Using a comprehensive index technology to analyze structural changes in the regions' economic development in a COVID-19 pandemic: the case of Ukraine

Pavlo Hryhoruk[1,*], *Nila* Khrushch[1,**], and *Svitlana* Grygoruk[1,***]

[1]Khmelnytskyi National University, 11 Instytutska Str., Khmelnytskyi, 29016, Ukraine

Abstract. The paper investigates the issues of evaluating structural changes in the regions' economic development based on the comprehensive index assessment technology. The impact of the COVID-19 pandemic on regional development and changes in the regional structure is considered. The authors propose the use of block convolution to design a comprehensive index based on a set of metric initial indicators that characterize the regions' economic development. Grouping the set of initial indicators is carried out based on the method of an extreme grouping of parameters and the method of principal components. A weighted linear additive convolution was used to develop partial composite indices and an economic development comprehensive index. The practical approbation was carried out for the regions of Ukraine according to the data of 9 months of 2019 and the same period of 2020. To establish the regions' structure, we used the division of the comprehensive index values into intervals and further distributing regions into classes according to the level of economic development. There is a general decrease in the value of the integrated indicator in 2020, caused by the impact of the COVID-19 pandemic. However, no significant changes in the structure of the regions were detected, which indicates an equally negative impact of the pandemic for all regions of Ukraine.

1 Introduction

One of the most significant problems of regional development is to ensure sustainable economic growth. The economic system of any country is a multifunctional regional entity, so the definition of long-term priorities of strategic planning of regional development should be based on comprehensive assessments of the level of their economic development. They allow tracking the dynamics and asymmetry of development, to establish inequalities and gaps in the region's structure, to provide an analytical basis for the preparation of strategic decisions on the transformation of socio-economic development policy of individual regions.

Global problems related to climate changes, financial crises, intensified competition in global and domestic markets, deepened in 2020 due to another global challenge – the COVID-19 pandemic [1, 2]. Its destructive impact has been reflected in all spheres of public life, destroying established socio-economic processes and relationships. Measures, severe restrictions, lockdowns aimed at curbing the spread of the pandemic, were reflected in the negative effects of slowing down the socio-economic development of both regions and the world economy as a whole. They were a prerequisite for a new financial and economic crisis. This is evidenced by the results of analytical studies and forecast estimates of basic macroeconomic indicators provided by global institutions, in particular, the

World Bank (WB), the International Labor Organization (ILO), the World Health Organization (WHO), the United Nations (UN), the European Bank for Reconstruction and Development (EBRD) and others.

In particular, according to the ILO, the loss of labor income for the three quarters of 2020 compared to the corresponding period of 2019 is estimated at 10.7%, or 3.5 trillion USD [3]. The baseline forecast calculated by World Bank analysts [4] predicts a reduction in world GDP by 5.2% in 2020. And although the world economy is expected to grow by 4.3% in 2021, the pandemic may hold back economic activity and income growth for a long time [5]. UBS Chairman Axel Weber also made a cautious forecast about the pace of global economic recovery, noting that "it would be at least a year to go back to pre-crisis levels of GDP. It'll take another year or two to be anywhere near getting unemployment and pre-crisis growth back and so it would be quite a long recovery that we're facing" [6].

The consequences of the pandemic were especially acute in developing economies countries, particularly in Ukraine. Thus, according to the State Statistics Service of Ukraine [7], real GDP in the third quarter of 2020 compared to the third quarter of 2019 decreased by 3.5%. The financial result before taxation of large and medium-sized enterprises in the III quarter of 2020 amounted to 93.3 billion UAH of profit, while for the corresponding period of 2019 – UAH 342.8 billion in profit, which is 73% less. Exports of goods for the period under review decreased by 3.6%, and imports – by 14.3%.

---

*e-mail: violete@ukr.net
**e-mail: nila.ukr@gmail.com
***e-mail: grygoruk.svitlana@gmail.com

The main forecast macroeconomic indicators for the end of 2020, presented by the Cabinet of Ministers of Ukraine, envisage a fall in GDP by 4.8%, the inflation rate – 11.6%; unemployment rate – 9.4%; reduction of the average salary – 4.5%; decrease in exports – 5.5%, imports – 10% [8]. According to the EBRD, by the end of 2020 GDP was expected to decline by 5.5%, but in 2021 it is predicted to grow by 3% [9]. The most optimistic about the resumption of production are construction companies, the most pessimistic – service companies that have suffered the most from the introduction of quarantine restrictions.

The decline in macroeconomic indicators is directly caused by negative changes in regional development. To reduce the negative socio-economic consequences of the COVID-19 pandemic, it is necessary to identify trends, assess different scenarios of regional development, identify existing structural changes and develop a system of measures within regional development strategies to stabilize the situation. The presented macroeconomic forecasts necessitate research aimed at estimating the real losses from COVID-19 pandemic in terms of socio-economic development of regions, identifying areas of rational use of endogenous factors to ensure their sustainable economic growth, which will contribute to the achievement of the goals reflected in the State Strategy for Regional Development for 2021–2027 [8].

2 Literature review

Currently, there is a large number of different scholar's approaches to assess the economic development level and the establishment of regional differences and imbalances.

These studies are based mainly on the use of quantitatively measurable indicators that allow sound mathematical processing to shape conclusions. One of the most commonly used approaches is research based on the analysis of the GDP indicator and indicators derived from it like the Hoover Concentration Index, the Theil index, the Herfindahl index, etc. [10–14]. In particular, the authors also use the Klassen typology to track the dynamics and nature of changes in regional development.

Given the natural multidimensionality of regions' economic development description, widely used methods of multidimensional statistical analysis for their structuring by the level of this characteristic and determination of disparities between regions, in particular, cluster analysis, factor analysis, multidimensional scaling, structural equation method, Solow-Swan, and Mankiw-Romer–Weil growth models [12, 15–23], which allows grouping regions into homogeneous aggregates based on various quantifiable indicators, to identify gaps in the development of individual regions. Among the shortcomings of these approaches, in our opinion, it is worth noting the difficulty of taking into account the importance of individual indicators. The authors of the study, who used these tools, also noted that the grouping results are significantly influenced by clustering methods, which is also a disadvantage. The further development of multidimensional statistics' methods is reflected in the application of fuzzy clustering meth-

ods for structuring regions and identifying imbalances in their economic development, which is presented in [24–27].

Another way to take into account the multidimensionality for the description of regional development processes is to use the technology of comprehensive index assessment [28–33]. The vast majority of scientists' approaches in the presented studies are focused on designing a composite indicator of economic development by linear convolution of a set of quantitatively measured indicators. The differences are in the information base chosen for the study and how the results are interpreted. Among the shortcomings, it is worth noting the lack of consideration of the weight of the initial indicators or proper justification of the proposed weights, which in most cases it is proposed to determine the expert method. Besides, either a linear relationship between the values of the composite indicator and these levels, or a desirability scale without proper conversion of the original data is usually used to interpret the results and establish levels of economic development [31].

The study of issues related to assessing the impact of the COVID-19 pandemic on the economic development of economic systems both at the global level and at the level of individual national economies is currently one of the most relevant and is quite intensively studied by scientists. The vast majority of researchers are inclined to believe that overcoming the crisis is possible only after a few years, even with the total vaccination of the population, which should curb the spread of viral infection. Such conclusions are supported by the results of economic and mathematical modeling and evaluation of current and future trends in the economic system development. Issues related to the application of mathematical modeling to assess the impact of a pandemic on economic development are reflected, in particular, in publications [34–37]. However, it should be noted that the authors of these studies provide short-term forecast estimates of macroeconomic indicators at the level of national economies. with an emphasis on trends and potential scenarios for their development. The main attention is paid to the assessment of GDP change as one of the most important macroeconomic indicators. In our point of view, insufficient attention is currently paid to research to identify changes in the trends of economic development of certain regions of the country.

Our study aims to develop an approach to building an economic development comprehensive index for analyzing the impact of COVID-19 on Ukraine's regions development and identifying structural changes by combining the technology of comprehensive index assessment, multidimensional statistical analysis, and projection of results on the desirability scale.

3 Problem description and methodology

The economic development of the regions is characterized by a large number of indicators. They usually reflect the quantitative results of the activities of regional business entities and therefore have a metric origin, i.e., measured on one of the quantitative scales. This significantly simplifies their further analytical processing, because for indi-

cators of this nature it is quite correct to use mathematical operations.

One of the difficulties that arise in the process of processing such data and interpretation of results is their internal inconsistency, diversity, and inequality of impact on the studied quality. To concentrate the information contained in the initial indicators and reduce the dimension of characteristics' space, various computing technologies are used. One of them is the technology of comprehensive index assessment, which allows reducing the description of the studied phenomenon, in this case, the economic development of the regions, to a single comprehensive indicator. This is usually done by weighted convolution of the initial units. At the same time, there are several methodological problems to realize this process. First, the economic development of regions, as a complex phenomenon, requires the use of a large number of baselines for their description. Thus, the relative impact of each indicator on the final result is reduced. Secondly, there is a problem of reasonable determination of the weight of each component when they are integrated into a composite indicator.

A possible solution to these obstacles is the use of block convolution. Under such conditions, the initial set of indicators is divided into subsets that don't intersect. A partial composite index is constructed for each subset. The final result is settled by convolution of the constructed partial composite indices taking into account the weight of each obtained subset.

One of the approaches that allow getting a solution to this problem is the method of an extreme grouping of parameters. It is based on the hypothesis that the set of initial characteristics can be divided into groups, each of which reflects the effect of a certain factor – the latent characteristics of the group. Therefore, the method focuses on the selection of groups of parameters such that the relationships between the parameters within the group are maximum under the assumption that the number of such groups is fixed. It is assumed that the relationships within the group are explained by the relationship between some generalized latent characteristic of the group (generalized index) and the initial indicators included in this group. Direct relationships between initial indicators are unknown and may be absent. Since the indicators within each of these groups must be more closely related than the indicators of different groups, the task is to identify highly correlated groups of indicators.

Denote by $G = X_1, X_2, \ldots, X_n$ the set of initial indicators. The initial data for the method's computational procedure is the correlation matrix R of these indicators. Let $G_1, G_2, \ldots, G_s$ be subsets into which the set of initial indicators is divided:

$$\bigcup_{i=1}^{s} G_i = G, \tag{1}$$

$$G_i \cap G_j = \varnothing, \tag{2}$$

$i \neq j, i, j = 1, 2, \ldots, s$.

Denote by $H_1, H_2, \ldots, H_s$ – the corresponding latent characteristics (indicators) of each group. The criterion that allows you to determine the best grouping of indica-

tors has the form:

$$\sum_{i=1}^{s} \sum_{X_j \in G_i} |r_{X_j, H_i}|, \tag{3}$$

where r_{X_j, H_i} is the correlation coefficient between initial indicator X_j, which belongs to subset G_i, and common indicator H_i of subset G_i.

To obtain a division of the original set of indicators into subgroups, you can use the method of principal components. It is known that the model of transition from the system of initial indicators to the set of latent characteristics, which are the principal components, is reflected by the dependence:

$$Z^T = WF^T, \tag{4}$$

where Z^T – transposed matrix of standardized initial indicators' values, F^T – transposed matrix of principal components' values, W – matrix of principal components factor loadings:

$$Z = \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ z_{m1} & z_{m2} & \cdots & z_{mn} \end{pmatrix}, \tag{5}$$

$$F = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{m1} & f_{m2} & \cdots & f_{mn} \end{pmatrix}, \tag{6}$$

$$W = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{pmatrix}, \tag{7}$$

where m – the volume of the sample, which is used to measure the initial set of indicators.

The relationship between the values of indicators and principal components (factors) can be written as follows:

$$z_{ji} = \sum_{k=1}^{n} w_{ik} f_{jk} \tag{8}$$

where z_{ji} – i-th component (value) of Z_j, w_{ik} – factor loadings for F_k, f_{jk} – j-th components of F_k, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$.

Let us calculate the correlation coefficient between the initial indicator Xi and the principal component F_j, taking into account the fact that the principal components are non-correlated:

$$r_{F_i, F_j} = 0, \tag{9}$$

$i \neq j$.

As a result, we obtain:

$$r_{X_i, F_j} = r_{Z_i, F_j} = \frac{1}{m} \left(\sum_{k=1}^{n} W_{ik} F_k \right) F_j = w_{ij}. \tag{10}$$

Therefore, the correlation coefficient between the initial indicator and the principal component is equal to the

factor load of this component for the corresponding indicator. This fact allows us to conclude that to get the desired grouping of indicators it is necessary to analyze the values of the factor loadings of the principal components for each initial indicator. In this case, as the latent characteristic H_j of the group G_j, we choose the corresponding principal component F_j. To avoid the formation of empty groups or all groups, each of which will contain only one initial indicator, for grouping, we choose not all the principal components, but only the first s most influential, which explain the given share of variance of initial indicators. The value of s is defined as the smallest value of the number of principal components for which the inequality is met:

$$\frac{\sum_{i=1}^{s} \lambda_i}{n} \geq \gamma \qquad (11)$$

where λ_i – eigenvalues, ordered by decreasing their values, γ – a predetermined explanation fraction of the initial indicators' variance by the principal components. Typically, this value is selected from 0.70 to 0.80.

In the group of homogeneous indicators, it is expedient to include those initial indicators for which the corresponding values of factor loadings for the principal components on absolute value will have the greatest values. To construct a partial composite index I_j for each formed group G_j, we use one of the formulas for weighted convolution [38]:

$$I_j = \sum_{i \in K_j} \alpha_i^{(j)} U_i^{(j)} \qquad (12)$$

$$I_j = \prod_{i \in K_j} \left(U_i^{(j)}\right)^{\alpha_i^{(j)}} \qquad (13)$$

$$I_j = -1 + \prod_{i \in K_j} \left(1 + U_i^{(j)}\right)^{\alpha_i^{(j)}} \qquad (14)$$

where $U_i^{(j)}$ – normalized values of those indicators X_i, that belong to subset G_j, $a_i^{(j)}$ – weight coefficients of appropriate indicators, K_j – set of indices for those indicators X_i, that belong to subset G_j, $j = 1, 2, \ldots, s$.

The initial indicators are transformed to normalize form according to the formula (15) or formula (16) [17]:

$$u_{ij} = 1 - \frac{|x_{ij} - x_j^*|}{x_{jmax} - x_{jmin}} \qquad (15)$$

$$u_{ij} = \begin{cases} \frac{x_{ij}}{x_{jmax}}, & \text{when } X_j \text{ is an incentive;} \\ \frac{x_{jmin}}{x_{ij}}, & \text{when } X_j \text{ is a disincentive;} \end{cases} \qquad (16)$$

where u_{ij} – normalized values of indicators, x_{ij} – initial values of indicators, $x_{jmin} = \min x_{ij}$, $x_{jmax} = \max x_{ij}$,

$$x_i^* = \begin{cases} x_{jmax}, & \text{when } X_j \text{ is an incentive;} \\ x_{jmin}, & \text{when } X_j \text{ is a disincentive;} \end{cases} \qquad (17)$$

$i=1, 2, \ldots, m$, $j=1, 2, \ldots, n$, m – number of units under study, n – number of initial indicators.

The normalization procedure is necessary to extract the units of measurement of the original indicators and reducing their values to a scale from 0 to 1. This step is aimed at simplifying the further interpretation of the calculation result. To calculate the weight coefficients, we propose to use the components of eigenvector V_j:

$$\alpha_i^{(j)} = \frac{\left(v_i^{(j)}\right)^2}{\sum_{i \in K_j} \left(v_i^{(j)}\right)^2}, \qquad (18)$$

where $a_i^{(j)}$ – weight coefficients of appropriate indicators, $v_i^{(j)}$ – components of j-th eigenvector V_j, that correspond to initial indicators X_i from the G_j, K_j – set of indices for those indicators X_i, that belong to subset G_j.

Equation (18) meets the condition, that the sum of weight coefficients should be equal to 1. This condition with the normalization procedure provides the location of partial composite indicators values in the range [0; 1].

We propose to calculate the final economic development comprehensive index I_{COM} using partial composite indicators I_j based on one of the convolution's forms like (12), (13), (14). For example, for linear weighted convolution appropriate expression has a form:

$$I_{COM} = \sum_{j=1}^{s} \beta_j I_j. \qquad (19)$$

Weight coefficients β_j are calculated in proportion to the eigenvalues λ_j that correspond with G_j, $j = 1, 2, \ldots, s$:

$$\beta_j = \frac{\lambda_j}{\sum_{k=1}^{s} \lambda_k}. \qquad (20)$$

Under such conditions, the values of the I_{COM} will also be in the range from 0 to 1. This approach to calculations simplifies the interpretation of the result.

To assess the studied objects' structure, the range of values of the comprehensive index should be divided into ranges. Dividing the range [0; 1] of values of the comprehensive index into intervals of the same length to achieve this goal is impractical.

First, ranges can be formed that don't cover any of the objects under study.

Second, the latent characteristic under study is usually nonlinear, and the use of intervals of the same length can disrupt the true structure of objects.

Third, such a division can be led to a situation where one group includes objects that have significant differences in the values of the integrated indicator, while two neighboring objects belonging to different groups may have a slight deviation of the values of the comprehensive index.

To solve the problem of grouping, you can also use the approach presented in [17], in which the definition of the boundaries of the ranges is carried out by calculating the ratios of two adjacent values of the integrated indicator:

$$\delta_j = \frac{I_j}{I_{j-1}}, \qquad (21)$$

$j = 2, 3, \ldots, m$.

The basis for the transition to a new range of values of the comprehensive index is a significant rise in the change

of values of δ_j. The grouping objects is executed according to the level of the corresponding values of the comprehensive index. This approach also has drawbacks. Given the slight difference in the values of the integrated indicator, which are in the middle of the range of all its possible values, one of the groups can have a very large number of objects, which will be significantly different from the content of other groups. Besides, in the case of a slight discrepancy in the values of the comprehensive index for neighboring objects, a significant rise in the values of δ_j may not be observed. Thus, all objects can belong to one group. It is also necessary to take into account the fact that the value of δ_j is also affected by the level of values of the comprehensive index for which this value is calculated. And the closer these values are to 0, the smaller should be the hike in the change of values of δ_j, which decides on the formation of a new range.

The iterative procedure presented in [31] can be used to determine the limit values of the comprehensive index' ranges. Its advantage is the "adjustment" of grouping ranges to the value of a specific sample, which makes its application more practical. However, the disadvantage of this approach is the use of a training sample.

Another approach that allows you to solve this problem is the use of desirability scales, which allow you to match the quantitative and qualitative levels and group objects according to the level of studied quality. One such scale is the Harrington scale. The use of this scale involves the transformation based on Harrington's function [39]:

$$H(Z_i) = exp(-exp(-Z_i)), \qquad (22)$$

where Z_i is the value of the indicator on the scale of partial indicators Z. The values $d = H(Z)$ of the Harrington's function form the desirability scale.

The correspondence between the values of Z_j and the values of the initial indicators I_j is determined by the formula:

$$Z_j = (Z^* - Z_*)\frac{I_{COM_j} - I_{min}}{I_{max} - I_{min}} + Z_*, \qquad (23)$$

where Z_j – current value of the Z-scale, corresponds with the value of I_{COM_j}; I_{COM_j} – current value of comprehensive index I_{COM}; Z_* and Z^* – low and high bounds of Z-scale, which define the workspace of Z_j; I_{min}, I_{max} – minimum and maximum of I_{COM}; $j = 1, 2, \ldots, m$.

Transformation (23) is required to match the value of the comprehensive index I_{COM} and Z-scale with the correspondence of the minimum and maximum values of both indicators.

Next step, we identify the value of $d_j = H(Z_j)$, $j = 1, 2, \ldots, m$, and distribute objects under study into five groups by qualitative development level of the group (table 1).

This approach allows taking into account the nonlinear nature of the studied characteristic, in this case, the economic development level, as well as to investigate changes in the structure of the objects under study by the values of the comprehensive index calculated for different periods.

Table 1. The relationships between the quantitative values of the desirability scale and qualitative development levels of group

Qualitative levels of development	The range of quantitative values on the desirability scale
relatively high	0.80..1.00
above average	0.63..0.80
average	0.37..0.63
below average	0.20..0.37
relatively low	0.00..0.20

Table 2. The relationships between the quantitative values of the desirability scale and qualitative development levels of group

Code	Region	Code	Region
r-01	Vinnytsia	r-13	Mykolaiv
r-02	Volyn	r-14	Odesa
r-03	Dnipro	r-15	Poltava
r-04	Donetsk	r-16	Rivne
r-05	Zhytomyr	r-17	Sumy
r-06	Zakarpattia	r-18	Ternopil
r-07	Zaporizhzhia	r-19	Kharkiv
r-08	Ivano-Frankivsk	r-20	Kherson
r-09	Kyiv	r-21	Khmelnytskyi
r-10	Kyrovohrad	r-22	Cherkasy
r-11	Luhansk	r-23	Chernivtsi
r-12	Lviv	r-24	Chernihiv

4 Findings

Let us consider the practical testing of the proposed approaches to the calculation of t economic development comprehensive index for Ukraine's regions, grouping regions based on their values, and the study of structural changes in the resulting grouping caused by the COVID-19 pandemic. We choose the data of the State Statistics Service of Ukraine [7] and the Ministry of Development of Communities and Territories of Ukraine [40] for the period of the first 9 months of 2019 and the first 9 months of 2020 as the information base for the calculations. We choose the following initial indicators:

X_1 – Volume of sold industrial products per capita, UAH;

X_2 – Volume of agricultural production per capita of the rural population, UAH;

X_3 – Volume of construction works performed per capita, UAH;

X_4 – Volume of capital investments per capita cumulatively since the beginning of the year, UAH;

X_5 – Exports of goods per capita, USD;

X_6 – An unemployment rate of the population aged 15-70 years (according to the ILO's Methodology), %;

X_7 – Employment rate of the population aged 15-70 years (according to the ILO's Methodology), %;

X_8 – Index of real wages, %;

X_9 – The volume of housing commissioned per 10 thousand people, sq. meters of the total area;

X_{10} – The volume of freight turnover of road and rail transport, thousand ton-kilometers per 1000 population, thousand ton-km.

Table 3. Indicator's values for data for first three quarters of 2019

Code	Values									
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
r-01	39228.9	20678.0	3774.4	5730.1	697.8	9.8	57.9	112.1	1248.9	2721.7
r-02	21736.3	11004.0	1322.5	8039.0	503.6	11.7	50.2	109.1	2978.8	2261.9
r-03	110916.2	25975.0	4163.1	13679.4	1971.8	7.7	59.5	112.9	858.1	3143.8
r-04	53475.4	17444.0	911.7	4294.1	804.4	13.7	50.9	108.9	105.4	2583.4
r-05	28004.0	15169.0	1301.6	3781.7	438.3	9.8	57.3	107.1	1121.4	2631.3
r-06	14401.8	3895.0	1008.4	3135.1	906.9	9.1	55.3	106.3	2852.6	3646.2
r-07	86201.6	22827.0	1456.1	5135.8	1369.6	9.5	57.8	111.7	351.4	2866.1
r-08	37051.0	5895.0	1522.7	4051.9	491.2	7.6	55.8	108.1	4151.6	1987.5
r-09	50780.4	17359.0	3627.6	17018.5	815.3	5.9	59.3	111.7	6897.8	2146.4
r-10	24742.5	28345.0	1361.0	4134.5	508.9	11.2	55.6	107.9	598.2	7529.0
r-11	7917.8	14675.0	175.7	898.1	57.2	14.2	58.4	108.8	59.5	5611.3
r-12	30915.6	8160.0	2744.3	6599.9	631.4	6.8	57.5	107.4	3842.8	1373.6
r-13	40158.9	24129.0	2075.6	7038.5	1451.0	9.7	58.9	112.4	598.3	4675.4
r-14	18915.3	11939.0	4653.9	5677.4	468.8	6.1	57.9	106.3	2844.4	5052.0
r-15	91151.2	22940.0	3673.1	10075.9	1121.6	11.1	56.1	109.7	1118.2	3515.0
r-16	26703.4	8958.0	1545.4	3649.4	277.9	8.6	57.9	111.0	2019.0	3480.3
r-17	33304.1	21871.0	878.7	4306.1	547.1	8.0	58.7	108.0	722.6	2697.3
r-18	14411.8	12093.0	1428.1	5437.1	308.9	10.4	53.4	109.9	2571.6	1119.1
r-19	51750.2	23955.0	3793.0	5187.6	360.9	5.0	62.2	108.3	984.1	1620.8
r-20	22025.6	24547.0	1092.2	5644.6	220.1	10.3	58.3	107.1	885.8	1692.7
r-21	25488.3	17527.0	1860.6	4677.0	356.6	8.7	56.6	109.7	1544.0	1870.5
r-22	45162.6	20718.0	1054.1	5544.1	483.4	8.5	58.7	109.6	944.5	3133.7
r-23	11193.4	6784.0	1330.5	2501.1	166.1	7.2	58.6	109.4	3952.4	1954.4
r-24	24933.9	20188.0	1212.6	5185.2	547.7	10.5	58.3	109.3	1150.5	1644.5

Table 4. Indicator's values for data for first three quarters of 2020

Code	Values									
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
r-01	36446.7	44134.0	4023.4	4345.4	679.2	10.4	56.9	109.0	720.5	3528.4
r-02	21419.8	25428.0	1376.7	7016.7	449.0	12.3	49.2	101.8	1806.3	2728.5
r-03	95582.3	58710.0	3737.0	10444.6	1755.7	8.2	58.4	106.5	235.0	5894.9
r-04	43994.1	43567.0	1300.5	3345.8	700.3	14.5	49.8	103.1	45.5	1188.8
r-05	26948.8	33675.0	1028.8	3118.3	397.7	10.5	55.8	108.7	438.6	5683.4
r-06	13637.5	7090.0	861.8	1956.5	769.9	10.2	54.3	104.3	1430.0	3680.1
r-07	78074.0	56801.0	1011.5	3907.7	1267.3	10.4	56.5	108.4	150.0	4121.6
r-08	30883.1	12247.0	1581.1	2212.1	400.2	8.1	54.4	107.4	1833.7	1031.1
r-09	49227.9	32672.0	4524.2	8784.3	762.2	6.6	58.2	105.1	3932.9	5024.2
r-10	26339.1	56737.0	958.0	3910.2	706.2	12.3	54.2	111.7	231.0	17025.4
r-11	6154.2	37587.0	169.3	620.4	45.8	15.2	57.3	112.5	51.0	299.9
r-12	30522.8	17182.0	3226.0	3772.5	646.3	7.4	56.2	106.4	1734.9	3386.5
r-13	40592.0	49388.0	1732.1	3632.5	1311.1	10.3	57.9	110.9	205.1	8912.5
r-14	20213.1	18536.0	6105.9	4248.0	400.4	6.9	57.0	109.4	1391.7	5460.3
r-15	80188.6	48838.0	3956.8	8287.4	1163.0	11.7	55.2	106.6	457.1	4618.8
r-16	28737.4	19118.0	1496.5	2010.7	297.3	9.1	56.7	109.2	1080.6	5255.5
r-17	30338.2	48048.0	1008.9	2983.4	596.6	9.1	57.2	110.1	282.9	2086.5
r-18	13649.8	26998.0	1472.1	3532.0	300.9	11.3	52.1	108.4	1543.0	1236.3
r-19	45981.1	61006.0	3608.1	3819.2	381.7	5.8	60.8	105.5	870.3	3537.0
r-20	22620.1	56237.0	763.4	2250.0	198.3	11.1	57.2	110.5	355.6	841.6
r-21	26941.7	39572.0	2763.3	4151.3	355.2	9.5	55.3	109.0	1069.1	3948.0
r-22	44996.7	43113.0	1112.1	3232.8	523.3	9.3	57.6	107.0	398.9	4440.9
r-23	10588.4	14647.0	1553.1	1448.2	126.7	8.6	57.2	107.7	1670.9	729.0
r-24	24360.0	43737.0	1687.1	3600.9	554.4	11.5	56.8	109.1	577.4	2961.3

We assigned to each of Ukraine regions' names the corresponding code which we used for the designation of each of them to further use (table 2).

The values of initial indicators to provide calculations are shown in tables 3 and 4.

Let's group the initial indicators by the method of an extreme grouping of parameters. To determine the correlations between the initial indicators and the latent characteristics of each group in the context of maximizing the expression (3), we use the method of principal components. Taking into account expression (10), it is necessary to calculate the factor loadings for the selected principal components and choose the largest from them in absolute value. The number of groups is defined as the number of principal components that explain a given level of variance of the initial indicators following expression (11).

We choose the level of explanation of the variance of the initial indicators as $\gamma = 0.80$. Under such conditions, it is necessary to choose the first four principal components. The values of the eigenvectors and the eigenvalues of the corresponding correlation matrices of the initial indicators are given in Tables 5 and 6, and the values of the factor loadings – in tables 7 and 8.

Table 5. Most significant eigenvalues of correlation matrix and values of appropriate eigenvectors for data 2019

Eigenvalues			
λ_1	λ_2	λ_3	λ_4
3.64	2.42	1.36	0.87
Eigenvectors			
V_1	V_2	V_3	V_4
0.45	-0.16	0.13	-0.11
0.29	-0.39	-0.28	-0.14
0.38	0.23	-0.14	0.32
0.42	0.16	0.27	0.15
0.42	-0.17	0.23	0.14
-0.20	-0.49	0.35	0.02
0.22	0.12	-0.69	-0.20
0.12	-0.12	0.21	-0.26
0.01	0.57	0.25	0.23
-0.02	-0.34	-0.25	0.81

Table 6. Most significant eigenvalues of correlation matrix and values of appropriate eigenvectors for data 2020

Eigenvalues			
λ_1	λ_2	λ_3	λ_4
3.35	2.53	1.76	0.92
Eigenvectors			
V_1	V_2	V_3	V_4
0.49	-0.11	0.14	-0.26
0.28	-0.44	-0.05	-0.20
0.34	0.33	-0.19	0.06
0.46	0.14	0.24	0.09
0.45	-0.15	0.19	0.05
-0.22	-0.38	0.43	0.06
0.20	-0.04	-0.64	-0.27
-0.11	-0.36	-0.49	0.20
0.00	0.57	-0.02	0.21
0.22	-0.21	-0.10	0.85

Analysis of Tables 7, 8 allows us to formulate a conclusion, that we have the following distribution of initial indicators between subsets G_j:

Table 7. Values of factor loadings for data 2019

Initial indicators	Values of factor loadings			
	w_1	w_2	w_3	w_4
X_1	0.86	0.24	0.16	0.10
X_2	0.56	0.61	0.32	0.13
X_3	0.73	0.36	0.16	0.30
X_4	0.81	0.25	0.31	0.14
X_5	0.79	0.27	0.27	0.13
X_6	0.38	0.77	0.41	0.02
X_7	0.42	0.19	0.80	0.19
X_8	0.24	0.19	0.25	0.24
X_9	0.02	0.88	0.29	0.22
X_{10}	0.04	0.53	0.29	0.76

Table 8. Values of factor loadings for data 2020

Initial indicators	Values of factor loadings			
	w_1	w_2	w_3	w_4
X_1	0.89	0.17	0.19	0.25
X_2	0.52	0.70	0.06	0.20
X_3	0.62	0.52	0.25	0.06
X_4	0.84	0.22	0.31	0.09
X_5	0.83	0.23	0.25	0.05
X_6	0.40	0.61	0.57	0.05
X_7	0.36	0.07	0.85	0.26
X_8	0.21	0.58	0.66	0.19
X_9	0.01	0.90	0.03	0.20
X_{10}	0.40	0.33	0.14	0.81

$G_1 = \{X_1, X_3, X_4, X_5\}$;
$G_2 = \{X_2, X_6, X_8\}$;
$G_3 = \{X_7, X_9\}$;
$G_4 = \{X_{10}\}$.

Note, that subset G_4 consists of one initial indicator X_{10}, so, partial composite index I_4 coincides with this indicator.

To calculate patrial composite indices I_j, $j = 1, 2, 3, 4$, we conduct a normalization procedure for initial data. In this case, we execute this step using formula (16), because this way allows keeping the proportions between the values of the indicator, which is important in the calculation of composite index's values.

We also take into account, that indicator X_6 is a disincentive, and other indicators are incentives. Weight coefficients we calculate, using formula (18). Values of composite indices I_j, $j = 1, 2, 3, 4$, have been calculated using linear convolution by expression (12). To calculate the comprehensive index, we also use weighted linear convolution like (19) with weight coefficients, obtained by the formula (20).

To correctly compare the results of calculations and identify changes in the levels of the comprehensive index for the relevant periods and the regions' structuring, the values for 2019 and 2020 will be combined into one sample. The normalization procedure is performed for the combined data.

Further calculations of both partial composite indices and comprehensive indices are executed for each period

separately. The values of the selected eigenvector elements for calculating the weights by formula (18) are different, as well as the corresponding eigenvalues that will be used to calculate the weights of the generalized indicator by formula (20). So, for the data of 2019 and the data of 2020, we obtain different values of weight coefficients, which means that both composite and comprehensive indicators will be calculated according to different rules.

Therefore, for a more accurate comparison of the results, we propose to calculate corresponding weights as the average values of the appropriate components obtained for 2019 and 2020.

Weight coefficients for calculation of composite indices I_1-I_4 in accordance with distribution initial indicators to G_j have such values: $w_{11} = 0.30; w_{13} = 0.18; w_{14} = 0.27; w_{15} = 0.25; w_{22} = 0.25; w_{26} = 0.28; w_{29} = 0.47; w_{37} = 0.77; w_{38} = 0.23; w_{4,10} = 1.00$. Values of weight coefficients for calculation of comprehensive index are: $w_1^{COM} = 0.42, w_2^{COM} = 0.28, w_3^{COM} = 0.19, w_4^{COM} = 0.11$. The results of the calculations of comprehensive index values are presented in table 9.

Table 9. Comprehensive index values

Code	2019	2020	Code	2019	2020
r-01	0.45	0.46	r-13	0.48	0.49
r-02	0.40	0.38	r-14	0.48	0.46
r-03	0.67	0.65	r-15	0.55	0.55
r-04	0.37	0.37	r-16	0.38	0.37
r-05	0.37	0.38	r-17	0.39	0.39
r-06	0.39	0.34	r-18	0.36	0.34
r-07	0.49	0.50	r-19	0.48	0.50
r-08	0.44	0.36	r-20	0.36	0.35
r-09	0.68	0.59	r-21	0.38	0.41
r-10	0.40	0.48	r-22	0.41	0.42
r-11	0.28	0.26	r-23	0.39	0.32
r-12	0.47	0.43	r-24	0.38	0.39

A comparison of the calculation results of the comprehensive index shows that for most of Ukraine's regions there is a decrease in its values. In our opinion, this fact indicates a negative impact of the pandemic COVID-19 on economic development. At the same time, for some regions, in particular, Vinnytsia, Zhytomyr, Zaporizhia, Kirovohrad, Mykolaiv, Kharkiv, Khmelnytskyi, and Chernihiv regions, there is an increase in the values of the indicator in 2020. This increase is especially noticeable for the Kirovohrad region. This can be explained by the fact that for a long time in this area was the best epidemiological situation in Ukraine. Also, the growth of industrial production, in particular pharmaceuticals one in Vinnytsia, Kirovohrad, Zaporizhia, and Kharkiv regions.

Let us consider the changes in the structure of Ukraine's regions in 2020 compared to 2019 in terms of the economic development comprehensive index. Given the relatively high density of values of the comprehensive index for different regions, the use of the approach to the grouping of regions, based on the analysis of the values of the delta, calculated by the formula (21), doesn't allow to determine their structure. Therefore, to solve this problem,

we apply an approach based on the use of the Harrington desirability scale. For this purpose, we transform the values of the integrated indicator according to formulas (22) and (22). The distribution of regions by groups is executed according to the rules given in table 1. The results of the calculations are listed in table 10.

Table 10. Identifying the structure of regions for data 2019 and 2020

Code	$d = H(Z)$		Level of development	
	2019	2020	2019	2020
r-01	0.72	0.75	above average	above average
r-02	0.47	0.35	average	below average
r-03	0.99	0.99	relatively high	relatively high
r-04	0.29	0.28	below average	below average
r-05	0.27	0.34	below average	below average
r-06	0.44	0.15	average	relatively low
r-07	0.86	0.88	relatively high	relatively high
r-08	0.68	0.25	above average	below average
r-09	0.99	0.97	relatively high	relatively high
r-10	0.46	0.84	average	relatively high
r-11	0.00	0.00	relatively low	relatively low
r-12	0.81	0.66	relatively high	above average
r-13	0.84	0.86	relatively high	relatively high
r-14	0.84	0.77	relatively high	above average
r-15	0.95	0.95	relatively high	relatively high
r-16	0.37	0.31	average	below average
r-17	0.42	0.43	average	average
r-18	0.24	0.13	below average	relatively low
r-19	0.83	0.88	relatively high	relatively high
r-20	0.22	0.18	below average	relatively low
r-21	0.37	0.56	average	average
r-22	0.56	0.58	average	average
r-23	0.39	0.06	average	relatively low
r-24	0.34	0.42	below average	average

The analysis of results obtained shows, that the first group with a relatively high level of economic development is quite large. Traditionally, this group includes Kyiv, Kharkiv, and Dnipropetrovsk regions, which in the "pre-pandemic" period had a fairly high level of economic development. These regions have a fairly high production potential, they account for a significant share of foreign investment in 2020 and therefore the pandemic has not had such a destructive impact on the economic development of these regions. Zaporizhia and Poltava regions also have significant potentials and were distributed to this group. The lowest level of economic development is in the Luhansk region, and in 2020 the situation has not changed.

It should be noted that for many regions there have been no changes in the level of economic development, although there has been a decrease in the value of the corresponding comprehensive index. For those regions where changes are taking place, they are usually associated with a decline in economic development. The only exception is the Kirovohrad region.

The most significant decrease in the level took place in Zakarpattia, Ternopil, and Chernivtsi regions. These are the regions that were the first to suffer from the pandemic

and were in the "red" zone for a long time, which negatively affected all indicators of economic development and led to a significant reduction in the corresponding comprehensive index values.

Thus, the results of the research demonstrate the fundamental possibility of applying the proposed approach to the study of economic development of regions by constructing an integrated indicator. The analysis of the structure of the regions showed the real impact of the pandemic on the development of almost all regions, which led to the corresponding structural changes.

5 Conclusions

The study of economic development trends both in the economic system of the country as a whole and at the level of individual regions remains the focus of the most significant research. The results of such studies are especially important in periods of global challenges, one of which at this stage of the world community development was the COVID-19 pandemic. Solving the problems of assessing the level of development of regions, their structuring, identifying gaps and breaks in the development of individual territorial units is complicated by the significant multidimensionality of their description. The use of analytical methods of information processing based on economic and mathematical models allows us to present it in a concentrated form without significant losses, which contributes to the adoption of sound management decisions and the development of strategic plans for regional development. Therefore, models that allow for a significant reduction in baseline and identify latent characteristics of the studied phenomena are important for studies. In particular, such approaches include models based on the comprehensive index assessment technology.

The approaches offered in the article allow estimating the level of economic development of regions by block convolution of the set of initial indicators into a single complex measure – an economic development comprehensive index. Thus, the toolkit which allows to carry out a grouping of initial indicators to take into account the weights of components at the construction of such indicators, and also the weights of partial composite indices at their convolution into the economic development comprehensive index is offered. The article proposes some approaches to grouping regions by the level of economic development. An approach based on the transformation of the comprehensive index values with the projection of the result on the desirability scale is chosen for practical implementation. This way allows to rank regions, determine their structure by this characteristic and assess structural changes over time.

According to the research outcomes, it can be concluded that the COVID-19 pandemic has a destructive impact on the economic development of the vast majority of Ukraine's regions, which was reflected both in changes in the values of the comprehensive index and in the regions' structure.

The direction of further research is the development of analytical tools to take into account indicators of non-

metric origins in the assessment procedures for the identifying level of economic development.

References

[1] S. Semerikov, S. Chukharev, S. Sakhno, A. Striuk, V. Osadchyi, V. Solovieva, T. Vakaliuk, P. Nechypurenko, O. Bondarenko, H. Danylchuk, E3S Web of Conferences **166** (2020)

[2] S. Semerikov, H. Kucherova, V. Los, D. Ocheretin, CEUR Workshop Proceedings **2845**, 22 (2021)

[3] International Labour Orgaization, *ILO monitor: COVID-19 and the world of work. Sixth edition: Updated estimates and analysis* (2020), `https://cutt.ly/GjAdvEa`

[4] The World Bank, *The Global economic outlook during the COVID-19 pandemic: a changed world* (2020), `https://cutt.ly/ff75GUY`

[5] The World Bank, *Global economy to expand by 4% in 2021; vaccine deployment and investment key to sustaining the recovery* (2021), `https://cutt.ly/OjAdKpT`

[6] H. Ellyatt, *'don't be under the illusion' that an economic recovery will come quickly with a vaccine, UBS' Weber says* (2020), `https://cutt.ly/ijAdCIO`

[7] *Derzhavna sluzhba statystyky Ukrainy (State statistics service of Ukraine)* (2020), `http://www.ukrstat.gov.ua/`

[8] *Derzhavna stratehiia rehionalnoho rozvytku na period 2021-2027 (State strategy for regional development for 2021-2027)* (2020), `https://cutt.ly/bjAfuSS`

[9] A. Williams, *EBRD revises down economic forecasts amid continuing coronavirus uncertainty* (2020), `https://cutt.ly/qjAd1F9`

[10] M.V. Moroshkina, Finance and Credit **10** (2020)

[11] E. Dhimitri, A. Gabeta, M. Bello, The USV Annals of Economics and Public Administration **15**, 51 (2015)

[12] A. Ramadhan, A. Sutandi, A.E. Pravitasari, IOP Conference Series: Earth and Environmental Science **556**, 012008 (2020)

[13] A. Soejoto, D. Fitrayati, L.R.N. Sholikhah, International Journal of Applied Business and Economic Research **14**, 9493 (2016)

[14] J. Li, Y. Chen, Z. Xi, X. Zou, Journal of Physics: Conference Series **1574**, 012170 (2020)

[15] T.A. Munandar, Azhari, A. Musdholifah, L. Arsyad, Journal of Computer Sciences **14**, 81 (2018)

[16] A.F. Doni, Y.D.P. Negera, O.A.H. Maria, Journal of Physics: Conference Series **1569**, 022078 (2020)

[17] P. Hryhoruk, N. Khrushch, S. Grygoruk, Periodicals of Engineering and Natural Sciences **7**, 712 (2019)

[18] L. Chagovets, S. Prokopovych, V. Kholod, Development Management **18**, 43 (2019)

[19] S. Eshugova, International Journal of Engineering & Technology **7**, 340 (2018)

[20] D. Cziraky, J. Sambt, J. Rovan, J. Puljiz, European Journal of Operational Research **174**, 427 (2006)

[21] P. Hryhoruk, N. Khrushch, S. Grygoruk, International Journal of Industrial Engineering & Production Research **31** (2020)

[22] B. Bal-Domanska, in *Advanced Studies in Classification and Data Science. Studies in Classification, Data Analysis, and Knowledge Organization*, edited by T. Imaizumi, A. Okada, S. Miyamoto, F. Sakaori, Y. Yamamoto, M. Vichi (Springer, Singapore, 2020), pp. 19–29

[23] V. Babenko, A. Panchyshyn, L. Zomchak, M. Nehrey, Z. Artym-Drohomyretska, T. Lahotskyi, WSEAS Transactions on Business and Economic **8**, 209 (2021)

[24] T.A. Munandar, R. Wardoyo, Intrnational Journal of Computer Applications **123**, 17 (2015)

[25] K. Gorbatiuk, O. Mantalyuk, O. Proskurovych, O. Valkov, SHS Web of Conferences **65**, 04008 (2019)

[26] K. Gorbatiuk, O. Mantalyuk, O. Proskurovych, O. Valkov, *Application of Fuzzy Clustering to Shaping Regional Development Strategies in Ukraine*, in *Proceedings of the 6th International Conference on Strategies, Models and Technologies of Economic Systems Management*, edited by P. Hryhoruk, N. Khrushch (Atlantis Press, Paris, 2019), Vol. 95 of *Advances in economics, business and management research*, pp. 271–276

[27] X. Wang, in *Big Data Analytics for Cyber-Physical System in Smart City*, edited by M. Atiquzzaman, N. Yen, Z. Xu (Springer, Singapore, 2020), Vol. 1303

[28] J. Gurna, K. Gurna, Statistics in Transition New Series **16**, 265 (2015)

[29] D.F. Meyer, J. De Jongh, N. Meyer, International Journal of Business and Management Studies **8**, 100 (2016)

[30] J. De Jongh, D. Meyer, Administratio Publica **27**, 122 (2019)

[31] P. Hryhoruk, N. Khrushch, S. Grygoruk, *An approach to design a composite index of economic development and identifying the bounds of its levels*, in *Proceedings of 9th International Conference on Advanced Computer Information Technologies* (Ceske Budejovice, 2019), pp. 48–51

[32] M. Stanichkova, L. Melecky, Regional Studies, Regional Science **5**, 231 (2018)

[33] V. Royuela, Social Indicators Research **148**, 831 (2020)

[34] A. Al-Baidhani, *COVID-19 economic impact on developed and developing economies* (2020), `https://ssrn.com/abstract=3700352`

[35] G. Gondwe, *Assessing the impact of COVID-19 on Africa's economic development* (2020), `https://cutt.ly/ijAkGWt`

[36] M. Gallagher, *COVID-19 economic impact assessment* (2020), `https://cutt.ly/1jAkKmX`

[37] M.A.R. Estrada, Estudios Economicos **37**, 139 (2020)

[38] P. Hryhoruk, N. Khrushch, S. Grygoruk, E3S Web of Conferences **166**, 13023 (2020)

[39] E. Harrington, Industrial Quality Control **21**, 494 (1965)

[40] *Ministerstvo rozvytku hromad ta terytorii Ukrainy. Reitynhova otsinka rehioniv (Ministry for communities and territories development of Ukraine. Rating assessment of regions)* (2020), `https://cutt.ly/gjJ5Hp0`