

NAR, 1(1): 42–61. DOI: 10.3934/NAR.2019.1.42 Received: 30 April 2019 Accepted: 02 August 2019 Published: 20 August 2019

http://www.aimspress.com/journal/NAR

Research article

Constructing a stable complex index of a system's quality for a series of observations

Tatyana V. Zhgun*

Department of Applied Mathematics and Computer Science, Novgorod State University, Velikiy Novgorod, Russia

* Correspondence: E-mail: Tatyana.Zhgun@novsu.ru.

Abstract: This paper discusses the solution to the problem of constructing latent complex indexes of a change in a system's quality for several observations in the absence of training. An analysis of the stability of such a solution is also provided. The algorithm for constructing complex indexes is implemented with the definition of non-random variables of the principal component characterizing the structure of the system under discussion. The algorithm uses a new approach to choose the principal component number, determine the weights of the considered variables and subsystems, and to determine the information content of the complex indexes of quality of life for Russia's constituent entities for 2007–2016. The analysis of the obtained solution's quality shows its high resistance to changes in the input data.

Keywords: composite index; data error; principal component analysis; signal-to-noise ratio; signal detection in noise; informativeness of the principal component analysis; stability of composite index; rank robustness tests

JEL Codes: C38, I38, O15, R13

1. Introduction

In most problems solved in the field of measuring technology, radar, astronomy, optical communication, positioning, navigation, television automation and many other broad areas of science and technology, one of the fundamental problems is finding the best way to recognize a signal in the presence of interference. The ideal reception of signals under the influence of noise and interference is based on simple yet deep ideas, set out in the most consistent and understandable form by Woodward (Woodward, 1955). Following this idea, the task of an ideal device, whose input receives a mixture of signal with noise, is the complete removal of unnecessary information contained in this mixture, and the preservation of useful information about those signal parameters that are of interest to the system user.

The task of calculating composite indicators characterizing the quality of a system can also be attributed to the problems of signal isolation against a background of interference. Comparison of the considered objects' integral indicators and ratings generated by their composite indices, makes it possible to judge the degree of achievement of the management objective. Although the scientific community has not come to a common opinion, if it is possible in principle to characterize a multifaceted phenomenon as a single scalar. International organizations, think tanks and social sciences at the turn of the millennium significantly increased the number of integrated indicators used to measure various latent characteristics of social and economic systems: social capital, human development, quality of life, quality of management, etc.

According to the UN, by 2011 there were 290 indices developed for the ranking or integrated assessment of countries. A review of the phenomenal growth in the number of complex indexes used by countries for integrated assessment performed by Bandura (Bandura, 2006, 2008, 2011) showed that only 26 of them (9%) were formed before 1991, 132 (46%) from 1991 to 2006, and 130 consolidated indices were formed in the period 2007–2011. A huge number of methods used to assess the latent characteristics of social systems highlight the dissatisfaction with the results and the need for further research in this area.

The rapid increase in the number of complex indexes is a clear sign of their importance in politics and the economy. All major international organizations, such as the Organization for Economic Cooperation and Development (OECD), the European Union, the World Economic Forum or the International Monetary Fund, are constructing composite indicators in various fields (Bandura, 2011; Nardo et al., 2005). The overall goal of most of these indicators is the ranking of objects (countries) and their comparative analysis for some aggregate measure (Bandura, 2011; Foa and Tanner, 2012; Saltelli et al., 2006; Sharpe, 2004). The use of a single indicator characterizing poorly formalized processes of social systems (quality of life, demographic situation, etc.) is the only possible solution to this problem. Therefore, improving the quality of composite indicators is relevant both from a theoretical and from a practical point of view. A discussion of the pros and cons of composite indicators is given in (Nardo et al., 2005; Foa and Tanner, 2012).

Despite the huge number of used composite indicators, unresolved methodological problems in their design lead to the fact that very often they raise more questions than they give answers. The huge number of applied methods for assessing the latent characteristics of social systems indicates the dissatisfaction of researchers with the results and the need for further research in this area.

The Organization for Economic Cooperation and Development (OECD) is constantly working to improve the methods for constructing complex indexes (Nardo et al., 2005; Saltelli et al., 2006;

Nicoletti et al., 2000; Saltelli, 2007). In 2008, the OECD, together with the Joint Research Center of the European Commission, prepared a (Handbook, 2008), which was the result of years of research in this field (Nardo et al., 2005; Saltelli et al., 2006; Nicoletti et al., 2000, Saltelli, 2007; Tarantola et al., 2002). The Handbook sets out a set of technical principles for the formation of composite indices. The authors choose a linear convolution of indicators for the primary method of the data's aggregation, while the main tool for constructing composite indicators is factor analysis.

So, to build a qualitative integral indicator, it is required: first, a thorough theoretical study of the measured phenomena's theoretical aspects, since "that which is not well defined is likely to be poorly measured" (Nardo et al., 2005), secondly, the choice of data quality, because the quality of the composite indicators largely depends on the quality of the main indicators and, thirdly, an adequate tool for working with multidimensional data.

Such a tool for working with multidimensional data is multidimensional statistical data analysis, namely, factor analysis and principal component analysis (PCA). Factor analysis was first used to combine a multitude of indicators into a single index in the development of the Health Index by (Hightower, 1978).

When calculating Socio-Economic Status Indices (SES), the principal component analysis was adopted as a standard for the construction method, and the calculated index was determined with a projection onto the first principal component (McKenzie, 2005; Vyas S and Kumaranayake, 2006). The same methodology was used by Lindman and Selin in creating the Environmental Sustainability Index (Lindman and Selin, 2011). Somarriba and Penna used PCA in measuring the quality of life in Europe (Somarriba and Pena, 2009). Mention should also be made of (Ajvazjan, 2003) on the definition of a population's quality of life index using the first principal component.

However, the first principal component approximates the simulated situation well if the maximum eigenvalue of the covariance matrix contributes at least 70% to the sum of all eigenvalues. This relationship is satisfied if a small number of features are considered (no more than five), and one of the properties of the system clearly dominates over the others. When describing socio-economic systems, the number of variables significantly exceeds five, and the structure of the system does not allow simple approximation. As a way out of this situation, according to (Ajvazjan, 2003), the information threshold is reduced to 55%, and the system is divided into subsystems described with a smaller number of variables. In the studies discussed above (Hightower, 1978; McKenzie, 2005; Vyas and Kumaranayake, 2006; Somarriba and Pena, 2009), the contribution of the largest eigenvalue ranged from 13% to 38%, with the exception of (Somarriba and Pena, 2009), in which an example of a model was considered, where this figure was 56%.

The authors followed the recommendations (Somarriba and Pena, 2009), which stated that the first principal component gives satisfactory weight coefficients when calculating the integral index even in cases where the largest eigenvalue makes a small contribution to the sum of all eigenvalues. However, such a statement cannot be called indisputable. In the study (Mazziotta and Pareto, 2016) the contribution of the largest eigenvalue of 74% the authors consider insufficient.

For a fixed t an integrated assessment is often recorded for each of the object numbered i in the form of additive data convolution with some weights. Researchers at the Organization for Economic Cooperation and Development hold a different point of view. For the formation of a composite index, factor analysis is used, where the analysis of the first principal components is used exclusively for extracting factors, so that the number of factors extracted explains more than 50% of the total variance. The value of the composite index in this case is determined only with significant loads of

45

the selected main factors after rotation. De facto, this method becomes the standard for calculating complex indexes (Nardo et al., 2005; Nicoletti et al., 2000; Tarantola et al., 2002). Although the authors of (Nardo et al., 2005) pay attention to the fact that different methods of extracting the principle components and different methods of rotation imply different significant variables and, therefore, different weights of variables when calculating the composite index, hence different values of the calculated integral indicator. In addition, factor analysis assumes there is sufficient correlation between the initial variables, which to some extent contradicts the idea of a complete description of the phenomenon under study by a set of independent variables.

Another circumstance should also be noted. Methods for determining weight coefficients using factor analysis cannot be used to compare the characteristics of objects described in dynamics. For different observations, the structure of factors is different, even for fixed methods of extracting factors and the rotation process. This makes intertemporal comparison meaningless (Zhgun, 2013). The reason for the different structure of the main factors may be the insufficient quality of the data used, namely the presence of errors in the data. Nevertheless, it is the statistical data containing fatal errors that currently represent the best estimates of the available real values in social systems (Nardo et al., 2005).

It is impossible to obtain exact characteristics of an object based on a single measurement, which inevitably contains an unknown error. However, based on a series of such measurements, it is possible to calculate the unknown characteristic. In particular, astrophotometry successfully solves this problem. It determines the basic numerical parameters of astronomical objects not on a single observation (image), but on a series of noisy images. Using the basic ideas underlying astrophotometry, we consider the construction of the complex indices changes in the quality of a complex system as a solution to the problem selection of the useful signal on a series of observations containing a description of the unknown parameter (in a multidimensional dataset with noise) with a priori uncertainty of the desired signals' properties based on a specified signal-to-noise ratio.

The goal of this manuscript is to construct a composite system quality index as a solution to the problem of extracting a signal against a background of noise. For this it is suggested to take linear convolution weights as the characteristics of a system's structure. While in order to determine the structure of the system according to a series of observations a modification of the PCA for a noisy signal is proposed. Further the stability and reliability of the composite index obtained is investigated.

2. Calculation of the integrated system indicator for a series of observations

To solve the control problem, it is required to give a motivated estimate of each observed object on the entire observation interval, i.e. to calculate in dynamics the integral characteristic of system quality according to the results of available measurements. Let's consider the construction of an integral estimation of a system of m objects for which tables of n descriptions of those objects for a series of observations are known. For each moment t the vector of integral indicators is written as

$$q^{t} = A^{t} \cdot w^{t} \tag{1}$$

For a fixed t an integrated assessment is often recorded for each of the object numbered i in the form of additive data convolution with some weights

$$q_i = \sum_{j=1}^n w_j \cdot a_{ij}, \quad i = 1, 2, \quad m,$$
(2)

where $q^t = (q_1^t, q_2^t, ..., q_m^t)^T$ —is the vector of integral indicators of the moment t, q^t —is the vector of the weights for the moment t, and A^t — is the data matrix for the time t. The numerical characteristics of the system are preliminarily subjected to unification—the reduction of the variables' values to the interval [0, 1]. If the initial indicator is related to the analyzed integral quality feature by monotone function, their putting to the interval [0, 1] will transform the variables x_{ij} for each time point of observation according to the rule:

$$a_{ij} = s_j + (-I)^{s_j} \cdot \frac{x_{ij} - m_j}{M_j - m_j},$$
(3)

where $s_j = 0$ if the optimal value of the *j*-th index is maximum and $s_j = 1$ if the optimum value of *j*-th index is minimal and

$$M_{j} = \max_{\substack{1 \le i \le n \\ 1 \le t \le T}} x_{ij}^{t}, \quad m_{j} = \min_{\substack{1 \le i \le n \\ 1 \le t \le T}} x_{ij}^{t}.$$
(4)

 m_j – minimum value of the *j*-th index for the whole sample (global minimum), M_j – maximum value of *j*-th index for the whole sample (global maximum). In case where there is another optimal (not minimum or maximum) index value, the following formula is used:

$$a_{ij} = \left(1 - \frac{\left|x_{ij} - x_{j}^{opt}\right|}{\max\left((M_{j} - x_{j}^{opt}), (x_{j}^{opt} - m_{j})\right)}\right).$$
(5)

To build an integral indicator of the systems' quality, it is required to find weights w_t for each time point. The method of expert assessments for determining weights is widely used due to the ease of obtaining information. However, not every complex system has a sufficient number of qualified experts. In addition, expert services are a commodity, so they are not cheap and cannot be objective. It is preferred to use formal methods that do not use human preferences (Nardo et al., 2005; Nicoletti et al., 2000; Ajvazjan, 2003). On the other hand, the extent of the importance (weights) of the variables calculated using formal methods often does not coincide with the intentions of the developers. Therefore, the weights are carefully investigated. (Becker et al., 2017a, 2017b; Paruolo et al., 2013).

The Principal Components Analysis (PCA) is perhaps the most-used method to obtain weights intrinsically. However, the application of this technique gives unexpected results. For example, in a study (Molchanova et al., 2014), the quality of life is determined by PCA (Ajvazjan, 2003). Obviously, geographically close regions should have comparable quality of life ratings. For example, Novgorod and Pskov are neighboring constitute entities and close in all parameters. However, the rating of the Novgorod province is 29, while the rating of the Pskov province is 54 (out of 83 positions). It is also obvious that the quality of life near a megapolis is higher than in Siberia and the Far East. Nevertheless, the rating of the Leningrad province is 63, lower than the ratings of

Novgorod and Pskov provinces, almost as low as the rating of Sakhalin 64. Complex indices, that give such ratings, cannot be called reliable.

If technique PCA is applied to a series of observations made at different times, then, in the condition of a stable situation, the ratings have strong fluctuations (Gajdamak and Hohlov, 2009). In this case as well, the complex indices, that determine such ratings, cannot be called reliable. The numerical evaluation of such changes for (Gajdamak and Hohlov, 2009) is given further in paragraph 3.

In the paper (Mazziotta and Pareto, 2016) the authors consider that PCA is a blindly empiricist method based on the observed correlations and it ignores the polarity of the individual indicators. Therefore, if the normalized indicators are not all positively intercorrelated, the results are not correct. It should be noted that the amount of variance accounted for, and the weights computed by PCA change over time, so the results of different PCAs are not easily comparable. Consequently, they believe the use of PCA for the construction of composite indices for assessing multidimensional phenomena is at all improper. It should be noted that this conclusion was made when constructing a composite index for a single observation.

However, any measurement, including statistical, is determined by the accuracy of the measuring device, therefore the measurement result always contains an unavoidable error. The construction of system's integral characteristic can be considered as the task of isolating a useful signal against a noisy background. This problem is analogous to the problem of restoring digital images distorted by Gaussian noise. The principal component analysis makes it possible to isolate the structure in a noisy array of data and it is successfully used to reduce noise in image restoration.

The quantitative characteristics of a particular system, related to its structural features, depend on the signal-to-noise ratio. SNR—signal-to-noise ratio (SNR) is the ratio of the signal (more precisely, the sum of the signal and noise) to noise. The value is calculated either as a dimensionless ratio of the signal amplitude to the noise amplitude $SNR = A_s / A_n$, or, in decibels, $SNR (dB) = 20 \log_{10} (A_s / A_n)$.

This ratio most fully describes the quality of the signal in technical systems: television, in means of control and diagnostics, in mobile communication systems, in astrophotometry, etc. This ratio most fully describes the signal quality in technical systems: television, in control and diagnostic tools, in mobile communication systems, in astrophotometry, etc. The choice of the SNR threshold value, which makes it possible to distinguish a signal against noise is justified (Zhgun, 2014).

Moving to another point in time means a change in the data, which is caused by both a change in the situation and random errors. The analysis of the principal components based on different values (for consecutive moments) of eigenvectors and eigenvalues describes the invariable structure of the system. Therefore, the undistorted values of the eigenvalues and eigenvectors will be a signal that must be extracted from the noisy data. For eigenvalues, such a signal is the average. Averaging works on the assumption that the noise is absolutely random. Such averaging of the values is used in astrophotography for noise suppression. The assumption that there is a general trend in the variation of input data is illustrated in Figure 1, where the values of the covariance matrices ordered in descending eigenvalues for different observations are presented. In the mean, a tendency (signal) and a deviation from it (noise) are clearly visible.



Figure 1. The eigenvalues of the covariance matrix of the variables for different observation times. Right yellow color is the mean of the eigenvalues. The variables considered are from the first block of Table 2. Source: Rosstat // http://www.gks.ru.

The eigenvectors in the PCA are determined up to the direction, in contrast to the eigenvalues, which are uniquely determined. The mean value of variables' factor loadings depends on the chosen direction and cannot unambiguously characterize the signal.



Figure 2. Coordination of eigenvectors directions for the first principal component. Left —the original directions of the eigenvectors, on the right—the choice of the direction of the eigenvectors, which ensures the maximization of the signal level. The variables considered are from the first block of Table 2. Source: Rosstat // http://www.gks.ru.

Based on the eigenvectors calculated for various observations (ordered in decreasing order of eigenvalues), it is necessary to recognize the random and nonrandom components of these vectors.

We assume that the nonrandom (significant) contribution of a variable to the structure of the main component is not a large value of the factor load after rotation, but an invariance of the factor load with data disturbances. A sign a variable's invariance is the signal-to-noise ratio, which is calculated from the values of this variable's factor loads. The signal amplitude is the absolute value of the mean of the load factors, the noise amplitude is the standard deviation of the load factors.

Approval of directions for the first variable												
Years					Variable	es						
	1	2	3	4	5	6	7	8	9			
2007	-0.07	-0.1	0.26	0.38	0.08	-0.21	0.67	-0.28	-0.44			
2008	-0.05	0.19	0.04	-0.14	-0.33	-0.18	-0.41	0.2	0.77			
2009	-0.16	0.22	0.19	-0.06	-0.49	-0.2	-0.24	0.04	0.74			
2010	-0.17	0.22	0.24	-0.06	-0.37	-0.25	-0.14	0.16	0.79			
2011	-0.24	0.18	0.31	0.07	-0.23	-0.2	0	0.21	0.82			
2012	-0.25	0.18	0.28	0.07	-0.4	-0.23	-0.03	0.17	0.76			
2013	-0.2	0.17	0.2	0.01	-0.47	-0.17	-0.03	0.12	0.79			
2014	-0.3	0.14	0.18	0	-0.38	-0.22	0.04	0.18	0.79			
Mean value, m	-0.18	0.15	0.21	0.03	-0.32	-0.21	-0.02	0.1	0.63			
Standard deviation, s	0.09	0.1	0.08	0.16	0.18	0.03	0.32	0.16	0.43			
Signal/noise Ratio, SNR	2.1	1.47	2.51	0.22	1.76	7.94	0.06	0.61	1.45			
Sum of SNRs for the line	18.13											
Sum of significant SNRs	10.54											
AI	oproval o	f direct	ions for t	the secon	d variab	le						

Table 1. Definition of the empirical principal component.

Approval of directions for the second variable											
Years	_				Variable	s					
	1	2	3	4	5	6	7	8	9		
2007	0.07	0.1	-0.26	-0.38	-0.08	0.21	-0.67	0.28	0.44		
2008	-0.05	0.19	0.04	-0.14	-0.33	-0.18	-0.41	0.2	0.77		
2009	-0.16	0.22	0.19	-0.06	-0.49	-0.2	-0.24	0.04	0.74		
2010	-0.17	0.22	0.24	-0.06	-0.37	-0.25	-0.14	0.16	0.79		
2011	-0.24	0.18	0.31	0.07	-0.23	-0.2	0	0.21	0.82		
2012	-0.25	0.18	0.28	0.07	-0.4	-0.23	-0.03	0.17	0.76		
2013	-0.2	0.17	0.2	0.01	-0.47	-0.17	-0.03	0.12	0.79		
2014	-0.3	0.14	0.18	0	-0.38	-0.22	0.04	0.18	0.79		
Mean value, m	-0.16	0.18	0.15	-0.06	-0.35	-0.15	-0.19	0.17	0.74		
Standard deviation, s	0.12	0.04	0.18	0.15	0.13	0.15	0.25	0.07	0.12		
Signal/noise Ratio, SNR	1.36	4.33	0.79	0.41	2.6	1.03	0.75	2.44	6		
Sum of SNRs for the line	19.7					-					
Sum of significant SNRs	15.37										

Note: The variables considered are from the first block of Table 2. Source: Rosstat//http://www.gks.ru/.

If the signal-to-noise ratio is above the threshold value, then such a variable is considered nonrandom. Otherwise, the variable characterizes the noise component of the signal and does not participate in further consideration. In Figure 2 the choice of the direction of the eigenvectors for the first principal component is shown.

The criterion for choosing the direction of the eigenvectors will be the maximization of the signal level—the sum of the calculated SNR values for significant variables. The principal component thus obtained will be called the empirical principal component (EPC), the significant variables, as in factor analysis, will participate in further consideration, and the insignificant variables are ignored. Table 1 gives an example of the determination of an EPC from eight observations. The second, fifth, eighth and ninth variables turned out to be significant (the calculated SNR is above the 2.2 threshold) and the load values for these variables will be non-zero loads in this EPC.

In the algorithms for computing the integral characteristic from the PCA method (Sharpe, 2004; Nicoletti et al., 2000; Saltelli, 2007; Handbook, 2008), the concept of informativeness, which is traditional for the PCA, is used, which determines the number of principal components 1 used to calculate the integral characteristic.

$$\gamma_{\sigma} = \frac{\lambda_1 + \lambda_2 + \dots + \lambda_l}{\lambda_1 + \lambda_2 + \dots + \lambda_n} \ge \theta \tag{6}$$

However, the dimensionality of the feature space in the problems of computing the integral quality characteristic of a complex system is not too large, and there are no computational problems in determining the eigenvalues and vectors. A qualitative description of the structure of the system requires either all the principal components or a sufficiently large number of them. It may turn out that information valuable for a particular task is contained solely in last principal components. For example, when creating a digital terrain model that is based on digitized images, the desired contour is given with the eighth and ninth principal components, and the principal components 12 and 13 in the method "caterpillar" estify to the availability of periodicals with a fractional period in the analyzed data (Golyandina et al., 2008)

Approaches to estimating the number of principal components with respect to the required fraction of the explained variance are formally applicable, but implicitly they assume that there is no separation into "signal" and "noise", and any predetermined accuracy makes sense. When the data is divided into useful signal and noise, the specified accuracy becomes meaningless and it is required to redefine the notion of informativeness. Analogously to the variance of information according to (2), it is possible to determine SNR- informativeness for a selected number of empirical principal components N:

$$\gamma_{SNR} = \frac{S_{11} + S_{12} + \dots + S_{1N}}{S_{21} + S_{22} + \dots + S_{2N}} , \qquad (7)$$

where S_{lk} —is the sum of the SNR values of the operating variables of the k-th EPC, S_{2k} —is the sum of the SNRs of all the variables of the k-th EPC. This value will be a posteriori estimate (from above) of SNR-informativeness. Unlike dispersion information, SNR-informativeness cannot reach 100% according to the logic of construction. The informativeness of the chosen system of attributes is determined by the variance and SNR-informativeness:

$$\gamma = \gamma_{\sigma} \cdot \gamma_{SNR} \tag{8}$$

The number of selectable EPCs involved in calculating the composite index maximizes the total informative value of the solution, determined both by the traditional cumulative dispersive information content and the accumulated SNR-informativeness characterizing the level of the EPC signal relative to the background level. Variance informativeness increases with increasing number of EPCs used, and SNR-informativeness decreases, since the younger components carry more noise. SNR - the informativeness of the EPC, defined in Table 1, is 15.37 / 19.7 = 0.78. The algorithm that realizes the foregoing is given (Zhgun, 2017a)

Figure 3 shows the definition of the number of empirical principal components for calculating the complex index of the system, described by 85 statistical indicators. The maximum information

content of the system $\gamma = 60.52\%$ is achieved, if nine EPCs are used for calculating the complex index of the system. If we use more EPCs to calculate the complex index, then the information content and the reliability of the computed characteristics is reduced.



Figure 3. Determination of the number of empirical principal components for calculating the complex index of the system.

3. Investigation of the quality of integrated indicators

According to the proposed algorithm, the complex indexes of quality of life for constituent entities for 2007–2016 are calculated. It is worth noting that the determined weighting factors characterize the structure of the system. Over the observed time interval, the structure should be constant. In Russia, there are changes in different areas, so a longer period should not be considered. Data for 2016 was the latest available at the time of this writing.

Variables from (Isakin, 2006) were used for the study (Table 2). All values of the variables are taken from the open directories of Rosstat. Among the listed variables, variables 1, 2, 5, 7, 9, 10, 12, 21, 22, 23 are related to the calculated characteristic by the monotonic increasing dependence, when the optimal *j*-th value is maximum. The remaining indicators, except for variable 27, the optimal index value is minimal. For variable 27 "net migration" x_j^{opt} value equals sample mean.

To obtain the weights of the second block, ten empirical EPCs were used. Used indicators and the resulting weight vectors for these blocks are presented in Table 2.

The number of EPCs for calculating the composite index should maximize the informational value of the obtained solution, defined by (7)—Table 3.

Variable number	Variable Name	Variable weight
	Block 1: Population's welfare	
1	Per capita GDP–living wage ratio, units	0.070
2	Per capita income purchasing power relative to living wage, %	0.780
3	Share of people with incomes below living wage, %	0.850
4	The ratio of average income of the richest 20% to the poorest 20% ($\mathbb{R}\P$ 20)	-0.430
5	Number of cars per 1 000 people	1.730
6	Share of families on waiting lists for housing, %	1.000
7	Total area of housing resources per resident (m2/10 people)	1.330
8	Share of dilapidated housing, %	0.420
9	Public road density (km/10,000 km2)	0.530
	Block 2: Population quality	
10	Life expectancy at birth, years	1.402
11	Mortality rate, infant (per 1,000 live births)	0.166
12	Population growth rate, per 1,000 people	0.667
13	Deaths caused by communicable, parasitic diseases and TB per 100,000 people	1.675
14	Deaths caused by neoplasms per 100,000 people	0.485
15	Deaths caused by cardiovascular diseases per 100,000 people	0.474
16	Deaths caused by respiratory diseases per 100,000 people	1.097
17	Deaths caused by digestive system diseases per 100,000 people	0.842
18	Incidence of injuries, intoxication and other external causes per 100,000 people	1.240
19	Number of disabled people per 1,000 people	0.000
20	Incidence of congenital anomalies per 1,000 people	0.381
21	Specialists with higher education employed in economy, %	1.344
22 23	Labor force productivity (GRP per average annual number of employed in economy, thousand rubles/person) Graduates from higher and vocational educational institutions per 1,000 people	-0.406 0.048
	Block 3: Social quality	
24	Unemployment, %	0.189
25	Employers engaged in harmful and hazardous working conditions in the	0.467
26	Number of employees injured at work resulting in death or loss of earning capacity for 1 or more days per 1.000employees	1.030
27	Net migration per 10,000 people	0.136
28	Intentional homicides per 100,000 people	0.980
29	Incidence of intentional infliction of grievous bodily harm per 100,000 people	0.933
30	Incidence of rape per 100,000 people	0.955
31	Incidence of robbery and theft per 100,000 people	0.833
32	Incidence of larceny or embezzlement per 100,000 people	0.486
33	Number of registered with drug and substance abuse per 100,000 people	0.761
34	Number of registered with alcohol abuse per 100,000 people	0.811
35	Number of infected with TB per 100,000 people	0.806
36	Mortality from external causes per 100,000 people	0.000
37	Number of people with mental disorders per 100,000 people	0.861

Table 2. V	Weights	of the	variables	of the	first,	second	and	third	blocks.
------------	---------	--------	-----------	--------	--------	--------	-----	-------	---------

	Number of EPCs													
	1	2	3	4	5	6	7	8	9	<u>10</u>	11	12	13	14
SNR informativeness	0.98	0.95	0.33	0.33	0.23	0.37	0.38	0.27	0.20	0.26	0.25	0.38	0.28	0.16
Variance informativeness	0.32	0.53	0.63	0.71	0.78	0.84	0.88	0.92	0.95	0.97	0.98	0.99	1.00	1.00
Total informativeness	0.31	0.51	0.58	0.63	0.67	0.70	0.72	0.728	0.729	<u>0.730</u>	0.726	0.72	0.70	0.69

Table 3. Determination of informativeness value of the complex index of Blok 2: Population quality.

Once the three intermediate composite indices had been constructed, they were aggregated by allocating a weight to each one of them equal to the proportion to the SNR-informativeness (to the signal level). The block weights are defined in Table 4.

Table 4. Block weights, determined by the SNR-informativeness.

Block	1	2	3
Sum of significant SNRs	172.2	348.1	251.7
Block's weight	0.243	0.464	0.294

The results are shown in Table 5. The values of the complex indexes obtained vary from 1 to 100 (in 2007). The Republic of Tyva has the minimum unit value; Ingushetia has the maximum value of 100.

The change in the quality of life for some of Russia's constituent entities is shown in Figure 4. It is interesting to trace the reflection of recent political events on the values of the computed composite index. The impact of the events of 2014 most affected the quality of life in the maritime regions—Kaliningrad and Murmansk—and the financial capitals of Moscow and St. Petersburg. In other regions, the quality of life index is less subject to fluctuations when the political situation changes. Nevertheless, the gap in the quality of life with the leaders is almost not reduced.

Region of the RF					Comple	ex index				
	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Belgorod reg.	73.6	76.8	80.8	83.2	85.1	85.8	87.3	87.3	90.4	91.6
Bryansk reg.	46.0	47.7	53.0	55.3	59.6	62.2	63.5	65.2	68.0	71.1
Vladimir reg	44.9	51.2	55.8	59.0	61.3	64.2	65.7	63.6	65.7	69.2
Voronezh reg.	61.3	62.1	64.2	67.3	72.8	77.2	78.3	80.9	83.8	85.7
Ivanovo reg.	37.2	43.1	46.4	47.0	52.8	55.8	59.3	60.8	69.9	72.9
Kaluga reg.	52.2	55.3	59.2	62.0	68.0	70.3	71.8	74.2	75.0	78.5
Kostroma reg.	50.0	51.7	55.7	58.5	60.6	67.1	70.4	72.2	73.0	74.8
Kursk reg.	50.5	54.3	59.2	64.1	65.7	70.5	73.9	74.2	79.5	81.2
Lipetsk reg.	62.0	60.5	65.7	69.5	74.2	75.1	78.5	82.5	82.3	85.4
Moscow reg.	61.6	67.4	72.9	77.4	82.2	84.8	86.9	89.1	93.8	96.0
Oryol reg.	53.8	59.0	61.6	62.8	69.2	70.5	73.0	73.4	75.5	77.9
Ryazan reg.	53.7	59.2	63.7	66.2	71.3	73.4	77.3	76.8	79.3	78.0
Smolensk reg.	39.1	44.4	46.9	51.7	58.9	61.9	66.8	69.3	70.7	73.7
Tambov reg.	54.4	58.0	61.6	64.3	67.8	69.2	74.4	73.3	78.0	81.5
Tver reg.	34.6	38.0	42.9	45.9	53.2	58.8	63.5	63.6	63.6	66.5
Tula reg.	38.7	44.2	51.3	56.7	63.0	64.5	64.4	67.8	69.4	72.9
Yaroslavl reg.	49.5	55.3	60.3	62.5	67.5	68.1	70.0	76.8	75.2	79.4
Moscow	94.6	94.7	98.4	100.9	104.3	104.7	106.9	102.9	105.5	104.7
Rep. of Karelia	44.0	48.0	52.5	53.5	60.6	63.7	68.9	67.6	67.4	71.7
Rep. of Komi	48.0	50.9	53.5	55.9	61.5	64.8	68.9	70.0	70.4	69.7
Arkhangelsk reg.	50.9	54.5	59.7	61.2	65.0	70.0	70.9	72.9	71.3	74.6
Nenets Aut. Okrug	40.0	49.2	52.8	53.2	60.0	66.7	63.1	73.9	73.7	80.4
Vologda reg.	49.2	50.2	52.4	52.8	59.5	64.5	68.8	73.0	74.6	76.7
Kaliningrad reg.	52.5	53.0	58.5	65.1	72.4	77.3	77.9	75.7	79.0	83.6
Leningrad reg	41.0	46.5	54.3	58.9	64.4	68.0	68.9	69.9	72.3	75.1
Murmansk reg.	68.0	68.7	73.7	77.2	78.6	81.4	86.2	82.9	84.0	86.9
Novgorod reg.	30.8	32.0	38.6	41.4	47.8	54.4	56.0	58.3	58.0	61.7
Pskov reg.	40.8	39.1	43.8	46.4	53.6	57.0	62.9	68.1	66.8	69.7
Saint Petersburg	70.4	73.2	78.8	83.9	87.8	91.1	93.9	91.4	92.0	94.3
Rep. of Adygeya	57.9	60.9	66.2	69.8	72.3	76.7	79.4	81.1	79.9	82.2
Rep. of Kalmykia	58.7	64.0	69.0	64.9	72.2	75.4	78.0	79.8	79.8	83.0
Crimea	-	-	-	-	-	-	-	33.4	34.1	33.8
Krasnodar reg.	65.3	69.0	77.9	75.3	76.4	80.5	85.3	85.5	87.9	91.9
Astrakhan reg.	50.1	53.9	59.2	59.7	65.2	68.7	73.9	75.4	77.9	82.3
Volgograd reg.	58.3	59.9	62.3	64.3	66.9	68.4	71.6	75.7	79.9	85.0
Rostov reg.	60.1	63.0	68.8	69.9	72.0	76.9	79.9	81.2	82.3	83.5
Sevastopol	-	-	-	-	-	-	-	49.8	48.9	48.8

Table 5. The complex indexes of quality of life and ratings for Russia's constituent entities for 2007–2016.

Continued on next page

Region of the RF					Comple	x index				<u> </u>
	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Rep. of Dagestan	82.9	85.0	85.6	88.4	89.4	89.9	91.0	94.2	92.6	98.2
Rep. of Ingushetia	100.0	104.9	99.9	97.6	96.5	96.9	98.1	92.3	94.8	94.6
Kabardino-Balkar Rep.	76.4	82.8	80.6	80.8	82.4	85.4	87.5	86.1	87.9	89.5
Karachay-Cherkessia	69.6	73.2	76.1	79.3	80.6	82.8	86.9	90.2	91.1	96.1
Rep. of North Ossetia	84.9	82.6	82.0	88.5	91.3	93.5	93.4	95.9	94.6	96.2
Chechen Rep.	87.7	93.0	94.2	94.3	93.0	89.5	94.3	94.0	96.2	98.1
Stavropol reg.	69.3	71.5	73.6	79.3	81.8	85.6	89.9	88.6	93.6	94.6
Rep. of Bashkortostan	63.2	66.4	69.0	69.2	72.7	74.2	76.4	74.1	74.3	77.9
Rep. Mary El	38.6	40.0	46.8	48.8	52.1	56.2	60.2	65.7	66.2	71.2
Rep. Of Mordovia	58.5	62.3	65.3	67.3	69.5	72.2	73.9	76.9	79.8	83.1
Rep. of Tatarstan	67.1	69.4	72.9	75.1	77.7	82.0	84.3	83.7	86.4	90.8
Udmurt Rep.	43.1	47.0	52.0	55.3	59.1	64.9	67.9	63.5	68.5	72.0
Chuvash Rep	43.8	46.9	50.9	52.4	57.8	61.3	63.1	65.4	67.1	71.2
Perm reg.	35.5	39.3	42.5	45.2	50.0	54.7	59.6	61.5	62.2	64.6
Kirov reg.	44.8	47.6	51.8	54.9	60.0	63.5	66.6	70.8	70.9	74.8
Nizhny Novgorod reg.	46.0	50.6	55.9	59.1	65.4	68.8	72.3	74.2	77.4	82.8
Orenburg reg.	48.1	49.9	54.8	58.1	60.2	62.1	64.2	64.3	68.1	73.9
Penza reg.	61.0	62.8	67.3	70.0	75.2	79.5	79.7	77.8	81.5	83.0
Samara reg.	56.9	56.9	59.1	60.3	64.0	67.4	66.5	72.8	74.2	77.0
Saratov reg.	57.8	61.1	63.6	65.6	68.2	72.0	74.9	77.0	77.9	80.9
Ulyanovsk reg.	53.1	57.5	60.8	60.7	64.6	70.3	69.1	69.3	69.2	71.4
Kurgan reg.	40.6	40.8	45.3	46.5	51.4	53.0	55.4	57.2	55.5	59.4
Sverdlovsk reg	51.5	53.5	56.4	59.6	61.7	64.5	68.7	70.8	69.7	71.8
Tyumen reg.	59.1	62.6	65.9	68.4	74.0	78.7	69.8	70.1	74.5	78.6
Khanty-Mansiysk. A.O.	71.4	73.5	74.6	75.2	80.7	85.9	87.5	87.7	92.0	96.0
Yamalo-Nenets A.O.	75.6	78.7	83.0	82.1	83.6	86.5	92.2	94.4	93.0	94.3
Chelyabinsk reg.	50.2	52.6	56.9	57.7	60.4	62.5	65.9	64.8	67.5	70.1
Rep. of Altai	39.6	38.6	39.9	43.9	49.9	54.1	53.1	54.9	58.4	67.9
Rep. of Buryatia	38.6	38.5	43.0	47.1	50.1	54.4	61.7	59.4	65.4	70.0
Resp. Tuva	1.0	5.1	8.2	12.0	20.6	20.5	25.6	31.4	29.2	29.2
Resp.Khakassia	45.0	46.4	43.5	50.9	51.9	52.4	58.2	59.0	59.7	63.8
Altai reg.	40.3	44.4	47.6	48.0	50.9	53.1	57.8	57.2	62.4	65.7
Transbaikal reg.	36.2	37.6	41.0	42.9	49.4	52.0	53.8	59.7	56.1	60.9
Krasnoyarsk reg.	48.4	50.4	52.5	54.9	58.8	62.1	63.6	64.5	65.5	68.0
Irkutsk reg.	32.2	33.5	34.0	35.2	38.6	43.0	44.7	46.3	48.8	52.5
Kemerovo reg	28.0	32.4	34.7	35.7	41.3	45.6	51.2	52.6	50.7	52.2
Novosibirsk reg.	46.0	51.0	56.7	59.6	64.3	66.5	68.7	69.7	70.8	73.1
Omsk reg.	46.8	52.2	55.9	59.7	65.6	66.8	64.8	64.6	64.1	70.1
Tomsk reg.	59.1	58.8	60.1	63.4	68.2	71.8	74.0	72.8	74.6	75.3
Yakutia	62.3	58.2	60.4	62.0	65.5	70.5	75.3	79.1	78.4	80.2

55

National Accounting Review

Continued on next page

Volume 1, Issue 1, 42-61.

Region of the RF	Complex index													
	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016				
Kamchatka Krai	57.1	61.0	64.7	64.1	66.0	71.2	75.6	76.1	78.0	77.6				
Primorsky Krai	34.6	39.0	45.3	48.4	50.5	56.4	59.8	63.4	64.0	68.6				
Khabarovsk reg.	37.3	40.3	46.6	50.3	53.4	60.1	64.2	63.4	66.6	69.2				
Amur reg.	34.5	32.2	33.8	36.8	39.3	46.9	55.5	60.6	60.1	63.0				
Magadan reg.	35.7	32.6	44.3	49.1	57.4	57.9	65.6	66.4	70.6	72.8				
Sakhalin reg.	34.7	36.4	38.2	42.3	46.6	50.1	56.6	59.5	57.6	61.3				
Jewish Authon. reg.	16.6	20.9	24.9	23.9	25.7	33.1	37.6	37.5	36.3	37.8				
Chukchi Aut.county	29.9	34.5	30.4	30.9	44.1	45.7	43.0	48.1	52.1	49.6				



Figure 4. The change in the quality of life for some of Russia's constituent entities for 2007–2016.

Consistently, the highest complex indices of quality of life for the entire observation period are shown by the constituent entities of the North Caucasus Federal District. Of the 37 indicators, when calculating the complex indices of quality of life, 20 reflect the human physiological well-being and explain (along with the characteristics of national statistics) high indicators for the national republics of the North Caucasus (North Ossetia, Ingushetia, Chechnya, Dagestan, etc.). In these constituent entities of the Russian Federation, it is less likely to get sick, die and undergo criminal violence. The population of these entities is not excessively involved in industry. Life expectancy in Ingushetia is the highest in Russia and exceeds the Novgorod region by 10 years.

4. Estimating the stability of a composite index

Researchers from Russia who compute integrated indicators for different observations complete their research at the stage of calculating and analyzing the ratings obtained, without analyzing the quality problems of the results, in particular robustness and sensitivity (Ajvazjan, 2003; Ajvazjan et al., 2009; Gajdamak and Hohlov, 2009)

The main area of application of composite indexes is the rating of objects. It is the position of the object relative to other objects that is the basis for attracting public attention, and for making political

decisions. If small changes in the input data (values of scores or weights) when calculating the composite index fundamentally change the ranking of objects made on the basis of the calculated complex indices, then such an integral indicator cannot be considered reliable. A necessary attribute of the reliability of a composite index is stability relative to perturbations of the original data. In particular, the consequence of this is a slight (on average) change in the rating of objects for different times.

The researchers, who calculate the complex indices for various observations, complete their research at the stage of calculating and analyzing the obtained ratings without analyzing the problems of the quality of the results, in particular, robustness and sensitivity.

It should be noted that the ranks of the 25 member countries of the European Union for 2009–2011, based on the values of the HDI (when calculating the value of the HDI by linear convolution with equal weights) give an average change in the rating for the year of 7.7% (Human Development Reports, 1990–2014). The change in the ratings of more than 15% makes up 14% of cases; more than 30% is 2% of cases. In other words, values are typical in the case of linear convolution with constant weights.

In a study that determines complex indexes of quality of life for the Russian Federation according to the methodology proposed by (Ajvazjan, 2003), ratings for constituent entities were exhibited according to two methods: the weights of the indicators for each year in the linear convolution were determined in one case by experts, in the other by the method proposed by the author. Cases when the rating change amounted to more than 15% are the following: 18 in the case of expert weights, one in the case of calculated weights. The average change in the rating for the first method is 3.6%, for the second method it is 1.6%. Hence, the proposed methodology showed a brilliant quality of the constructed composite index.

However, with careful application of this technique by other authors, the results turn out to be quite different. For example, in the paper (Gajdamak and Hohlov, 2009) the complex indices of the quality of life for municipal formations of the Tyumen region (2005–2008) were calculated. In this case, a change in the rating of more than 15% of the maximum is 48.9% of the total number of cases. In 17.9% of cases, this value exceeds 30%. The average rating change is 16.9%. In the paper (Ajvazjan et al., 2009) values of ratings for municipal formations of the Samara area, calculated by the same method, are given. 45% of objects significantly changed their position in the rating (by 15% or more from the maximum possible rating change). In 21.6% of cases the rating change exceeds 30%. The average rating change is 16.9%.

If small changes in input data while calculating the composite index dramatically change the ranking of objects, then such a composite index cannot be considered reliable. A necessary attribute to the reliability of a composite index is stability relative to perturbations of the original data. In particular, the consequence of this is a slight (on average) change in the rating of objects for different measurements. Schemes for determining weights using factor analysis or the principal component analysis do not have this property.

Let $R_t = (r_{t1}, r_{t2}, ..., r_{tm})$ —the ratings of *m* objects for the moment *t*. Known values of the rating sets R_t , R_{t+1} , ..., R_T for the moments *t*, t + 1, ..., *T*. R_t —can be considered as a random value uniformly distributed on the interval [1, m]. Numerical characteristics R_t correspond to the numerical characteristics of uniform distribution.

Values of ratings for object i at successive instants of time r_{ti} , $r_{t+1,i}$, $r_{t+2,i}$ represent a numerical implementation of a complex functional dependence, a formal description of which is not possible. Changes in ratings over time are dictated mainly by this dependence and to a lesser extent by random factors. The degree of linear connection between the sets of R_i , R_{t+1} ratings is high, the Pearson and

Spearman correlation coefficients are close to one and do not allow the making of unambiguous conclusions about the quality of the ratings. We propose a method for estimating the stability of a composite index.

If the random variables R_t , R_{t+1} are independent, then the variance of the difference of independent random variables is equal to the sum of the variances $D(\Delta_t)=D(R_{t+1}-R_t)=D(R_{t+1})+D(R_t)$.

Because X is even distribution on [a,b], then the variance $D(X) = (b-a)^2/12$. Therefore, $D(R_t) = D(R_{t+1}) = (m-1)^2/12$ and

$$D(\Delta_t) = D(R_{t+1} - R_t) = D(R_{t+1}) + D(R_t) = \frac{(m-1)^2}{6}$$
(9)

It is possible to estimate the stability of the complex indexes by estimating the randomness of the difference in the exposed ratings of R_t . Such an estimate of randomness is the fraction of the variance of the realization of the quantity R_t with respect to the variance (9), which reaches a maximum if the ratings R_t , R_{t+1} for two consecutive moments of time are absolutely independent of each other and are completely random. Since randomness is not the main reason for rating change, this share should be small.

Complex index Source		Period	Weights,	Num	ıber	Evaluation	
			source	Variables,	Objects,		
				n	m		
HDI of the EU countries	Human Development Reports, 1990–2014	2009– 2011	expert	5	25	6.9	
Quality of life of the Russian Federation	Ajvazjan, 2003	1997, 1999	expert	9	79	6.3	
Quality of life of the Russian Federation	Ajvazjan, 2003	1997, 2000	expert	9	79	1.2	
Quality of life in the Samara Region	Ajvazjan et al., 2009	2002– 2004	PCA	11	37	31.5	
Quality of Life of the Tyumen Region	Gajdamak and Hohlov, 2009	2005– 2008	PCA	17	26	22.7	
Quality of life of the Russian Federation	Zhgun, 2017a	2007– 2014	PCA	37	83	1.7	
Quality of life of the Russian Federation	(*)	2007– 2016	PCA	37	85	1.6	
SSI. Human Wellbeing	http://www.ssfindex.com/data- all-countries/	2006– 2016	expert	9	154	1	
SSI. Environmental Wellbeing	http://www.ssfindex.com/data- all-countries/	2006– 2016	expert	8	154	3.5	
SSI. Economic Wellbeing	http://www.ssfindex.com/data- all-countries/	2006– 2016	expert	5	154	4.5	

Table 6. Comparison of quality indicators of complex indexes.

Note: Integral indicators (*) are calculated by the author's method in this paper.

Table 6 shows the results of assessing the quality of different complex indexes. As a conditional unit of sustainability, we can consider an estimate of the integral index of the HDI, which is about 7% of the possible randomness.

The magnitude of the estimate, comparable to this value, will indicate a relatively good stability of the integral indicator with respect to the input data. Values that significantly exceed this value characterize the instability of the complex indexes and, consequently, it's of poor quality. In the table below, these are studies Gajdamak and Hohlov (2009) and (Ajvazjan et al., 2009), where the method Ajvazjan (2003) is implemented. The author's method shows good stability.

5. Conclusion

In this paper we consider a solution to the problem of constructing latent complex indexes for the change of a systems quality on the basis of registered measurements for a number of observations in the absence of training. An analysis of the stability of such a solution is provided.

The construction of a system's complex indexes can be considered as a task of extracting the useful signal from a background noise. The signal in this case is the weights of the linear convolution of the indicators. Determined weights should reflect the structure of the system being evaluated. Successful application of the PCA to describe the structure of various system types suggests that the method will also give adequate results for describing social systems. However, the principal component method and factor analysis (even with fixed methods for extracting factors and the method of rotation) determine the structure of the principal components and the main factors for differently. Hence, the method of determining weights using multidimensional analysis cannot be used to compare the characteristics of objects in dynamics.

The reason for this may be the presence of irremovable errors in the data used. Even a small perturbation in the source data can cause a significant change in the weights when using multidimensional analysis methods. As a way out of this situation, a modification of the principal component method is proposed, taking into account the presence of errors in the data used. The algorithm uses a new approach to choosing the number of principal components, to determine the significant loads of the principal components describing the structure of the system, to determining the weights of the considered subsystems, and to determining the informativeness of the selected principal components based on the signal-to-noise ratio parameter.

The solution of the problem requires a detailed understanding of the influence of the errors in the data used on the calculated characteristics. The consequence of stability is on average a slight change (increment) of the ratings of objects for different measurements. This increment can be a posteriori estimated by a sequence of observations using the proposed dispersion criterion. Estimates of different complex indexes' stability by the variance criterion are given. Complex indexes calculated using the proposed modification of the principal component method, which takes into account the presence of errors in the data used, show good resistance to changes in input data.

Conflict of interest

The author declares no conflict of interest in this paper.

References

- Ajvazjan SA (2003) Empirical analysis of synthetic categories of life quality of the population. *Econ Math Methods* 3: 19–53.
- Ajvazjan SA, Stepanov VS, Kozlova MI (2009) Measurement of synthetic categories of quality of life of the population of the region and identify key areas of improvement in socio-economic policy (in the Samara region and its municipalities an example). *Appl Econom* 3: 18–84.
- Bandura R, del Campo CM (2006) Indices of National Performance: A Survey. Office of Development Studies, United Nations Development Programme, New York 91.
- Bandura R (2008) A Survey of Composite Indices Measuring Country Performance: 2008 Update. Office of Development Studies United Nations Development Programme, New York 96.
- Bandura R (2011) Composite Indicators and Rankings: Inventory 2011. Working Paper.
- Foa R, Tanner JC (2012) Methodology of the Indices of Social Development (ISD Working Paper Series from International Institute of Social Studies of Erasmus University Rotterdam (ISS), The Hague) 4: 1–66.
- Becker W, Paruolo P, Saisana M, et al. (2017) Weights and Importance in Composite Indicators: Mind the Gap, In: Ghanem R., Higdon D., Owhadi H. (eds) *Handbook of Uncertainty*, Quantification Springer International Publishing, 1187–1216.
- Becker W, Saisana M, Paruolo P, et al. (2017) Weights and Importance in Composite Indicators: Closing the Gap. *Ecol Indic* 80: 12–22.
- Gajdamak IV, Hohlov AG (2009) Modeling of integral indicators of quality of life of the population of the South of the Tyumen region. *Vestnik TjumenSU* 6: 176–186.
- Golyandina NE, Usevich KD, Florinsky IV (2008) The Analysis of a singular range for a filtration of digital models. *Geod cartography* 5: 21–28.
- Hightower WL (1978) Development of an index of health utilizing factor analysis. *Medical Care* 16: 245–255.
- United Nations Development Programme (1990–2015) Human Development Reports.
- Joint Research Centre-European Commission (2008) Handbook on Constructing Composite Indicators: Methodology and User Guide, OECD Publication, Paris CEDEX, 16: 1–162.
- Isakin MA (2006) Modification of a Method to *k*-Means with Unknown Number of Classes. *Appl Econom* 4: 62–70.
- Lindman C, Sellin J (2011) Measuring Human Development: The Use of Principal Component Analysis in Creating an Environmental Index. University essay from Uppsala universitet, Uppsala, 45.
- Mazziotta M, Pareto A (2016) On The Construction Of Composite Indices By Principal Components Analysis. *Rivista Italiana di Economia Demografia e Statistica* 70: 103–109.
- Molchanova EV, Kruchek MM, Kibisova ZS (2014) Building of the rating assessments of the Russian Federation subjects by the blocks of socio-economic indicators. *Econ Soc Changes Facts Trends Forecast* 3: 196–208.
- McKenzie DJ (2005) Measuring Inequality with Asset Indicators. J Pop Econ 18: 229-260.
- Nardo M, Saisana M, Saltelli A, et al. (2005) Tools for composite indicators building. Joint Research Centre European Commission. *Institute for the Protection and Security of the Citizen Econometrics and Statistical Support to Antifraud Unit, Ispra,* 1–134.

- Nicoletti G, Scarpetta S, Boylaud O (2000) Summary indicators of product market regulation with an extension to employment protection legislation. Economics department working papers ECO/WKP (99)18) 226.
- Saltelli A, Munda G, Nardo M (2006) From Complexity to Multidimensionality: the Role of Composite Indicators for Advocacy of EU Reform. LI *Rev Bus Econ Lit* 51: 221–235.
- Sharpe A (2004) Literature Review of Frameworks for Macro-indicators, Centre for the Study of Living Standards, Ottawa, CAN.
- Saltelli A (2007) Composite indicators between analysis and advocacy. Soc Indic Res 81: 65-77.
- Somarriba N, Pena B (2009) Synthetic Indicators of Quality of Life in Europe. Soc Indic Res 94: 115–133.
- Paruolo P, Saisana M, Saltelli A (2013) Ratings and Rankings: voodoo or science?. *J Royal Stat Soc* A 176: 609–634.
- Tarantola S, Saisana M, Saltelli A (2002) A Internal Market Index 2002: Technical details of the methodology. JRC Eur Comm Ispra (VA) Italy.
- Vyas S, Kumaranayake L (2006) Constructing socio-economic status indices: how to use principal components analysis, Oxford University Press, London.
- Woodward PM (1955) Probability and Information Theory with Applications to Radar, London, Pergamon Press, 128.
- Zhgun TV (2013) Calculation of an Integrated Indicator of Efficiency of Functioning of Dynamic System on the Example of an Integrated Assessment of Demographic Development of Municipalities of the Novgorod Region. *Vestnik NovSU* 75: 11–16.
- Zhgun TV (2014) Constructing the integral characteristic of the system quality changes on the basis of statistical data as a solution to the problem of signal allocation under conditions of prior uncertainty. *Vestnik NovSU* 81: 10–16.
- Zhgun TV (2017a) An algorithm for constructing integral quality indicator of complex systems for a sequence of observations. *Bul South Ural St Univ* 1: 5–25.
- Zhgun TV (2017b) Building an integral measure of the quality of life of constituent entities of the Russian Federation using the principal component analysis. *Econ Soc Changes Facts Trends Forecast* 10: 214–235.



© 2019 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0)