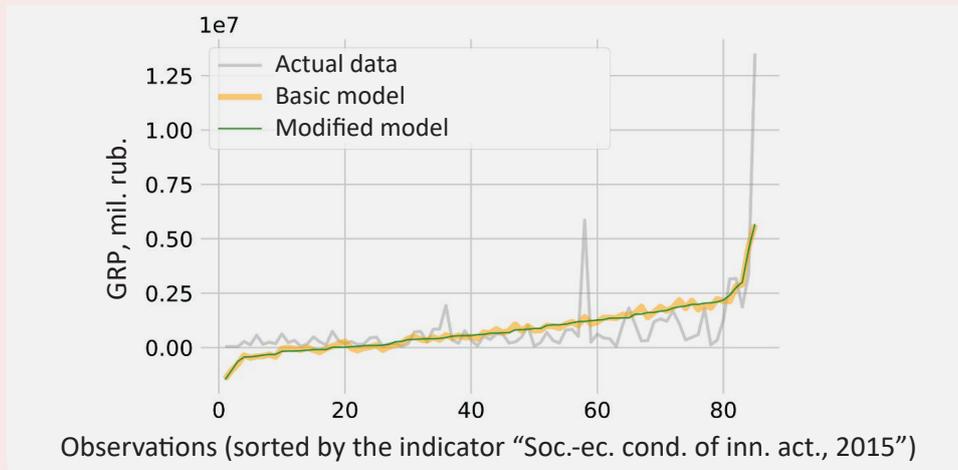
There is a situation similar to the case with the correlation analysis conducted according to the rating of the assessment of innovation development of the RF subjects. Specifically, it is necessary to highlight the relationship of the indicators in the "Technology" and "Innovation" groups with "Research and development". Nevertheless, there is still a strong interdependence of some metrics with others used to calculate the final resulting estimate. It indicates the need to use an effective method of searching for weights when forming the final calculation, or reducing the dimension of a number of predictors by eliminating irrelevant factors.

Multiple linear regression

Table 6 shows the main statistical metrics of the results obtained according to the statistical models of multiple linear regression constructed in the work. Namely, parameters of the regression model, both basic and modified; coefficients of determination and their adjusted estimate; calculated and critical Fisher F -statistics at a given reliability level of 5% and assessment of the significance of the obtained model parameters using t -statistics.

The graph of actual 2015 values of the regions' GRP, sorted by ascending values of the sub-index of the Rating of Innovation Development of the Russian subjects (*Fig. 1*) that characterize the socio-economic conditions for the implementation and carrying out of innovative activities in the country's territorial subjects, is given below. This

Figure 1. Graphical visualization of regression modeling using the HSE methodology



Compiled according to Table 3.

ranking is caused by the presence of the GRP values approximated by the only parameter, specified before, in the model. In this regard, we can try to graphically display the dependence of the studied value (gross regional product) on a specifically established predictor (integral assessment of socio-economic conditions of innovation activity).

Since the value F of the calculated statistics ($(F_f)_p = 16.3967$) exceeds the critical value ($(F_k)_p = 2.4859$) at the given significance level of 5%, then the hypothesis that all predictors in the regression model are simultaneously equal to zero is rejected, i.e. the constructed basic statistical model is statistically significant (Tab. 6). However, with a more detailed examination of it and tests for the obtained parameters, biased estimates were identified into t -statistics. Among five coefficients, including “intercept”, only two were significant: “intercept” and the parameter β_1 , which characterizes the socio-economic conditions of innovation activity. Insignificant parameters (β_2 – scientific and technical potential; β_3 – innovation activity;

β_4 – quality of innovation policy) were excluded from the modified model, and a new statistical model, corresponding to the conditions of F and t statistics was obtained. It was also possible to increase the value of the adjusted R^2 from 0.42 to 0.44.

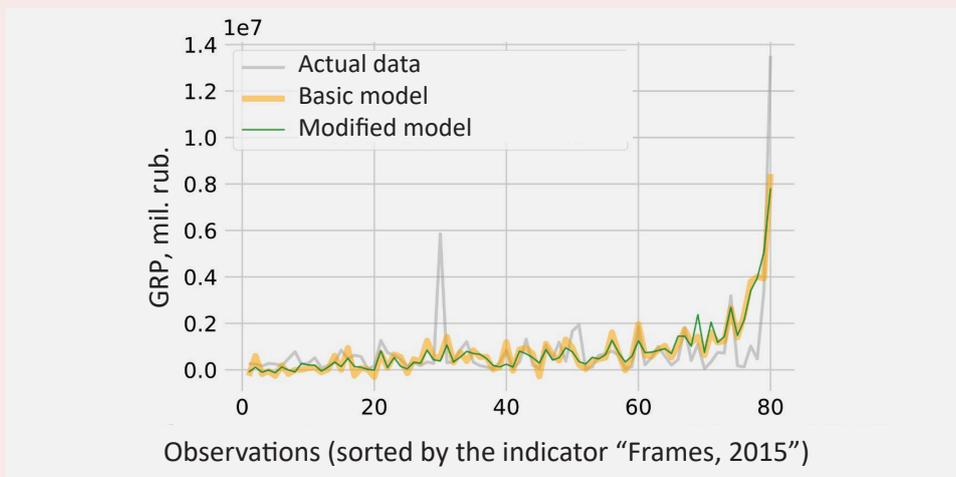
Similarly to the previous one, the results of modeling according to the methodology of VolRC RAS (Fig. 2) are presented. The sorting of the resulting and actual GRP values was carried out according to the “Frames” indicator in accordance with the highest value obtained when calculating its t -statistics (3.1554).

During the initial examination, we may note that the regressions constructed according to the second studied indicator (I_2) more accurately repeat the dynamics of actual GRP values. This is also evidenced by the adjusted coefficient of determination $R_{adj}^2(I_2) = 0.4962$ larger $R_{adj}^2(I_1) = 0.4380$. It may be caused by the decrease in the dimension of the indicators included in the second methodology. That is why the averaging of the studied estimates is not so pronounced. To some extent, this is

Table 6. Regression modeling results

Indicator's name	Symbolic designation	Calculated value	
		I_1	I_2
1. Basic regression model: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$			
$y_{I_1} = -3964780.87 + 11310999.28x_1 + 1854358.59x_2 - 173667.55x_3 - 50769.69x_4$			
$y_{I_2} = -663095.39 + 564364.88x_1 + 533221.86x_2 - 335361.73x_3 + 201959.81x_4$			
intercept	β_0	-3 964 780.9	-663 095.39
Calculated t -statistics for β_0	t_{β_0}	6.2250	2.7310
Socio-economic conditions of innovation activity / Research and development	β_1	11 310 999	564 364.88
Calculated t -statistics for β_1	t_{β_1}	5.8615	3.0278
Scientific and technical potential / Personnel	β_2	1 854 358.6	533 221.86
Calculated t -statistics for β_2	t_{β_2}	0.9157	3.1671
Innovation activities / Technologies	β_3	-173 667.55	-335 361.73
Calculated t -statistics for β_3	t_{β_3}	0.1326	2.1266
Quality of innovation policy / Innovations	β_4	-50 769.69	201 959.81
Calculated t -statistics for β_4	t_{β_4}	0.0564	1.4406
Critical point of Student's distribution	t_k	1.9901	1.9921
Coefficient of determination	R^2	0.4505	0.5437
Adjusted coefficient of determination	R_{adj}^2	0.4230	0.5194
Fisher's F -criterion, calculated	$(F_f)_p$	16.3967	22.3439
Critical value of Fisher's distribution	$(F_k)_p$	2.4859	2.4936
2. Modified regression model: $y'_{I_1} = \beta_0' + \beta_1' x_1$			
$y'_{I_1} = -3813320.76 + 12243723.31x_1$			
New intercept	β_0'	-3 813 320.76	
Calculated t -statistics for β_0'	t_{β_0}'	6.5117	
Socio-economic conditions of innovation activity (unbiased assessment)	β_1'	12 243 723.31	
Calculated t -statistics for β_1'	t_{β_1}'	8.1531	
New critical point of Student's distribution	t_k'	1.9890	
New coefficient of determination	R^2'	0.4447	
New adjusted coefficient of determination	$R_{adj}'^2$	0.4380	
New Fisher's F -criterion, calculated	$(F_f)_p'$	66.4736	
New critical value of Fisher's distribution	$(F_k)_p'$	3.9560	
3. Modified regression model: $y'_{I_2} = \beta_0' + \beta_1' x_1 + \beta_2' x_2$			
$y'_{I_2} = -393469.1 + 346573.61x_1 + 543161.88x_2$			
New intercept	β_0'	-393 469.10	
Calculated t -statistics for β_0'	t_{β_0}'	2.0592	
Research and development (unbiased assessment)	β_1'	346 573.61	
Calculated t -statistics for β_1'	t_{β_1}'	2.4416	
Personnel (unbiased assessment)	β_2'	543 161.88	
Calculated t -statistics for β_2'	t_{β_2}'	3.1554	
New critical point of Student's distribution	t_k'	1.9912	
New coefficient of determination	R^2'	0.5090	
New adjusted coefficient of determination	$R_{adj}'^2$	0.4962	
New Fisher's F -criterion, calculated	$(F_f)_p'$	39.9059	
New critical value of Fisher's distribution	$(F_k)_p'$	3.1154	
Complied according to Table 3.			

Figure 2. Graphical visualization of regression modeling by the VoIRC RAS methodology



Compiled according to Table 3.

an advantage of the VoIRC RAS methodology in relation to the one developed by HSE researchers.

By modified regression y'_{I_2} in comparison with y'_{I_1} , we should note that it has a greater number of degrees of freedom (per predictor). It provides appropriate flexibility and independence in predicting the resulting estimate.

Conducted research once again showed the absence of a significant influence of factors appearing in the Russian Regional Innovation Index and the Index of Scientific and Technological Potential, which, in turn, determine scientific and technological development and innovative activity in Russia. Although, in accordance with the provisions of the economic theory, they should have a direct impact on the dynamics of economic growth and development of the business entity. Considering all this, we may conclude that artificially derived indicators do not adequately characterize the territories' scientific and technological potential, innovative activities, and innovative policies carried out in a particular region.

Logistic regression

Table 7 provides the main estimates based on the developed and tested logit model that predicts the probability of the GRP growth in accordance with the absolute change in the sub-indexes of the rating of the Russian Regional Innovation Index and the assessment of scientific and technological potential. The basic logistic regression includes all four key indicators available in the rating, but due to the model's inconsistency (insignificance according to Fisher statistics and coefficients for the indicators included in the t -statistics model), there is no modification of it.

According to the obtained values of the F -test (Table 7), the compiled models are statistically insignificant, i.e. the cumulative change in the model parameters characterizing the change of socio-economic conditions of innovation activity, scientific and technical potential, innovation activity, quality of innovation policy or innovations, technology, personnel and research, and development in the regions does not manifest itself in a change in the GRP dynamics.

Table 7. Results of regression modeling of the logit model

Indicator's name	Symbolic designation	Calculated value	
		I_1	I_2
Logistic regression model: $\hat{P}' = \frac{1}{1 + e^{-(p_0 + p_1 \Delta x_1 + p_2 \Delta x_2 + p_3 \Delta x_3 + p_4 \Delta x_4)}}$			
$\hat{P}'_{I_1} = \frac{1}{1 + e^{-(0.7466 - 1.4252 \cdot \Delta x_1 + 1.8070 \cdot \Delta x_2 - 1.7907 \cdot \Delta x_3 + 0.3162 \cdot \Delta x_4)}}$			
$\hat{P}'_{I_2} = \frac{1}{1 + e^{-(0.7111 + 0.0775 \cdot \Delta x_1 + 0.1452 \cdot \Delta x_2 - 0.1321 \cdot \Delta x_3 + 0.1373 \cdot \Delta x_4)}}$			
intercept	p_0	0.7466	-0.7111
Calculated t -statistics for p_0	t_{p_0}	13.4386	11.3365
Changes of socio-economic conditions of innovation activity	p_1	-1.4252	0.0775
Calculated t -statistics for p_1	t_{p_1}	1.0734	0.3926
Changes of scientific and technical potential	p_2	1.8070	0.1452
Calculated t -statistics for p_2	t_{p_2}	1.2301	0.4324
Changes of innovation activity	p_3	-1.7907	-0.1321
Calculated t -statistics for p_3	t_{p_3}	1.8590	2.0185
Changes of the quality of innovation policy	p_4	0.3162	0.1373
Calculated t -statistics for p_4	t_{p_4}	0.4394	1.9119
Critical point of Student's distribution	t_k	1.9908	1.9921
Fisher's F -criterion, calculated	$(F_f)_p$	1.9378	1.4182
Critical value of Fisher's distribution	$(F_k)_p$	2.4889	2.4937
Complied according to Table 3.			

Thus, we may draw the following conclusion: considered artificially derived estimates do not effectively reflect the processes and phenomena of scientific and technological development, which, in turn, should directly affect the values of economic growth indicators.

Regarding partial coefficients of the obtained model, we may note that their values also turned out to be untenable in most cases in accordance with the t -test. The only exception is the “intercept” element, which does not carry any interpretation, but it is responsible for optimizing the spatial location of the resulting model’s graph. Once again, it indicates the absence of a relationship between GRP changes and indicators of scientific and technological development. It is necessary to look for other parameters to determine the probability of a positive change of the gross regional product.

We should mention that one of the parameters included in the basic model \hat{P}'_{I_2} , characterizing “Technologies” (p_3), has t -statistics acceptable

(2.0185) in comparison with the calculated critical level of 1.9921. However, the hypothesis of its difference from zero was not confirmed with the further exclusion of insignificant predictors from the model. In this case, only the “intercept” parameter remained significant. It is characterized by other conditions that are not included in the model and cannot be adequately interpreted.

Conclusion

To conclude, we might say the following about integral assessments characterizing the territory’s aggregate level of scientific and technological development: such metrics do not adequately characterize the processes and phenomena they are intended for. Their reliability is doubtful, and it is caused by the fact that the studied object is often an artificially derived category, which is interpreted differently by scientists and researchers.

It could be compared to a diagnosis, which allows saying that a patient is sick. But what exactly is the problem? Such an approach does not provide

Figure 3. Assessment of the reliability of integrated assessments of socio-economic objects

1. Linking the integrated assessment to the resulting indicator, which will be influenced by changing integral assessment				
2.1. One-way analysis of variance	2.2. Correlation analysis	2.3. Multiple linear regression	2.4. Logistic regression of changes in dynamics	2.5. Other methods of statistical analysis and data processing

Source: own compilation.

an answer. As a result, we cannot take appropriate measures to eliminate negative effects.

We should also mention the specifics of socio-economic indicators. Unlike the natural sciences, simulations based on such estimates cannot be reproduced for testing in laboratory conditions. In this regard, the reliability and adequacy of such models is based only on theoretical propositions that may turn out to be both correct and refuted over time.

In addition, let us note that the nature of socio-economic metrics is unstable and changeable. That is, the same indicator can mean different things at different times. These indicators are narrow-profile, and they characterize various unique objects, which cannot be stated about the mass of a material object, the speed of movement, the presence of a certain set of genes and other things, inherent in the objects of natural sciences research.

Largely, the bias of the estimates obtained is caused by the fact that duplicate (highly correlated) indicators are used in their compilation. With their significant share in the final calculation, the value of the integral indicator is averaged, and it does not reflect the differentiation in the totality of the studied objects.

To avoid such problems, discussed statistical methods can be used. At the same time, they can be

applied separately and in combination, complementing and confirming the corresponding statistical hypotheses, or indicating the need for additional research (*Fig. 3*).

In accordance with figure 3, a stable indicator should be determined first. It is recorded in statistical reports without any changes and will potentially be used. Then we need to establish a connection between it and the integral assessment (as shown in the work, it can be done by means of variance, correlation, and regression analysis), which will allow assessing the reliability of the developed integral methodology. The more statistical tests are applied to the assessment of the relationship between the resulting metric and the integral one, and the more coincidences there are between them, the more stable and reliable the conclusions are, which can be obtained when comparing socio-economic objects evaluated using integral methods, are.

Methods designed to reduce the dimensionality and the number of studied processes and phenomena are also important in solving this problem. The final model should include significant factors, which is shown by the methods discussed.

We should also mention the approach associated with the use of narrow-profile econometric models that consider spatial locations of the studied object;

time lag of the manifestation of an event; the nonlinear nature of the process, etc. These problems and questions are not trivial, and they are more applicable to metrics and indicators that can be removed from the studied object. In case of assessments similar to integral or comprehensive one, such approach is unlikely to be successful, since it includes things that occur independently

in various manifestations. It will just confuse the researcher during mathematical modeling.

The materials of the publication might be used by experts specializing in scientific and technological development, innovation, and technological promotion, as well as by scientists and researchers engaged in statistical processing and data analysis.

References

1. Abalkin L.I. *Logika ekonomicheskogo rosta* [Logic of Economic Growth]. Moscow: Inst. of Econ. RAS, 2002. 228 p.
2. Uskova T.V. et al. *Problemy ekonomicheskogo rosta territorii: monografiya* [Problems of Economic Growth of a Territory: Monograph]. Vologda: ISEDT RAS, 2013. 170 p.
3. Shumpeter J.A. *Capitalism, Socialism and Democracy*. London & New York: Routledge, 2003. 460 p.
4. Foster R., Kaplan S. *Creative Destruction: Why Companies That Are Built to Last Underperform the Market – and How to Successfully Transform Them*. Crown Business, 2001. 384 p.
5. Schumpeter J.A. *The Theory of Economic Development*. Moscow: Progress, 1982. 401 p.
6. Grasmik K.I. Innovative activity of firms during the economic crisis. *Problemy teorii i praktiki upravleniya=Problems of Management Theory and Practice*, 2017, no. 2, pp. 58–64. Available at: <https://elibrary.ru/item.asp?id=28289539> (in Russian).
7. Kornai J. Innovation and dynamism. Interaction between systems and technical progress. *Voprosy Ekonomiki*, 2012, no. 4, pp. 4–31. DOI: 10.32609/0042-8736-2012-4-4-31 (in Russian).
8. Solow R. Contribution to the theory of economic growth. *Quarterly Journal of Economics*, 1957, no. 70 (1), pp. 65–94.
9. Rebelo S. Long-run policy analysis and long-run growth. *Journal of Political Economy*, 1991, no. 3, pp. 500–521 (in Russian).
10. Komkov N.I. Scientific and technological development: limitations and opportunities. *Problemy prognozirovaniya=Studies on Russian Economic Development*, 2017, no. 5, pp. 11–21. Available at: <https://ecfor.ru/publication/nauchno-tehnologicheskoe-razvitie-ogranicheniya-i-vozmozhnosti/> (in Russian).
11. Chernykh S., Frolova N. On the participation of Russian business in the financing of scientific and technological research and development (economic and ideological aspects). *Obschestvo and ekonomika=Society and Economy*, 2018, no. 11, pp. 86–97 (in Russian).
12. Ross A. *Industries of the Future*. Moscow: AST, 2017. 351 p.
13. Schwab K. *The Fourth Industrial Revolution*. Moscow: Eksmo, 2018. 288 p.
14. Schwab K. *Shaping the Fourth Industrial Revolution*. Moscow: Eksmo, 2018. 320 p.
15. Kaneva M.A., Untura G.A. Evolution of theories and empirical models of a relationship between economic growth, science, and innovations (Part I). *Mir ekonomiki i upravleniya=World of Economics and Management*, 2017, vol. 17, no. 4, pp. 5–21. DOI: 10.25205/2542-0429-2017-17-4-5-21 (in Russian).
16. Kaneva M.A., Untura G.A. Evolution of theories and empirical models of a relationship between economic growth, science, and innovations (Part II). *Mir ekonomiki I upravleniya=World of Economics and Management*, 2018, vol. 18, no. 1, pp. 5–17. DOI: 10.25205/2542-0429-2018-18-1-5-17 (in Russian).

17. Arrow K. Economic welfare and allocation of resources for invention. In: *The Rate and Direction of Inventive Activity*. Princeton Uni. Press, 1962. Pp. 609–625.
18. Romer P.M. Increasing returns and long-run growth. *Journal of Political Economy*, 1986, no. 94 (5), pp. 1002–1037.
19. Lucas R. On the mechanics of economic development. *Journal of Monetary Economics*, 1988, no. 22, pp. 3–42.
20. Grossman G.M., Helpman E. *Innovation and Growth in the Global Economy*. Cambridge: MIT Press, 1991. 384 p.
21. Freeman C. *Technology Policy and Economic Performance: Lessons from Japan*. London: Pinter, 1987. 155 p.
22. Griliches Z. Issues in assessing the contribution of research and development to productivity growth. *The Bell Journal of Economics*, 1979, no. 10, pp. 92–116.
23. Zadumkin K.A., Kondakov I.A. *Nauchno-tehnicheskii potentsial regiona: otsenka sostoyaniya i perspektivy razvitiya: monografiya* [Scientific and Technical Potential of the Region: Assessment of the State and Prospects of Development: Monograph]. Vologda: ISEDT RAS, 2010. 205 p.
24. Gulin K.A. et al. Scientific and technological potential of a territory and its comparative appraisal. *Problemy Razvitiya Territorii=Problems of Territory's Development*, 2017, no. 1 (87), pp. 7–26. Available at: <http://pdt.vsc.ac.ru/article/2102> (in Russian).
25. Epstein N.D., Karmanov M.V., Vasileva A.V. Problems of integrated assessment and security demographic security. *Statistika i ekonomika=Statistics and Economics*, 2015, no 3, pp. 233–237. DOI: 10.21686/2500-3925-2015-3-233-237 (in Russian).
26. Taylor G. *An Introduction to Error Analysis*. Translated from English Moscow: Mir, 1985. 272 p.
27. Sveshnikov A.A. *Osnovy teorii oshibok* [Fundamentals of Error Theory]. Leningrad: LGU Publ., 1972. 122 p.
28. Volkova E.N. Problems of integral statistical evaluation of social-economic development region. *Ekonomika i statistika=Statistics and Economics*, 2015, no. 3, pp. 170–175. Available at: <https://core.ac.uk/download/pdf/234096888.pdf> (in Russian).
29. Klyushnikova E.V., Shitova E.M. Methodological approaches to calculation of integral index, ranking methods. *InnoCenter*, 2016, no. 1 (10). Available at: <http://innoj.tversu.ru/number10.html> (in Russian).
30. Saib M-S., Caudeville J., Beauchamp M. et al. Building spatial composite indicators to analyze environmental health inequalities on a regional scale. *Environ Health*, 2015. Nj. 68. DOI: 10.1186/s12940-015-0054-3
31. Sidorov A.A. Methodological approach to composite estimation of status and dynamics of multi-dimensional objects of social and economic nature// *Problemy upravleniya=Control Sciences*, 2016, no. 3, pp. 32–40. Available at: http://pu.mtas.ru/archive/pu_316.php (in Russian).
32. Aivazian S.A. On the methodology of measuring synthetic categories of the quality of life of the population. *Ekonomika i matematicheskie metody=Economics and Mathematical Methods*, 2003, vol. 39, no. 2, pp. 33–53 (in Russian).
33. Aivazian S.A. Empirical analysis of synthetic categories of the quality of life of the population. *Ekonomika i matematicheskie metody=Economics and Mathematical Methods*, 2003, vol. 39, no. 2, pp. 19–53 (in Russian).
34. Aivazian S.A., Stepanov V.S., Kozlova M.I. Measuring the Synthetic Categories of Quality of Life in a Region and Identification of Main Trends to Improve the Social and Economic Policy (Samara Region and its Constituent Territories). *Prikladnaya ekonometrika=Applied Econometrics*, 2006, no. 2, pp. 18–84. Available at: <https://ideas.repec.org/a/ris/apltrx/0086.html> (in Russian).
35. Zhou P., Ang B.W., Zhou D.Q. Weighting and aggregation in composite indicator construction: A multiplicative optimization approach. *Social Indicators Research*, 2010, no. 96 (1), pp. 169–181. DOI: 10.1007/s11205-009-9472-3

Information about the Authors

Dmitrii A. Alfer'ev – Candidate of Sciences (Economics), Researcher, Vologda Research Center, Russian Academy of Sciences (56A, Gorky Street, Vologda, 160014, Russian Federation; e-mail: alferev_1991@mail.ru); Assistant, Higher School of Engineering and Economics, Institute of Industrial Management, Economics, and Trade, Peter the Great St. Petersburg Polytechnic University (29, Polytechnicheskaya Street, Saint Petersburg, 194064, Russian Federation)

Aleksandr E. Kremin – Candidate of Sciences (Economics), Vologda Research Center of the Russian Academy of Sciences (56A, Gorky Street, Vologda, 160014, Russian Federation; e-mail: akremin@inbox.ru); head of section, Vologda Institute of Law and Economics, Federal Penitentiary Service of Russia (2, Schetinin Street, Vologda, 160026, Russian Federation)

Dmitrii G. Rodionov – Doctor of Sciences (Economics), Professor, director, Higher School of Engineering and Economics, Institute of Industrial Management, Economics, and Trade, Peter the Great St. Petersburg Polytechnic University (29, Polytechnicheskaya Street, Saint Petersburg, 194064, Russian Federation; e-mail: drodionov@spbstu.ru)

Darya S. Velichenkova – Candidate of Sciences (Economics), 2nd Category Engineer, Higher School of Engineering and Economics, Institute of Industrial Management, Economics, and Trade, Peter the Great St. Petersburg Polytechnic University (29, Polytechnicheskaya Street, Saint Petersburg, 194064, Russian Federation; e-mail: velichenkova_ds@spbstu.ru)

Received April 23, 2021.