

Computational method for risk assessment of regional socio-economic development

Oleg Pursky^{1,*}, Tatiana Dubovyk^{1,**}, Iryna Buchatska^{1,***}, Iryna Lutsenko^{2,****}, and Hanna Danylchuk^{3,†}

¹Kyiv National University of Trade and Economics, 19 Kioto Str., Kyiv, 02156, Ukraine

²National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, 37 Peremohy Ave., Kyiv, 03056, Ukraine

³The Bohdan Khmelnytsky National University of Cherkasy, 81 Shevchenko Blvd., Cherkasy, 18031, Ukraine

Abstract. In this study, we present the computational method for risk assessment of the socio-economic development of regions. An attempt has been made to develop a method for the determination of integral risk indicators of socio-economic development based on the joint use of the methods of factor analysis and expert evaluation. This approach has increased the reliability of the calculations and made it possible to analyze the influence of socio-economic indicators on the risk level of socio-economic development. The integral risk indicator shows the effect of the inconsistency in the level of factor provision on the socio-economic development of the j -th region (district) in comparison with the general situation in the country (regions). The closer the value of integral risk indicator is to 1, the higher the level of risk in this region. Using Kyiv region districts as an example, the process of risk assessment for regional socio-economic development has been considered. The results obtained in this investigation demonstrate that the presented computational method solves the problem of formalization of risk assessment for the socio-economic development of regions.

1 Introduction

The market today is functioning in a turbulent environment facing continuous change because of hyper-competition, changing demands of customers, regulatory changes, and technological advancement [1]. Modern world economic conditions, economic globalization, acceleration of market development processes, information technologies, socio-political factors require public administration new approaches to the formation of socio-economic strategies, development of adequate methodological solutions, and tools in the field of governance, especially it concerns socio-economic development management of regions [2, 3]. Using modern information technologies and new electronic communication channels significantly reduce costs related to organization and support social activity and business, and the new possibilities allow re-designing socio-economic development strategy at any moment [4]. In connection with the globalization and the processes of post-industrial economy development, the effect of unpredictability appears in the change of socio-economic systems state due to the increasing influence of economic crises, suddenly emerging threats and risks [5–9]. One of the urgent problems in the risk analysis of socio-economic systems is the construction of adequate methods. This is due to the multidimensionality of socio-economic systems, the stochasticity of their behavior, as well as the

complex interaction between the elements of the systems [6, 8].

In socio-economic studies, to improve the reliability of the procedure for assessment of socio-economic development using mechanism for determining the integral indicators based on factor analysis, taking into account consider the knowledge and experience of experts [10]. Accordingly, the aim of this study is to develop a reliable computational method for risk assessment of regional socio-economic development on the basis of the joint use of the methods of factor analysis and expert evaluation. This article poses and solves the problem of formalization of risk assessment for regional socio-economic development with the Kyiv region districts as an example.

2 Computational method of risk assessment

This section presents a method for determining the integral risk indicator of regional socio-economic development. The method is based on a model for determination of the socio-economic development integral indicator which is described in detail in [10]. In this model, methods of factor analysis and expert evaluations are used to determine integral indicators. To reducing the dimension of the feature space (socio-economic indicators), one of the methods of factor analysis is used [11–13], the principal component analysis (PCA) [14, 15]. Based on the reduced set of independent factors, a single integrated indicator is obtained, which combine all these factors in the

* e-mail: pursky_o@ukr.net

** e-mail: tatiana_dubovyk@i.ua

*** e-mail: i.buchatska@knute.edu.ua

**** e-mail: lutsenkois0802@gmail.com

† e-mail: abdaniichuk@gmail.com

best way [15, 16]. The main disadvantage of factor analysis methods is the reliability of the conclusions, in particular, in this model of determining the integral indicators [10], the weight of the factor is determined by the dispersion of initial indicators, which is not always reliable in socio-economic studies, since in this case the importance of indicators for the socio-economic system is not taken into account [10, 16, 17]. Therefore, within the framework of this model, in order to increase the reliability of the algorithm for determining the integral indicators based on factor analysis, expert evaluation procedures have been introduced in the mechanism of determining the weight of the factors [10, 18, 19]. In this case, the generalized weight of factors that takes into account both the weight of the factor, determined on the basis of expert evaluations, and the weight of the factor determined statistically, can be obtained as the weighted average of these two evaluations [10, 20]:

$$w_i = (\bar{q}_i + \bar{v}_i) \sqrt{\sum_{i=1}^n (\bar{q}_i + \bar{v}_i)}, \quad (1)$$

where $\bar{q}_i = q_i \sqrt{\sum_{i=1}^n q_i}$, $\bar{v}_i = v_i \sqrt{\sum_{i=1}^n v_i}$ are expert and statistical (factor analysis) weighted coefficients of the factor, respectively. Thus, the complex indicator of socio-economic development for j -th region is calculated as the sum of factors with the corresponding weighted average weight coefficients w_i [10]:

$$I_j = \sum_{i=1}^n w_i F_{ij} \quad (j = 1, 2, \dots, n), \quad (2)$$

where n is a number of factors; F_{ij} is the value of the i -th factor for the j -th object (region). Taking into account the proposed complex indicators of socio-economic development (2), the value of the integral risk indicator R in the region can be calculated using the following formula [21, 22]:

$$R_j = 1 - \frac{(I_j / I_{avg})}{m}, \quad (3)$$

where R_j is the integral risk assessment for socio-economic development of j -th region; I_{avg} is the numerical value of the complex indicator I in average in study regions; m is the number of factors. In this formula (3), the risk assessment is calculated taking into account the average value of the complex indicator I_{avg} for all regions and the number of factors. In this case, the risk R_j varies in the range from 0 to 1. The integral indicator R_j shows the effect of the inconsistency in the level of factor provision on the socio-economic development of the j -th region (district) in comparison with the general situation in the country (regions). The closer the value of R_j is to 1, the higher the level of risk in this region.

Figure 1 illustrates a general scheme of the developed computational method for determining the regional socio-economic development integral risk indicator. This method makes it possible to develop a procedure for au-

tomated data processing of socio-economic research and includes the following stages:

- 1 stage. The entering of values of indicators of socio-economic development and expert evaluations in the form of the matrix of indicators and the matrix of expert evaluations of indicators, followed by its normalization to a single scale of measurements.
- 2 stage. The calculation of the pairwise correlations matrix and determination of its eigenvalues and eigenvectors.
- 3 stage. Obtaining a matrix of factors by multiplying the normalized matrix of indicators and the matrix of eigenvectors, and normalization of factors and calculation of their variance.
- 4 stage. Determining the number of N factors included in the integral risk indicator (figure 1) on the basis of eigenvalues of the matrix of pairwise correlations of indicators and the given boundary value L dispersion of normalized indicators or in other words – a sampling of the minimum number of factors with maximal eigenvalues λ_i is made, the sum values of which are not less than nL [10].
- 5 stage. Determining the relative contribution $\%(F_i)$ of each of the N factors in the description of the total dispersion of all n indicators as the ratio of the eigenvalue λ_i of the factor F_i to the total dispersion of the features, which is also equal to n :
$$\sum_{i=1}^N \lambda_i \geq nL, \%(F_i) = \lambda_i / \sum \lambda_i = \lambda_i / n \quad (4)$$
- 6 stage. The determination of the experts competence and calculation of Kendall's coefficient of concordance for evaluation of the consistency of their conclusions.
- 7 stage. The determination of weighting coefficients of factors (1) included in the integrated indicators (2).
- 8 stage. The calculation of integral risk indicators of regional socio-economic development (3) and the visualization of the results of data processing.

3 Risk assessment for regional socio-economic development of the Kyiv region districts

In this section, we consider the process of risk assessment for regional socio-economic development in accordance with the developed computation method for integral risk indicators (figure 1) on the example of the Kyiv region districts. Listed in a single scale of measurements and normalized values of socio-economic indicators of districts are presented in table 1.

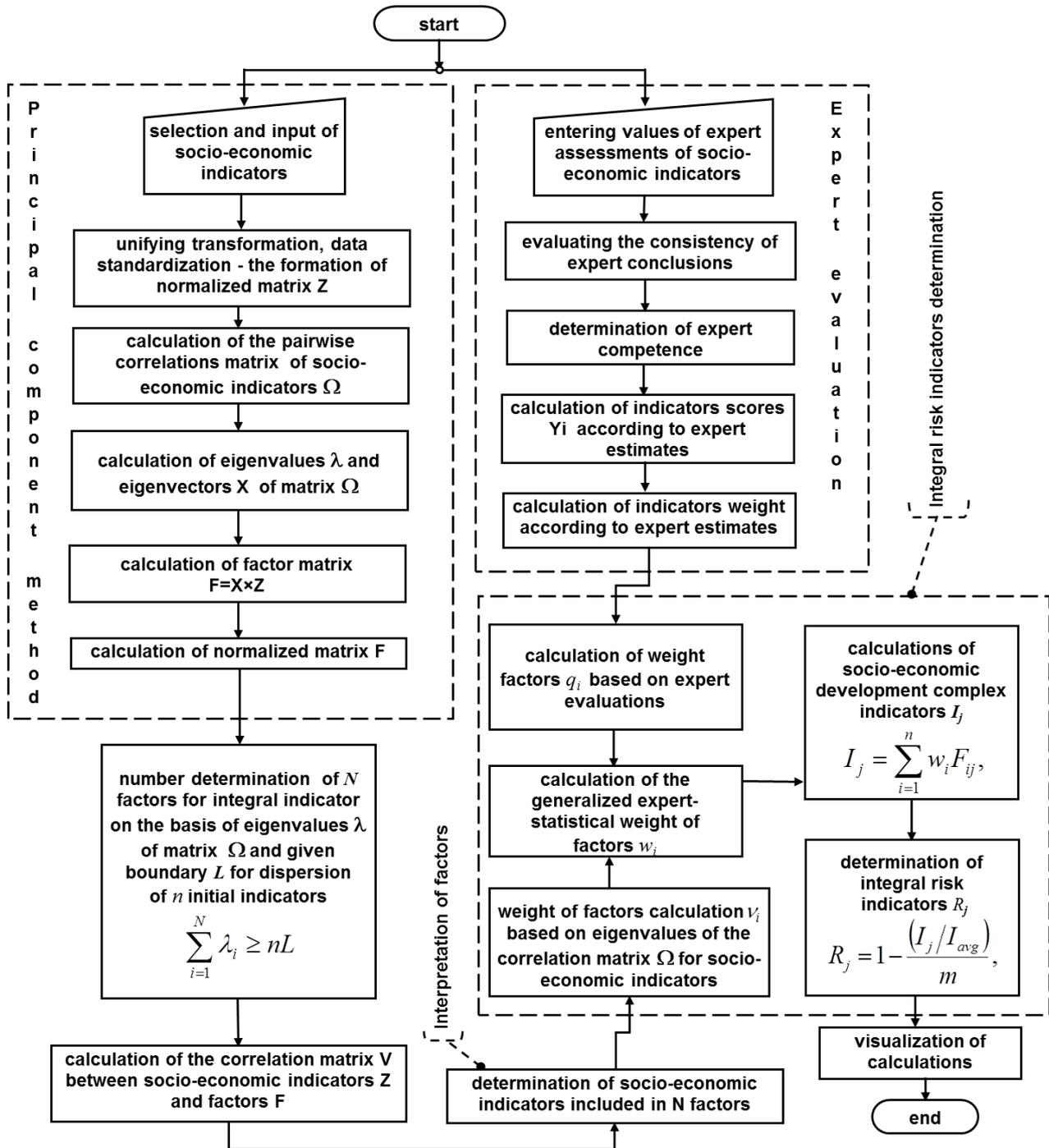


Figure 1. General scheme of the computational method for risk assessment of regional socio-economic development

According to the National State Statistics Service of Ukraine [23], one of the main indicators that characterize the level regional socio-economic development are (table 1):

- (P1) Number of cars per 1000 people;
- (P2) Services rendered per unit of population, UAH;
- (P3) Natural increase (reduction) of the population;
- (P4) Registered unemployment rate;

- (P5) Average monthly salary, UAH;
- (P6) Provision of housing by the population, m^2 per person;
- (P7) The ratio of m^2 of built housing to the population;
- (P8) Preschool establishments per unit of population;
- (P9) General educational institutions per unit of population;
- (P10) Number of crimes per 1000 people;

Table 1. Normalized values of socio-economic indicators

Districts	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
D1	-0,209	-0,054	-0,094	-0,129	0,077	-0,270	-0,063	-0,157	-0,171	0,065	-0,310
D2	-0,069	-0,075	-0,461	0,186	-0,005	0,294	-0,049	-0,027	-0,062	-0,240	-0,301
D3	0,044	-0,050	-0,088	0,300	0,253	-0,108	-0,068	-0,252	-0,298	-0,465	-0,236
D4	-0,077	0,016	0,005	0,174	0,223	-0,005	0,005	-0,142	-0,221	-0,287	-0,328
D5	0,468	-0,026	0,027	0,199	0,004	0,046	-0,031	0,111	0,058	-0,220	0,129
D6	0,036	0,013	-0,130	-0,129	0,255	-0,228	0,011	0,018	-0,053	0,097	-0,063
D7	-0,109	-0,071	-0,312	0,174	-0,173	-0,330	-0,001	-0,084	-0,044	-0,167	-0,312
D8	0,048	0,219	0,127	0,161	0,026	0,251	0,038	-0,211	0,021	-0,172	0,004
D9	-0,035	-0,086	0,131	-0,192	0,360	-0,142	-0,066	-0,043	-0,072	0,110	-0,090
D10	-0,170	-0,064	0,110	0,123	-0,219	0,054	-0,069	-0,219	-0,139	0,076	-0,175
D11	-0,125	-0,061	-0,131	0,035	-0,259	0,037	-0,064	-0,243	-0,014	0,066	0,087
D12	0,475	-0,067	-0,030	0,110	-0,058	-0,288	-0,049	0,071	0,020	-0,071	0,209
D13	-0,193	0,928	0,466	-0,205	-0,129	0,089	0,948	0,594	0,671	0,200	0,250
D14	-0,128	-0,052	-0,102	0,186	0,331	0,037	-0,037	0,014	0,027	-0,147	-0,071
D15	0,085	-0,041	-0,006	-0,381	-0,219	0,183	-0,061	0,110	0,229	0,306	0,181
D16	-0,170	0,073	0,110	-0,230	0,067	-0,176	0,162	0,198	0,256	0,254	0,313
D17	-0,003	-0,066	-0,292	-0,230	-0,166	-0,091	-0,069	0,250	0,129	0,156	0,089
D18	-0,189	-0,086	0,338	0,060	-0,285	0,106	-0,069	-0,222	-0,288	0,159	0,206
D19	0,043	-0,077	0,226	0,161	0,151	-0,228	-0,068	-0,088	-0,190	-0,002	0,101
D20	-0,166	-0,066	-0,169	-0,167	0,270	-0,031	-0,066	0,108	0,076	0,096	-0,118
D21	0,274	-0,080	0,189	-0,066	-0,126	0,114	-0,069	0,034	-0,057	0,097	-0,081
D22	0,090	-0,067	0,043	0,174	0,084	-0,168	-0,069	-0,097	-0,166	-0,132	-0,045
D23	0,389	-0,057	0,165	0,211	-0,044	0,285	-0,069	0,056	-0,035	-0,268	-0,003
D24	-0,182	-0,082	-0,088	-0,419	-0,067	0,465	-0,063	-0,154	0,188	0,231	0,275
D25	-0,125	-0,020	-0,036	-0,104	-0,352	0,106	-0,062	0,375	0,136	0,259	0,291

(P11) Emissions of pollutants.

It should be noted that the list of indicators, depending on the goals and objectives of the risk assessing, may change, thereby changing its emphasis. Thus, for the Kyiv region we have a matrix of initial socio-economic indicators in the size of 25×11 (25 districts of the Kyiv region: Baryshivsky (D1), Bilotserkiivsky (D2), Boguslavsky (D3), Boryspilsky (D4), Borodyansky (D5), Brovarsky (D6), Vasylkiivsky (D7), Vyshgorod (D8), Volodarsky (D9), Zguriivsky (D10), Ivankiivsky (D11), Kagarlytsky (D12), Kyiv-Sviatoshiivsky (D13), Makariivsky (D14), Myroniivsky (D15), Obukhovsky (D16), Pereiaslav-Khmelniivsky (D17), Polissya (D18), Rokytniivsky (D19), Skviivsky (D20), Stavyshechensky (D21), Tarashchansky (D22), Tetiivsky (D23), Fastiivsky (D24), Yahotiivsky (D25)).

On the basis of the normalized matrix of socio-economic indicators (table 1), the pairwise correlations matrix of indicators is dimensioned 11×11 . For the pairwise correlations matrix of indicators, we determine eigenvalues λ (table 2) and eigenvectors \mathbf{X} . The matrix of factors F is obtained by multiplying the normalized matrix of socio-economic indicators (table 1) into the matrix of the eigenvectors of the pairwise correlations matrix. The obtained factors are normalized. The normalized factor matrix is used to calculate the matrix of correlations between factors and indicators of socio-economic development that is required for the interpretation of factors.

On the basis of the calculated eigenvalues of the pairwise correlations matrix (table 2) and the given threshold

L of the dispersion for normalized socioeconomic indicators (table 1), the formula (4) determines the number of N factors in the integral risk indicator. In this case, the number of main components (factors) must be used, which exhaust at least 60-70% of the variance of the initial random variables. For example, at a given threshold of 0,6 from table 2 it is necessary to select N factors with maximal eigenvalues, the sum of values of which is not less than $0,6 \times 11 = 6,6$. The sum of the first three eigenvalues λ is 7,49 that is the integral index consists of the first three factors ($N = 3$) that explain approximately 68% (see formula 4) of the variance of the initial data (table 2). The calculate matrix of correlations between the normalized socio-economic indicators and the factors shows, which indicators are included in the given three factors (with the value of the variance of the indicators should not be less than the given limit value of 0,6). Table 3 shows the structure of factors: the coefficient of correlation between the indicators and factors in which they are included, statistical and expert weights coefficients and weighted average weight coefficient of factors. The first factor included the first four indicators: 1) the number of cars per 1000 people; 2) services rendered per unit of population; 3) natural increase (reduction) of the population; 4) the level of registered unemployment. The second factor included the eleventh indicator – emissions of pollutants. The third factor entered the seventh indicator – the ratio m^2 of the built housing to the population.

By multiplying the obtained factors by the corresponding weighted average weight coefficients of factors, by for-

Table 2. Eigenvalues of the pairwise correlation matrix of indicators

λ	4,565	1,733	1,193	0,931	0,764	0,587	0,495	0,364	0,244	0,097	0,022
$\%(F_i)$	41,51%	15,76%	10,85%	8,46%	6,95%	5,34%	4,50%	3,31%	2,22%	0,89%	0,20%
$\Sigma \%$	41,51%	57,27%	68,11%	76,58%	83,53%	88,87%	93,37%	96,69%	98,91%	99,80%	100,00%

Table 3. Structure of factors and their weight coefficients

Factors	Socio-economic indicators	Correlation coefficient	Statistical weight coefficient	Experts weight coefficient	Weighted average weight coefficient
F1	P1	0,912	0,456	0,269	0,363
	P2	0,873			
	P3	0,741			
	P4	0,647			
F2	P11	0,794	0,293	0,086	0,191
F3	P7	0,663	0,251	0,644	0,446

mula (2) we obtain the values of integral risk indicators according to formula (3) that allow ranking the regions in terms of their risk assessment of socio-economic development (figure 2). In table 4 calculations results of integral risk indicators R_j of social and economic development of Kyiv region districts are presented. It is also worth noting that for a better understanding of the calculation procedures should be carefully study the model that presented in [10].

4 Conclusions

The obtained results demonstrate that the presented computational method solves the problem of formalization of risk assessment for the socio-economic development and can be used to analyze and predict the socio-economic situation in the region. In the framework of the presented method by changing the values of socio-economic indicators, it is possible to analyze and model the socio-economic situation in the region, which undoubtedly provides management with valuable information on possible risks and directions of effective strategies for socio-economic development. The management receives not only an adequate assessment of the risk level of socio-economic development of the region, but also the opportunity to determine the immediate causes and consequences that shape the current and future socio-economic situation in the region. The results of modeling the process of assessing the level of socio-economic development of Kyiv region showed that the main advantages of the method of determining the integrated risk indicators are the possibility of studying correlations between socio-economic indicators, between indicators and factors, interpretation of factors, determining negative and positive characteristics of socio-economic situations in context of research objects. The proposed method for risk assessment makes it possible to implement a unified approach to data analysis and to ensure the efficiency of constructing integral risk indicators.

This study was supported by the Ministry of Education and Science of Ukraine, Project No. 0112U000635, "Development and

implementation of the modern information systems and technologies in the socio-economic activities".

References

- [1] A. Raman, S. Bharadwaj, J. Mukherjee, *International Journal Business Information Systems* **27** (2018)
- [2] S. Belew, J. Elad, *Starting an Online Business All-in-One for Dummies*, 3rd edn. (For Dummies, New Jersey, 2017)
- [3] O. Pursky, T. Dubovyk, I. Moroz, I. Buchatska, A. Savchuk, *CEUR Workshop Proceedings* **2422**, 15 (2019)
- [4] O.I. Pursky, B.V. Grynyuk, D.A. Shestopal, *Aktual'ni Problemy Ekonomiky = Actual Problems in Economics* p. 407 (2016)
- [5] D. Chernov, D. Sornette, *Critical risks of different economic sectors* (Springer, 2020)
- [6] K. Ptilakis, P. Franchin, B. Khazai, H. Wenzel, *SYNER-G: systemic seismic vulnerability and risk assessment of complex urban, utility, lifeline systems and critical facilities: methodology and applications*, Vol. 31 (Springer, 2014)
- [7] M. Machina, W.K. Viscusi, *Handbook of the Economics of Risk and Uncertainty* (North Holland, Amsterdam, 2014)
- [8] S. Ashmarina, M. Vochozka, *Sustainable Growth and Development of Economic Systems* (Springer, 2019)
- [9] H. Danylchuk, L. Kibalnyk, O. Serdiuk, *CEUR Workshop Proceedings* **2224**, 211 (2019)
- [10] O. Pursky, T. Dubovyk, I. Gamova, I. Buchatska, *CEUR Workshop Proceedings* **2393**, 919 (2019)
- [11] K. Adachi, *Matrix-based introduction to multivariate data analysis* (Springer, 2020)
- [12] K.A. Pituch, J.P. Stevens, *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS* (Routledge, 2015)
- [13] K. Vehkalahti, B.S. Everitt, *Multivariate analysis for the behavioral sciences* (CRC Press, 2018)

Table 4. The value of integral risk indicators of socio-economic development calculated for Kyiv region districts

Districts	R_j								
D9	0,805	D22	0,744	D15	0,684	D19	0,669	D8	0,567
D16	0,803	D20	0,722	D25	0,677	D3	0,661	D14	0,553
D23	0,758	D17	0,711	D2	0,675	D24	0,636	D6	0,529
D10	0,755	D18	0,709	D1	0,671	D7	0,620	D13	0,502
D21	0,752	D12	0,704	D5	0,670	D11	0,608	D4	0,483

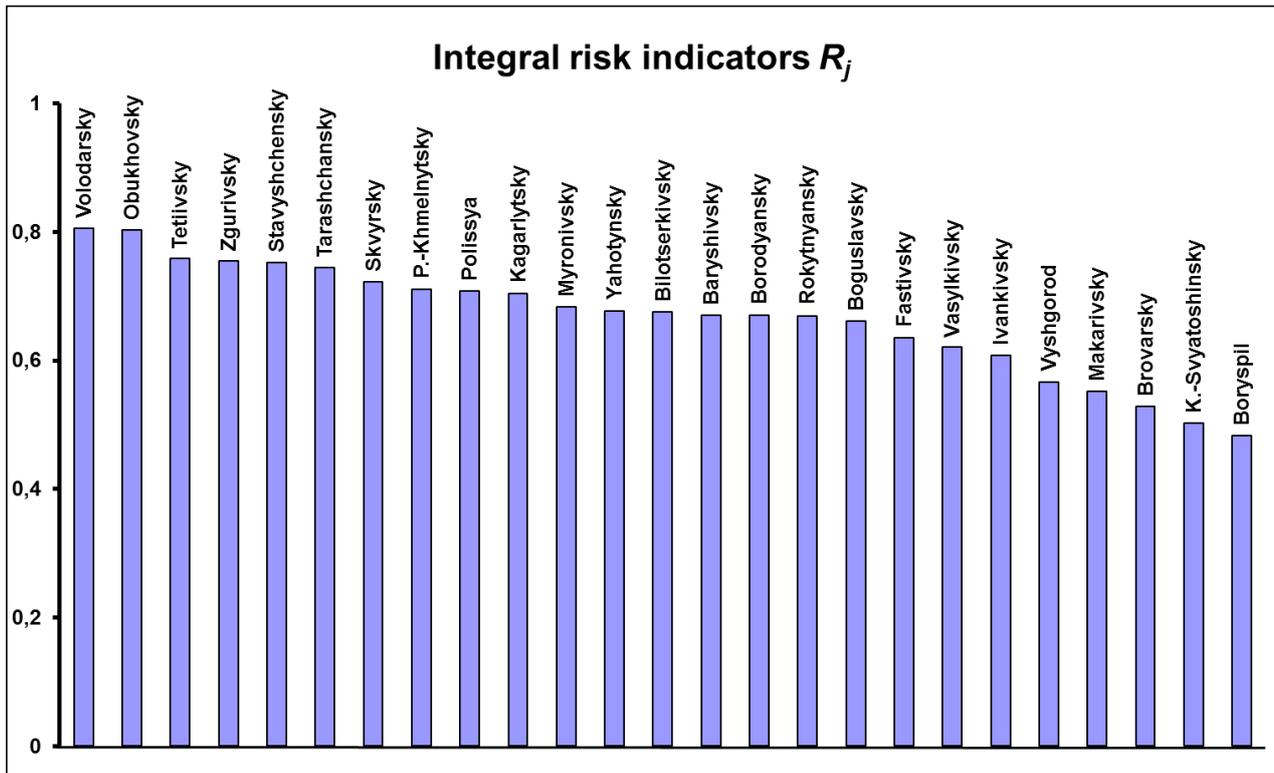


Figure 2. Ranking of regions in terms of their risk assessment of socio-economic development

- [14] V. Gray, *Principal Component Analysis: Methods, Applications and Technology* (Nova Science Pub. Inc., 2017)
- [15] S. Okafor, I. Okafor, *Factor Analysis in Economics: Theory and Applications* (LAP Lambert Academic Publishing, 2017)
- [16] G. Di Franco, A. Marradi, *Factor analysis and principal component analysis* (FrancoAngeli, 2013)
- [17] W.H. Finch, *Exploratory factor analysis* (SAGE Publications, 2019)
- [18] S. Farhady, *Expert Evaluation Game Factors* (LAP Lambert Academic Publishing, 2013)
- [19] J.J. Phillips, P.P. Phillips, *Handbook of training evaluation and measurement methods* (Routledge, 2016)
- [20] L. Ustinovieius, *Journal of Civil Engineering and Management* **7**, 321 (2001)
- [21] M. Rausand, S. Haugen, *Risk Assessment: Theory, Methods, and Applications (Statistics in Practice)* (Wiley, 2020)
- [22] A. Kuklin, M. Pecherkina, A. Tyrsin, A. Surina (2017)
- [23] *Statistical yearbook of Ukraine for 2017* (2017)