Original scientific paper
Journal of Central European Agriculture, 2024, 25(2), p.588-605

DOI: /10.5513/JCEA01/25.2.4176

Assessment of regional development level in Romania through Principal component analysis

Evaluarea nivelului de dezvoltare regională din România folosind analiza componentelor principale

Ancuţa Simona ROTARU[1], Cristina HEGEDUŞ[1], Ioana Delia POP[2], Sorin VÂTCĂ[3], Anamaria VÂTCĂ[1] (✉), Luisa ANDRONIE[2]

[1] Department of Fundamental Science, Faculty of Animal Science and Biotechnology, University of Animal Sciences and Veterinary Medicine Cluj-Napoca, 400372 Cluj-Napoca, Romania

[2] Department of Land Measurements and Exact Sciences, Faculty of Forestry and Cadastre, University of Animal Sciences and Veterinary Medicine Cluj-Napoca, 400372 Cluj-Napoca, Romania

[3] Department of Plant Physiology, Faculty of Agriculture, University of Animal Sciences and Veterinary Medicine Cluj-Napoca, 400372 Cluj-Napoca, Romania

✉ Corresponding author: anamaria.vatca@usamvcluj.ro

Received: November 14, 2023; accepted: February 25, 2024

ABSTRACT

This paper primarily aims to address aspects related to regional development in Romania, which has become a challenge both at the national and global level. The way regional development is approached is different, depending on the particularities of each country, as well as the governing ideologies in each state. However, the goal is common, one aims at shaping a competitive and strong economy, a high standard of living, access to education and medical assistance. The purpose of the paper is to compare socio-economic indicators from two time periods, to highlight similarities and differences and to observe their evolution within a 10-year timespan. The method used is Principal Component Analysis (PCA) to identify groups of statistical indicators (variables) that explain the level of development within a sample, in our case the development regions of Romania. In the context of regional development, PCA can be applied to extract relevant information about economic, social or demographic variability in different regions of the country. Several objectives of regional development analyzed through the lens of this method were pursued: identification of regional disparities, optimization of resource allocation, evaluation of the determining factors of economic development, monitoring of the impact of regional policies. Following the PCA analysis, it was observed that regional development cannot be evaluated by means of variables belonging to a single sector. This method allowed the identification of the variables that have the greatest weight describing economic, social, educational and infrastructure aspects. PCA was applied for two time periods and similar situations were obtained, the first main component called by the authors the economic statistics includes for both periods the indicators: Gross average nominal salary gain, GDP by development regions, Employees from research and development activity at the end of the year, Beds in health facilities (state and private). The second component was called social statistics and includes the indicator Employment rate in working age (15-64 years), and the third component called agricultural statistics includes the indicator Population served by the public water supply system. As a conclusion, we consider that the situations are not overlapping because there are indicators that have undergone changes during the 10 years, which is expected and normal in regional development. For example, the number of graduates decreased due to the continuously decreasing birth rate, the vacancy rate decreased in 2018-2019 compared to 2008-2009 because Romania went through the financial economic crisis.

Keywords: regional development, principal components, correlation matrix, PCA

REZUMAT

Această lucrare îşi propune în primul rând să urmărească aspecte ce ţin de dezvoltarea regională în România care a devenit o provocare atât la nivel naţional cât şi mondial. Modul în care este abordată dezvoltarea regională este diferit, în funcţie de particularităţile fiecărei ţări, precum şi de ideologiile de guvernare din fiecare stat. Cu toate acestea scopul este comun, se doreşte o economie competitivă şi puternică, un nivel de trai cât mai ridicat, acces la educaţie şi asistenţă medicală. Scopul lucrării este de a compara indicatori socio economici din două perioade de timp, pentru e evidenţia asemănări, deosebiri şi a observa evoluţia acestora în decurs de 10 ani. Metoda folosită este analiza în componente principale (PCA) pentru a identifica grupuri de indicatori statistici (variabile) care explică nivelul de dezvoltare în cadrul unui eşantion, în cazul nostru regiunile de dezvoltare ale României. În contextul dezvoltării regionale, PCA poate fi aplicată pentru a extrage informații relevante despre variabilitatea economică, socială sau demografică în diferitele regiuni ale țării. S-au urmărit câteva obiective ale dezvoltării regionale analizate prin prisma acestei metode: identificarea disparităţilor regionale, optimizarea alocării resurselor, evaluarea factorilor determinanţi ai dezvoltării economice, monitorizarea impactului politicilor regionale. In urma analizei PCA s-a observat că dezvoltarea regională nu poate fi evaluată prin intermediul unor variabile care aparţin unui singur sector. Această metodă a permis identificarea variabilelor care au ponderea cea mai mare descriind aspecte de ordin economic, social, educaţional, de infrastructură. S-a aplicat PCA pentru două perioade de timp şi s-au obţinut situaţii asemănătoare, prima componentă principală denumită de autori *statistica economică* cuprinde pentru ambele perioade indicatorii: Câştigul salarial nominal mediu brut, PIB pe regiuni de dezvoltare, Salariaţii din activitatea de cercetare dezvoltare la sfârşitul anului, Paturi în unităţile sanitare (stat şi privat). Componenta a doua a fost denumită *statistică socială* şi cuprinde indicatorul Rata de ocupare în vârsta de muncă (15-64 de ani), iar componenta a treia numită statistică agricolă cuprinde indicatorul Populaţia deservită de sistemul public de alimentare cu apă. Ca o concluzie, considerăm că situaţiile nu sunt identice deoarece există indicatori care în decursul celor 10 ani au suferit modificări, lucru care este de aşteptat şi normal în dezvoltarea regională. De exemplu, numărul absolvenţilor a scăzut datorită natalităţii în continuă scădere, rata locurilor de muncă vacante a scăzut în 2018-2019 comparativ cu 2008-2009 deoarece România a treversat criza economică financiară.

Cuvinte cheie: PCA, matricea de corelaţie, dezvoltare regională, componente principale

INTRODUCTION

Regional development is a concept that aims to boost and diversify economic activities, stimulate private sector investment, contribute to reducing unemployment and, last but not least, improve living standards and welfare.

Regional development policy is a set of measures planned and promoted by local and central public administration authorities and bodies, in partnership with various stakeholders (private and public actors, volunteers). The aim of this policy is to provide a dynamic and sustainable growth, through the efficient use of regional and local potential, in order to improve the living conditions of people (Popa and Popescu, 2013).

The main areas that can be targeted by regional policies are public and private sector development, labour market, attracting investment, technology transfer, improving infrastructure, environmental quality, rural development, health, agriculture, education, training, and culture.

Regional development policy is one of the most important and complex policies of the European Union. Sustainable regional development is a concept that raises challenges for the regions of the European Union through the 1977 Amsterdam Treaty (Sedlacek and Gaube, 2010). Its aim is to close the economic and social gaps between the various regions of Europe, which means that a large number of institutional stakeholders are involved in the development and implementation of such policies. The European Commission is directly responsible for preparing and ensuring the implementation of the EU’s regional development policy. Its role is to initiate and complete new legislation in this field, and to ensure that the measures thus adopted are implemented by the Member States.

On the Romanian territory, ancient initiatives were taken to set up territorial entities, namely: between 1859 and 1918 there were several historical regions, between 1938 - 1940 there were established incipient regions called Lands, between 1950 - 1968 these areas were called Regions, between 1947 and 1989 an *Authoritarian regime* that adopted a centralized territorial planning system was in place, in 1995 the Abscission Agreement between Romania and the EU was signed and since 1998 the Romanian regions have been called Development Regions (The Government of Romania, 2017).

To date, Romania developed National Sustainable Development Strategy versions for the time horizons 2008, 2020, and 2030, covering all 17 Sustainable Development Goals (SDGs) of the Sustainable Development Agenda. At national level, the key objectives can only be achieved by having a detailed knowledge of the state of play at the level of each administrative region. In order for a state to achieve sustainable development, effective tools are needed to enable the implementation of many strategies and programs. These programs must adopt a top-down implementation scheme, from the highest level to the bottom of the pyramid of administrative structures (Law 315/2004).

Through this paper we aim to make a comparative analysis of the level of regional development in Romania within a 10-year timespan, starting from indicators selected by authors in the economic, social, education, health, and agriculture which are important regional development related fields. By this approach, we want to check the applicability of the principal component analysis in the field of regional development, and to find the pros and cons of this analysis method in the above-mentioned field.

In the first stage, authors selected the variables that will be the subject of the article in such a way as to cover as many areas of regional development as possible. The variables (i.e., socio-economic indicators in our case) should be correlated with each other in order to be able to apply the principal component analysis successfully. The next step is to identify the principal components according to Kaiser's criteria and name them. The names of these components must be chosen in such a way as to best describe the nature of the variables with which they are related. The paper shows that this method can be successfully applied in the field of regional development, and similar situations were obtained for the two periods studied, yet there are differences that in the opinion of the authors are because these socio-economic indicators have change over time.

The paper is structured in several sections: the first part i.e., the introduction presents the concept of regional development and the historical perspective, the aim of the paper, and the relevance of the chosen topic. Section 2 presents the main theoretical considerations on regional development at the EU and national level. There are also mentioned here the most important studies on sustainable regional development in Romania. Section 3 presents the main techniques used in the analysis, data presentation and related methodology. Section 4 is dedicated to the results and discussions, and the last part contains the paper's conclusions.

LITERATURE REVIEW

Regional development has many definitions, one of the most well known being given by the United Nations World Commission on Environment and Development which states that "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (WCED, 1978).

In their book, the authors address issues related to what kind of local and regional development should be carried out and to whom it is addressed. Stakeholders and institutions in all regions are looking for development and prosperity in a constantly changing world, driven by increasingly globalized capitalism, global financial instability, climate change and a lack of resources (Pike et al., 2016).

Principal component analysis (PCA) was first used in 1901 by Pearson and later developed by Hotelling in 1931, for this reason, it is also called the "Hotelling transform" or "Karhunen-Loeve (KL) Method" (Hotelling, 1931). It is one of the most common analysis methods of multivariate data. It was first applied in psychology, to evaluate variables such as intelligence, and affectivity, and will later be applied in as many fields as possible to extract relevant information from various data sets. It has applications in areas such as facial recognition and image compression (XLSTAT, 2024).

It is successfully applied in financial analysis following which a score function can be constructed to testing the financial health of companies (Sabău-Popa et al., 2020). The method also found its application in agriculture, by diminishing the amount of data, thus, being able to analyse variables that represented the land area cultivated with various crops in the development regions of Romania, and only two factors concentrating over 80% of the information provided were maintained (Rotaru et al., 2012).

In their paper, Petrişor et. al. (2012), introduced the use of PCA based methodology together with Geographic Information System (GIS) modelling to assess the level of development within the territorial subunits of different dimensions of a given region, to test the hypothesis according to which the level of development cannot be accurately described from a single standpoint, be it economic, social, cultural, etc. From a methodological standpoint, the approach is a useful tool in decision-making for underdeveloped areas, as long as there is a consistent database.

In his doctoral thesis, Andrei Gina seeks to provide a new perspective using new techniques of multivariate analysis (PCA), in the processing of geophysical data, and the results obtained are then compared with those of conventional interpretation (Andrei, 2015).

In their paper, Lulu et al. (2011), selected statistical data from several cities within the Henan Province in China, and with the help of the principal component analysis, they obtained scores and rankings of urban competitiveness in this province, noting huge differences between the north and the south. In that sense, results are obtained for cities according to their different stages of development that would lead to the improvement of their competitiveness. Also in China, specifically in Fujian Province and starting from the dependent variable gross domestic product and using PCA, Wangzi Xu analyses in his paper the development level for nine cities of the province above and provides solutions for the local government of Fujian for a harmonious development of the province (Xu, 2021).

Moreover, the method is successfully applied in genetics, being commonly used for data analysis in single-nucleotide polymorphism (SNP) in order to detect population structure and potential abnormal values (Abraham and Inouve, 2014).

In Romania, economic policy priorities encompass closing economic and social gaps between regions, as well as between Romania and other EU Member States.

In his paper, Bălăcescu et al. (2016) consider that it is essential to find those factors that play a positive or negative role in regional development. The influence of endogenous factors on the economic development of the eight development regions of Romania is subjected to analysis. One conclusion of the paper is that GDP per capita is a key indicator for measuring regional disparities; however, a single factor is not sufficient.

Moreover, in terms of economics, in their paper, Manole et al.(2015) analyse the influence of the following factors: the number of tourists that characterize tourism and the road network density that features transportation. The study shows that nominal GDP is a factor that directly influences the density of the national road network, while GDP per capita and the number of employees per 1,000 people are relevant factors that influence tourism.

Boldea et al. (2012) performed an analysis of the level of regional development in Romania using indicators such as GDP, productivity and employment rate. They found that there are differences between the regions of Romania, although the regions with lower incomes benefit from a higher percentage of the European investment and structural funds. In Romania, as in most EU countries, there has been noticed an increased development pace in the counties around the capital, due to investments that prefer to develop as many areas as possible.

Due to foreign direct investment, there are visible discrepancies between the developing regions of Romania, with most of the large investments being made around the country's capital i.e., the city of Bucharest (Zaman et al., 2011).

Another study conducted in the development regions of Romania shows a direct link between the evolution of GDP per capita and spending on research and development. Thus, the most developed regions invest more in R&D than the least developed ones (Dachin and Postoiu, 2015).

Pîrvu et al. (2018) characterizes the development regions in Romania and measures the territorial development gaps, starting from the aspiration of Romania and the EU to promote more economic and social policies tailored to the different particularities of the regions. In that regard, a synthetic index was obtained by combining several sub-indices (economy, health, standard of living, and environment). As of 1998 and making use of cluster analysis, models of regional development have been found over time.

The evaluation of the progress made by the South-Muntenia Region towards sustainable regional development in the period 2010-2017 was assessed by Davidescu et al. (2020). Using PCA, the main determinants of regional development in this region were found. The empirical results highlighted the importance of the business environment, public service infrastructure, education, and social protection as determinants in regional development.

MATERIALS AND METHODS

Romania is organized into eight main development regions, comprising 42 counties together with the Municipality of Bucharest totalling 320 cities (of which 103 municipalities), 2,861 communes, and 12,957 villages.

Development regions of Romania are statistical territorial units made up of four to seven counties (with the exception of the Bucharest-Ilfov Region), through the free association of the County Councils (Figure 1).

Figure 1. Regions in Romania (source: Wikipedia)

They correspond to the NUTS II level according to the EUROSTAT classification and are the framework for collecting specific statistical data at the NUTS II territorial level.

Development regions are the framework for the development, implementation, monitoring and evaluation of regional development policies, including the regional development strategies and economic and social cohesion programmes (European Commission). The indicators that measure the level of development are many, therefore we selected 17 indicators in the first phase, 12 remaining for the subsequent phase, while trying to cover the economic, social, educational, and agricultural sectors. These indicators are intended to provide an overview of Romania's progress towards reaching sustainable development objectives in relation to the targets it assumed.

Two time series were considered in the paper i.e., 2008-2009 and 2018-2019, respectively, the first representing the period immediately following Romania's accession to the European Union and the beginning of the economic-financial crisis that affected the whole of Europe, implicitly our country, and the second was chosen in order to prevent entering data from the year 2020, a period that marks the onset of the global health crisis due to the Covid-19 pandemic. The latest data show that the Covid 19 pandemic triggered an unprecedented crisis, halting economic growth and standard of living, with the poorest and most vulnerable communities being the most affected (Administrative Capacity Operational Programme - Competence Makes a Difference).

12 variables were considered in the analysis: Gross average monthly wages, GDP by development regions, the Employment rate (15-64 years), Research and development employees at the end of the year, Job vacancy rate, Total unemployment rate, Average monthly pension of state social insurance pensioners and farmer pensioners, Total number of graduates, Number of beds in (public and private) hospitals, Area under cultivation with principal crops, Number of individual farms, Population served by the public water supply system (Table 1).

The data were taken from the website of the National Institute of Statistics of Romania (1998-2018), and we performed the statistical processing thereof using the SPSS software, version 19 (Statistics Package for the Social Sciences).

The Principal Components Analysis method includes both a preliminary statistical processing of the observation data and a mathematical and numerical processing thereof.

Since it is about a statistical technique used to reduce data, it is applied to a single set of variables when the researcher wants to find out which variables from the set form coherent subsets that are relatively independent of each other. In a first stage, the associations (correlations) between the variables and the determination of the less variables (latent variables) that lie behind the measured variables (several) are highlighted. These latent variables are called factors or components, hence the name of the method.

Table 1. Name of variables

Gross average monthly wages (Lei) Câstigul salarial nominal mediu lunar brut (Lei)	V1
GDP by development regions (mil.Lei) PIB pe regiuni de dezvoltare (mil lei)	V2
Employment rate (15-64 years) (%) Rata de ocupare in varsta de munca (15-64 ani) - total	V3
Research and development employees at the end of the year (no.) Salariatii din activitatea de cercetare-dezvoltare la sfarsitul anului	V4
Job vacancy rate (%) Rata locurilor de munca vacante (%)	V5
Total unemployment rate (%) Rata somajului - total (%)	V6
Average monthly pension of stat social insurance pensioners and of farmer pensioners Pensia medie lunara de asigurari sociale de stat si agricultori	V7
Total number graduates (no.) Total absolveti	V8
Number of beds in (public and private) hospitals (no.) Paturi in unitatile sanitare (stat si privat)	V9
Area number cultivation with principal crops (ha) Suprafata cultivata cu principalele culture	V10
Number of individual farms (no.) Nr. exploatatii agricole individuale	V11
Population served by the public water - supply sistem (no.) Populatia deservita de sistemul public de alimentare cu apa.	V12

This mathematical procedure transforms a number of correlated variables into a number of uncorrelated variables called principal components. The first principal component represents most of the variability, with each successive component representing as much of the remaining variability as possible (Jolliffe et al., 2016; Paul et al., 2013).

It can be considered as a projection method that projects observations from an n-dimensional space with *n* variables into an m-dimensional space with *m* variables ($m < n$), so that the maximum amount of information is maintained. When analysed from a mathematical perspective, we consider a lot of variables x_1, x_2, ..., x_n, we want to determine a new set of variables c_1, c_2, ..., c_m, where $c_i = w_{i1}x_1 + w_{i2}x_2 + ... + w_{in}x_n$, provided the condition $m < n$.

The inversion requirement is also necessary, i.e., the possibility to find the variables x with the help of the components, i.e., $x_i = v_{i1}c_1 + v_{i2}c_2 + ... + v_{im}c_m$. By reducing these components, we try to reduce the number of variables without losing the variance of the initial variables.

For this, a new variable *Z* is introduced, as a linear combination of the initial variables, as follows:

$Z = a_1x_1 + a_2x_2 + ... + a_nx_n$, where a_1, ..., a_n are weights associated with the initial variables.

The previous equation is only apparently similar to a regression equation, since no observed values are known for the variable *Z*, there is no free term and no errors (residues).

The principal components analysis determines those weights ai that maximize the variance of the variable *Z*. As the variance can tend to infinity for values of the conveniently chosen weights, the method determines only the weights subject to the restriction that the vector a is normalized, i.e.

$$\sum_{i=1}^{n} a_i^2 = 1$$

Once the weights ai have been calculated, the variable *Z* is called the f*irst principal component*.

Denoting with *C* the covariance (correlation) matrix of the variables *X*, in fact by transforming the analysis data into principal components $C = X'X$, it results that the dispersion of *Z* is $a'Ca$. It is aimed at maximizing the variance of *Z* provided the restriction $a'a = 1$. The general problem below is thus reached:

$max\ a'X'Xa$ provided the restriction $a'a = 1$

The method of Lagrange multipliers will seek for the maximum of the function $F(a) = a'Ca - \lambda(a'a - 1)$ from where it results, as in the general method, that a is an eigenvector of the matrix *C* corresponding to the eigenvalue λ and $a'Ca = \lambda$. Since, $Var\ (Z) = a'Ca$, it results that $Var\ (Z) = \lambda$, i.e., a is the eigenvector that corresponds to the largest eigenvalue λ.

The second principal component is defined as the linear combination of the variables *X* with the following largest variance: $Z_2 = a_{12}x_1 + a_{22}x_2 + ... + a_{p2}x_p$.

This leads to the second eigenvalue in terms of size, etc. It is worth mentioning that a_{ij} represents the share of the variable i in the principal component with the number *j*.

A consequence of the fact that the variances of the principal components are the eigenvalues while the weights (coefficients of the linear combinations) are the eigenvectors is that the obtained factors (i.e., the principal components) are uncorrelated with each other.

Thus, from the matrix expression $z = Ax$ of the principal components and from the fact that the matrix of eigenvectors is orthogonal, $A'A = I$, it results $A'z = A'Ax = Ix = x$, i.e., the initial variables can also be expressed as linear combinations between the principal components. By denoting with *Czz* the covariance matrix of the principal components, the previous relation generates $C = A'CzzA$, from where, using the known result $C = A'\Lambda A$, where Λ is the diagonal matrix of eigenvalues, it follows that *Czz* is a diagonal matrix, i.e., all the principal components are uncorrelated with each other. It is thus observed that the transition to the principal components eliminates data redundancy.

In a nutshell, the first principal component extracted is a maximum share of the total variance among the observed variables. This means that the first component will be correlated with some of the observed variables, and it can be correlated with many of them. The second component has two characteristics, namely it will explain a maximum share of variance in the data set that was not accounted for by the first component, this means that the second component will be correlated with some of the variables that do not manifest strong correlations with the first component. The other characteristic is that the second principal component is not correlated with the first component, and if the correlation between them were to be calculated, it would be zero.

The remaining components that are extracted in the analysis have the same two characteristics: each component has a maximum weight of variance in the observed variables that were not considered by the previous components and is not correlated with any of these components. A principal components analysis takes place in this way, with each of the following components there are decreasing weights of variance, this is also the reason why only the first components are retained and subjected to interpretation (Abdi et Williams, 2010).

RESULTS

Next, we analyse the coefficients correlation matrix to learn whether the calculated indicators are independent. To apply factor analysis, there must be sufficiently large correlations between the variables in order for the lower-dimensional data problem to make sense. Therefore, if one variable is not correlated with the others it will have to be removed from the analysis. At the same time, even very large correlations (multicollinearity) do not lead to easy-to-interpret results, the extreme situation being that of singularity, of the existence of perfectly correlated variables.

In the first stage of the analysis, the descriptive statistics (mean, standard deviation) for the 8 development regions and consecutive years (2008 and 2009, respectively 2018 and 2019) are computed (Table 2).

Variables in the paper appear in a variety of units of measurement, and because the standard deviation shows very large differences, the data will be standardized.

Standardization is the most common method used to normalize variables, i.e., converting variables to a common scale that entails a normal distribution (Paul et al., 2013).

The formula for data standardization is:

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma}$$

where z_{ij} is the standardized variable, x_{ij} is the initial variable in the sample *i*, and (x_j) is the mean value and σ is the standard deviation.

Before the data standardization, the correlation matrix was calculated for all the 12 variables in both cases (series 1 i.e., 2008-2009 and series 2 i.e., 2018-2019). Table 3 and Table 4 show strong positive correlations between *Gross average monthly wages (V1)* and *GDP by development regions (V2)* and *Research and development employees at the end of the year (V4)* for both series of time subjected to analysis in our paper. Moreover, a strong positive correlation between the values of indicators *GDP by development regions (V2)* and *Research and development employees at the end of the year (V4)* is found. Different correlations are found for the two series of time, for example in 2008-2009 a correlation of 0.528 between the *Employment rate (15-64 years) (V3)* and *Total graduates (V8)* while in 2018-2019 the correlation between the two variables is 0.921. This can be explained by the fact that the *Employment rate (15-64 years)* increased on average in 2018-2019 compared to 2008-2009 while the number of graduates is decreasing (the gap is about 25,000 people). The same situation is encountered between variables V3 and V9, with the number of hospital beds being lower in the years 2018-2019 compared to the previous time series. There is a negative correlation between the *Job vacancy rate (V5)* and the *Total unemployment rate (V6)*, in both situations.

Table 2. Descriptive Statistics

	N	2008-2009		2018-2019	
		Mean	Std. Deviation	Mean	Std. Deviation
V1	16	1740.3750	298.45041	4412.8125	741.03893
V2	16	66865.0250	28167.39753	113012.3563	53358.49211
V3	16	61.7813	2.98959	64.8063	5.05272
V4	16	5370.1250	5970.35562	5544.1250	6715.47429
V5	16	1.3669	.61919	1.1031	.37494
V6	16	6.1813	2.66439	3.2500	1.50510
V7	16	924.7500	96.94431	1671.1250	128.38997
V8	16	90225.2500	19746.87968	62805.5000	14280.71357
V9	16	17137.1875	3416.52984	16686.8750	3695.13152
V10	16	980136.0000	571994.44032	1075245.8750	576252.25013
V11	16	656388.7500	306429.54823	638028.1250	292875.20307
V12	16	1445448.1250	263765.50604	1702735.6250	296531.92934

Source: author's calculation

Table 3. Correlation Matrix for years 2008-2009

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
V1	1.000											
V2	0.905	1.000										
V3	0.216	0.213	1.000									
V4	0.937	0.980	0.304	1.000								
V5	0.143	0.303	0.457	0.310	1.000							
V6	-0.431	-0.631	-0.202	-0.616	-0.708	1.000						
V7	0.496	0.364	-0.241	0.368	-0.639	0.277	1.000					
V8	0.608	0.790	0.528	0.798	0.426	-0.499	0.104	1.000				
V9	0.425	0.663	0.464	0.666	0.260	-0.457	0.187	0.924	1.000			
V10	-0.534	-0.553	-0.019	-0.636	-0.183	0.467	-0.359	-0.466	-0.501	1.000		
V11	-0.717	-0.776	0.086	-0.798	-0.207	0.586	-0.448	-0.509	-0.489	0.913	1.000	
V12	0.369	0.617	-0.143	0.476	-0.029	-0.332	0.256	0.520	0.531	0.056	-0.233	1.000

Table 4. Correlation Matrix for years 2018-2019

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
V1	1.000											
V2	0.933	1.000										
V3	0.495	0.568	1.000									
V4	0.923	0.971	0.559	1.000								
V5	0.431	0.437	0.120	0.459	1.000							
V6	-0.686	-0.643	-0.140	-0.614	-0.841	1.000						
V7	0.865	0.686	0.166	0.633	0.276	-0.627	1.000					
V8	0.394	0.556	0.921	0.531	0.116	-0.139	0.064	1.000				
V9	0.658	0.790	0.849	0.775	0.304	-0.442	0.322	0.915	1.000			
V10	-0.700	-0.662	-0.346	-0.732	-0.416	0.646	-0.492	-0.309	-0.635	1.000		
V11	-0.835	-0.805	-0.156	-0.843	-0.647	0.836	-0.679	-0.127	-0.499	0.840	1.000	
V12	0.362	0.586	0.179	0.444	0.309	-0.463	0.277	0.425	0.507	-0.095	-0.336	1.000

Source: author's calculation

The next step in our research is the accurate selection of the number of principal components so as not to lose information after the application of PCA compared to the original list of characteristics. Table 5 show the eigenvalues for the 12 variables in both situations and the value of the Bartlett's Test of Sphericity which compares Pearson's correlation matrix with the identity matrix. The value of the Bartlett's Test of Sphericity in both situations (299,662, Sig = 0.000 and 353,398, Sig = 0.000) is small enough to reject the hypothesis that the variables are uncorrelated. These values show the presence of at least one common factor, which substantiate the application of PCA (Cărbureanu, 2010; Saporta and Stefănescu, 1996).

According to Kaiser's Criterion (superunit value criterion) which states that only components with eigenvalues greater than 1 can be retained in the analysis, it is observed that in the first case 4 components are maintained while in the second three components are maintained (Kaiser, 1960). In addition, this KMO index is used to show the validity of the analysis, being relevant when it has values between 0.5 and 1 (for the first series the value is 0.58 and for the second one the value is 0.55 (Snedecor et al., 1989).

The first principal component explains 51.6% of the information contained in the correlation matrix (first case) and 59.1% in the second case. The second principal component explains 19.4% in the case of 2008-2009 time series and 18.4% in the case of 2018-2019 time series. From the last column (Cumulative%) we read how much of the total variance is explained by retaining the four components i.e., in the first case about 92%, namely three components in the second case (86%), which is a very good result for this analysis.

The number of principal components maintained is also highlighted by making eigenvalues graphs, but these were not added in our paper (Cattell, 1996; Hatcher et al. 1994; Freudenberg 2003).

The study continues with the calculation of the factor matrix for the principal components resulting from the analysis. The matrix factor is a very important one since its elements also known as factor loadings are the correlation coefficients between the original variables (rows) and the principal components (columns).

Table 5. The Kaiser–Meyer–Olkin (KMO) test and eigenvalues of the principal components

Years 2008-2009				Years 2018-2019			
KMO			0.584	KMO			0.552
Bartlett's Test of Sphericity	Approx. Chi-Square		299.662	Bartlett's Test of Sphericity	Approx. Chi-Square		353.398
	Df Sig.		66 .000		Df Sig.		66 .000
Component	Initial Eigenvalues			Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %		Total	% of Variance	Cumulative %
1	6.193	51.611	51.611	1	7.095	59.122	59.122
2	2.339	19.488	71.099	2	2.215	18.459	77.581
3	1.376	11.463	82.562	3	1.056	8.803	86.384
4	1.074	8.95	91.511	4	0.869	7.238	93.621
5	0.675	5.627	97.138	5	0.445	3.712	97.333
6	0.191	1.591	98.729	6	0.259	2.16	99.493
7	0.112	0.932	99.662	7	0.039	0.324	99.817
8	0.029	0.243	99.905	8	0.013	0.111	99.928
9	0.006	0.051	99.955	9	0.007	0.055	99.983
10	0.003	0.023	99.978	10	0.001	0.011	99.994
11	0.002	0.018	99.996	11	0.001	0.004	99.998
12	0	0.004	100	12	0	0.002	100

Source: author's calculation

A better view of the data is obtained after "rotating it". We used the Varimax rotation which minimizes the number of variables with high loads on each factor, which simplifies the interpretation of the factors.

The loading matrix was also displayed after data rotation, and the loading differences of the factors are thus better highlighted. Factor loads are the basis for naming factors, an important aspect of factor analysis. A factor, as a latent variable, should bear a name to be understood, used, referred to, and so on.

The load structure of a factor can provide suggestions in that regard, and loads higher than 0.5 are considered important, while those below 0.4 are considered low (are written in bold in Table 6).

High-load variables are a combination of initial variables that determine the factor, and therefore its name, which must have a suggestive name (Table 6).

We may notice strong correlations between variables in the graphical representation of the principal components, after which the nature of each component is determined. In Figure 2 we can highlight the variables correlated with the principal components, namely:

Component 1 (marked in red) strongly correlates with the variables V1 - Gross average monthly wages, V2 - GDP by development regions, V4 - Research and development employees at the end of the year, V8 - Total graduates, V9 - Beds in hospitals, for the time series 2008-2009.

Table 6. Rotated Component matrix

	Years 2008-2009				Years 2018-2019		
	Component				Component		
	1	2	3	4	1	2	3
V1	0.841	-0.253	-0.03	0.086	0.931	-0.119	-0.249
V2	0.961	-0.101	0.091	-0.121	0.959	0.059	-0.029
V3	0.311	0.607	0.265	0.634	0.584	0.735	-0.122
V4	0.971	-0.08	0.006	0.031	0.948	0.035	-0.127
V5	0.386	0.83	-0.244	-0.148	0.58	-0.397	0.478
V6	-0.691	-0.444	0.268	0.406	-0.78	0.466	-0.29
V7	0.308	-0.884	0.097	0.236	0.694	-0.37	-0.297
V8	0.865	0.228	0.326	0.156	0.573	0.796	0.121
V9	0.786	0.139	0.351	0.174	0.83	0.524	0.051
V10	-0.699	0.194	0.572	-0.202	-0.779	0.161	0.31
V11	-0.828	0.281	0.456	0.061	-0.863	0.444	0.074
V12	0.508	-0.183	0.626	-0.553	0.523	0.113	0.665

Extraction Method: Principal component analysis.
Rotation Method: Varimax with Kaiser Normalization
Loads greater than 0.5 are considered important, they are written in bold

As for the 2018-2019 time series, the situation is similar except one difference i.e., the component 1 strongly correlates with V1 - Gross average monthly wages, V2 - GDP by development regions, V4 - Research and development employees at the end of the year, V5 – Job vacancy rate, V7 - Average monthly pension of state social insurance and farmers, V9 - Beds in hospitals (Figure 3).

We could call this component 1 as a component that largely describes economic statistics by development regions, but there are also indicators related to education (V8) and health infrastructure (V9).

Component 2 (marked in blue) correlates well with the variables *V3 - Employment rate* and *V8 - Total graduates* for the time series 2018-2019, and with *Variable V5 - Job vacancy rate* for the time series 2008-2009. For these years, in Table 6 Variable V3 has a high load on the second component i.e. 0.607 compared to 0.634 for component 4, therefore it should be considered in the analysis as being correlated with component 4; however, comparing the two time series and since the time series 2018-2019 has only 3 components, while the difference being very small, we will consider the Variable V3 as correlated with component 2.

Component 2 includes social statistics related indicators (i.e., labour force, standard of living).

Component 3 (marked in yellow) strongly correlates with the variables *V10 - Area under cultivation with principal crops* and with *V12 - Population served by the public water-supply system* (years 2008-2009) and with the Variable V12 for the years 2018-2019. This component contain the variable related to agriculture. Using this type of factor analysis, we can obtain useful information on the factors that have a great influence on regional development, giving statisticians the opportunity to follow the upturns or downturns of the level of regional development. The

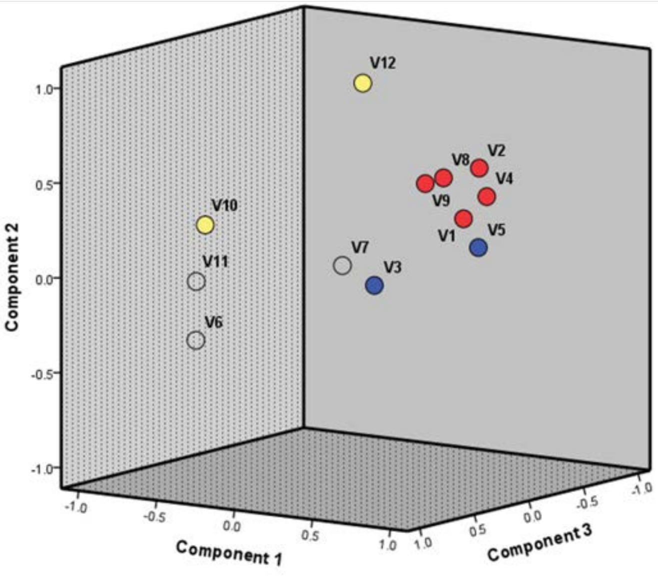

Figure 2. Graphical representation of principal components for the years 2008-2009

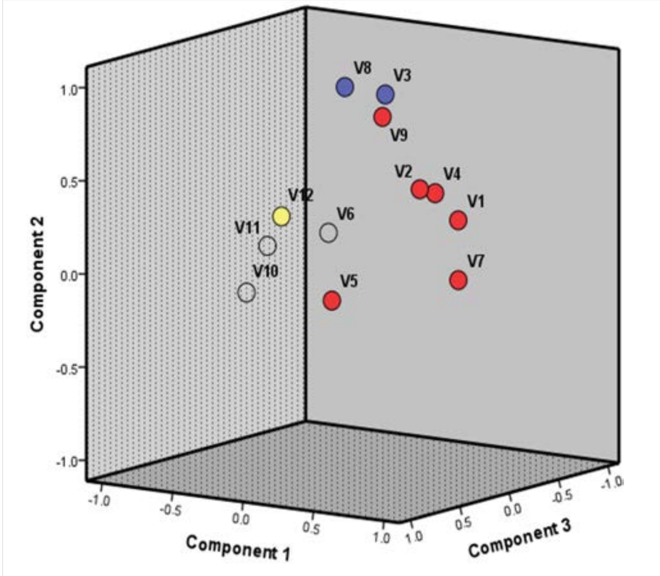

Figure 3. Graphical representation of principal components for the years 2018-2019

immediate result of this type of analysis is finding and then optimizing those factors responsible for regional development.

A brief representation of the principal components for the two time series and for the 12 variables is presented in Table 7.

Next, we carried out a comparative analysis of the indicators for each region over each of the time series. We generated "100% Stacked Line" charts in which the percentage of 100% was considered for the Bucharest-Ilfov Region, this being the region for which most indicators record the highest values. Figure 4 and Figure 5 provide such an analysis, and we may notice that the trend is broadly the same, highlighting similarities and differences materialized in the following conclusions:
GDP by development regions have increased significantly between the two time series analysed (%), an increase of about 59% being recorded. The economic and financial crisis affected the real economy in 2008, the value for this indicator decreasing by 1.65% in 2009 compared to 2008. It is observed that the Bucharest-Ilfov Region has the highest GDP allocated, which leads to additional investments in infrastructure and public services, at the opposite pole is the South-West Oltenia Region. Analysis of GDP by region can help identify economic areas in which certain regions have competitive advantages. The promotion of these areas can contribute to the increase of regional competitiveness and the creation of specialized areas in certain industries.

The total unemployment rate increased from 4.5% in 2008 to 7.82% in 2009, which is also explained by the effects of the economic crisis. Due to the fragile economy recovery that followed, there is a decrease of this indicator from an average of 6.18% in the development regions in 2008-2009 down to 3.25% in the 2018-2019 period.

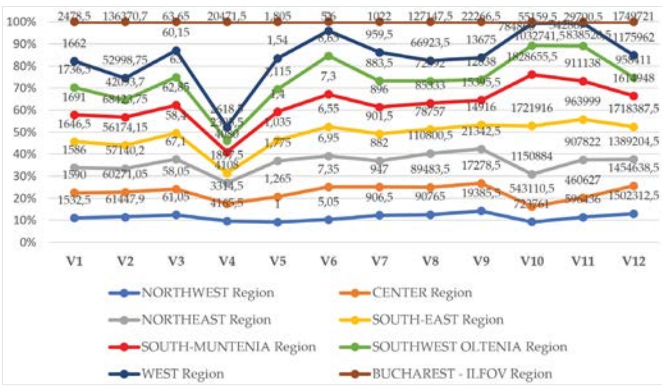

Figure 4. The evolution of the 12 variables on the development regions (2008-2009)

DOI: /10.5513/JCEA01/25.2.4176

Table 7. The main components of regional development

Statistical indicators (variables) analysed	2008-2009			2018-1019		
	Economic statistics [1]	Social statistics [2]	Agricultural statistics [3]	Economic statistics [1]	Social statistics [2]	Agricultural statistics [3]
Gross average monthly wages – V1	0.841			0.931		
GDP by development regions – V2	0.961			0.959		
Employment rate (15-64 years) (%) - V3		0.607			0.735	
Research and development employees at the end of the year (no.) - V4	0.971			0.948		
Job vacancy rate (%) - V5		0.830		0.580		
Total unemployment rate (%) - V6						
Average monthly pension of stat social insurance pensioners and of farmer pensioners - V7				0.694		
Total number graduates (no.) - V8	0.865				0.796	
Number of beds in (public and private) hospitals (no.) - V9	0.786			0.830		
Area number cultivation with principal crops (ha) - V10			0.572			
Number of individual farms (no.) - V11						
Population served by the public water - supply sistem (no.) - V12			0.625			0.665

Source: author's calculation.

[1] The data in this column refer to component 1- Economic statistics (the one in red)
[2] The data in this column refer to component 2 – Social statistics (the one in blue)
[3] The data in this column refer to component 3- Agricultural statistics (the one in yellow)

The lowest unemployment rate in 2009-2009 was recorded in the Bucharest-Ilfov region i.e., 2% and the highest in the South-West Oltenia Region i.e. 8.65%. In 2018-2019 period, the lowest unemployment rate was also recorded in Bucharest-Ilfov i.e., 1.15% and the highest also in the South-West Oltenia Region (i.e. 5.5%).

A low unemployment rate in a region can lead to increased income and quality of life for residents. Conversely, a high rate can contribute to poverty and lower living standards. In order to reduce the unemployment rate, it is essential to have effective education and training programs. Regional development can involve initiatives to provide resources and training opportunities that align with local labor market requirements.

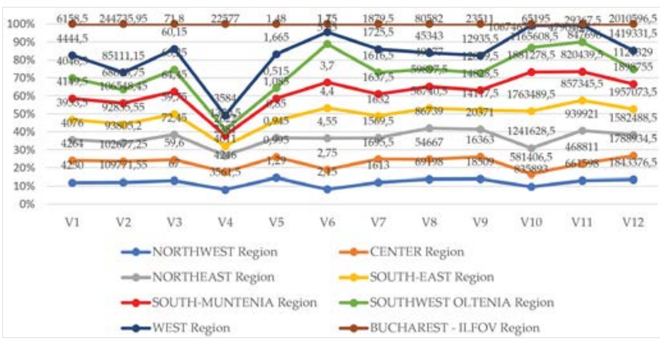

Figure 5. The evolution of the 12 variables on the development regions (2018-2019)

ISSN 1332-9049

The Total graduates indicator has faced a substantial decrease, the mean on the development regions in 2008-2009 being of 90,225 graduates, while in the period 2018-2019 it was only of 62,805 graduates (i.e. a roughly 30% decrease). This decrease is due both to the declining birth rate and to the problems related to the migration abroad of young people. Thus, while the number of births is steadily declining, the number of deaths and emigrants has been steadily increasing. Thus, Romanians are fewer and fewer from year to year. Low birth rates are the main cause of changing demographics.

The same decreasing trend is being observed in the case of the indicator Number of beds in (public and private) hospitals which has decreased on average by 2.62% (the mean in the development regions in 2008-2009 was 17,137 beds and in 2018-2019 this value was 16,686). The only regions that have recorded an increase in terms of this indicator is the Bucharest-Ilfov Region with a plus of 1,245 beds, explainable by the fact that here is the capital of Romania, where are the largest hospitals in the country, and the South West Oltenia Region with a number of plus 21 beds.

From the data presented in the paper, we may notice that the values for both indicators - the Area under cultivation with principal crops and the number of individual farms - are distributed according to the specifics of the regions. We mention that in the period 2008-2009 most individual farms were located in the South-East Region (18.3%), followed by the South-Muntenia Region (17.4%) and the North-East Region (17.3%) and they also own most of the agricultural area used. In the period 2018-2019, the number of individual farms has decreased compared to 2008-2009 by about 2.8%, the North East Region being the only one that recorded a higher number of farms, followed by the South-East and South Muntenia Region, both the latter having decreased in terms of absolute number of individual farms.

CONCLUSIONS

The paper tried to introduce the PCA based to notice an evolution of the regional development level in Romania, starting from the hypothesis that the level of development cannot be accurately described individually only from an economic, social, cultural point of view, etc. The approach shows, in addition to its usefulness as a research method, it could be an important tool in regional decision-making by finding those indicators that lead to progress in sustainable development.

By applying this method, we may follow the indicators (variables) that strongly correlate with the principal components highlighted in the PCA. Thus, for the two time series considered there are similarities described by the three components: component 1 which describes economic indicators, component 2 which describes indicators of a social nature, and component 3 which target agricultural statistics specific indicators.

We should bear in mind that the two periods considered in our paper i.e., the period immediately following Romania's accession to the EU (which coincides with the beginning of the economic crisis), in which a full connection of our country to the new philosophy of the European Union in terms of sustainable development was attempted. During this period, Romania had to recover considerable gaps compared to the other Member States of the European Union, it had an economy based on intensive consumption of resources, a society and a public administration still searching for an own vision.

The second period addressed in the paper is Romania after 10 years in which the Romanian society has felt a significant improvement in terms of standard, this being observed by the Gross average monthly wages (V1), GDP by development regions (V2), Average monthly pension of state social insurance pensioners and of farmer pensioners (V7), Population served by the public water-supply system (no) (V12) which increased considerably, through the Total unemployment rate (V6) which have decreased in 2018 compared to 2008 at the level of all the eight development regions of Romania.

Therefore, some conclusions can be generalized regarding the regional development of Romania:

1. From the point of view of regional disparities, the data show that there is considerable inequality between Romania's regions. Regions in the west of the country, such as Timiş or Cluj, often have higher GDP per capita and a higher quality of life, while regions in the east, such as Moldova or Muntenia, face greater economic challenges.
2. From the point of view of urbanization and economic development, the Bucharest-Ilfov area was a center of significant economic growth, due to the presence of the capital and foreign investments. This concentration can accentuate regional divisions and negatively influence the development of other regions.
3. Access to education and health, there are significant differences between regions in terms of access to education and health services. More developed regions often have stronger education and health systems, which contribute to economic growth and quality of life.
4. Rural development, rural regions, especially those in the east of the country, face significant challenges in terms of infrastructure, public services and access to economic opportunities. Promoting rural development can be a crucial aspect for reducing regional disparities.
5. The impact of regional policies, the implementation of various regional policies, including European funds intended for regional development, had varied impacts. Monitoring the effectiveness of these policies and adjusting them according to the needs of the regions can help reduce the gaps.
6. Sustainability and environment, regional development should also consider sustainability aspects, such as natural resource management and environmental protection. Regions that adopt sustainable practices can benefit in the long run.

Discussions on regional development in Romania must be integrated in the broader context of sustainable development goals, including equitable economic growth, improving the quality of life and reducing disparities between regions. It is important to emphasize collaboration between different levels of government, the private sector and civil society to address complex challenges and promote sustainable and inclusive development.

Principal component analysis (PCA) is a technique used in data analysis to reduce dimensionality and highlight important patterns in complex data sets. In the context of Romania's development areas, the application of PCA can have both advantages and disadvantages. Here are some of them:

Advantages of Principal Component Analysis (PCA):

1. Dimensionality reduction: PCA allows reducing the number of variables in a data set without losing significant information. This can facilitate data analysis and interpretation, especially for large and complex datasets.
2. Highlighting Latent Patterns: PCA identifies and highlights latent patterns or structures in the data. This can help to identify the key factors influencing the development of an area and to understand the correlations between different variables.
3. Simplifying interpretation: dimensionality reduction can make data analysis more accessible and easier to interpret, which can be useful in the context of decision-making in Romania's development areas.

Disadvantages of Principal Component Analysis (PCA):

1. Loss of detail: while PCA simplifies data, it can result in loss of fine details of information. This can be problematic in the context of regional development, where details such as the specific characteristics of certain areas can be critical to planning. It is observed that the variable V11, which represents the number of individual agricultural holdings, was not included in any component of the PCA method. The number of these agricultural holdings is the highest in the EU, but with all that, Romania cannot be considered more developed from an agricultural point of view.

2. Linear Assumptions: PCA is based on the assumption that the relationships between variables are linear. If the relationships are more complex or non-linear, PCA may provide an inadequate representation of the data.
3. Sensitivity to extreme data: PCA can be sensitive to extreme data (outliers), which means that unusual or atypical data points can significantly influence the results of the analysis.
4. Interpretation can be difficult: in some cases, the principal components obtained by PCA can be difficult to interpret in physical or economic terms. This can make it difficult to connect the results to practical regional development decisions.

Therefore, we intend to continue this scientific approach, for at least another period of time after 2020-2021, the latter period being affected by the Covid 19 health crisis, in order to follow these structural changes throughout the country.

REFERENCES

Abdi, H., Williams, L. J. (2010) Principal component analysis. Wiley interdisciplinary reviews: computational statistics, 2 (4), 433-459.

Abraham, G., Inouye, M. (2014) Fast principal component analysis of large-scale genome-wide data. PloS one, 9 (4), e93766. DOI: https://doi.org/10.1371/journal.pone.0093766.

Andrei G., (2015) Contribuţii la valorificarea superioară a datelor geofizice prin procedee moderne de analiză matematico-numerică. Teză de doctorat. Universitatea din Bucureşti, Facultatea de Geologie şi Geofizică, România.

Bălăcescu, A., Gogonea, R. M., Zaharia, M., Babucea, A. G. (2016) The influence of endogenous factors on the sustainable economic development at regional level in Romania. Cleaner Production and Green Economy, 4, 456-460.

Boldea, M., Parean, M., Otil, M. (2012) Regional disparity analysis: The case of Romania. Journal of Eastern Europe Research in Business & Economics, 1.

Cărbureanu, M. (2010) O metodă de analiză factorială aplicată în domeniul dezvoltării. Annals of the „Constantin Brâncuşi" University of Târgu Jiu, Economy Series, Issue 1/2010, pp. 187-194.

Cattell, R. B. (1996) The screen test for the number of factors. Multivariate behavioral research, 1 (2), 245-276.

Dachin, A., Postoiu, C. (2015) Innovation and regional performance in Romania. Theoretical & Applied Economics, 22 (2).

Davidescu, A. A., Apostu, S. A., Pantilie, A. M., Amzuica, B. F. (2020) Romania's South-Muntenia region, towards sustainable regional development. implications for regional development strategies. Sustainability, 12 (14), 5799. DOI: https://doi.org/10.3390/su12145799.

Freudenberg, M. (2003) Composite indicators of country performance: a critical assessment. Paris: Organisation for Economic Co-Operation and Development (OECD).

Hatcher, L., Stepanski, E.(1994) A step by step approach to using the SAS system for univariate and multivariate statistics. Cary, NC: SAS Institute Inc.

Hotelling, H. (1931) The economics of exhaustible resources. Journal of political Economy, 39 (2), 137-175.

National Institute for Statistics of Romania (1998-2018) Available at: http://statistici.insse.ro:8077/tempo-online [Accessed 1 August 2023].

Jolliffe, I. T., Cadima, J. (2016) Principal component analysis: a review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374, 20150202. DOI: https://doi.org/10.1098/rsta.2015.0202.

Kaiser, H. F. (1960) The application of electronic computers to factor analysis. Educational and psychological measurement, 20 (1), 141-151.

Law 315/2004 on regional development in Romania. Available at: https://www.fonduri-structurale-europene.ro/legislatie/legea-315-2004 [Accessed 5 November 2023].

Lulu, S., Zhengsheng, Z., Feng, L., Xiaoya, W., Lijuan, X. (2011) Research on the Evaluation of Urban Competitiveness in Henan Province Based on Principal Components Analysis [J]. Henan Science, 4.

Manole, S. D., Tache, A. (2015) Panel data models for regional development. Urbanism. Arhitectura. Constructii, 6 (1), 71-82.

Paul, L. C., Suman, A. A., Sultan, N. (2013) Methodological analysis of principal component analysis (PCA) method. International Journal of Computational Engineering & Management,16 (2), 32-38.

Petrişor, A. I., Ianoş, I., Iurea, D., Văidianu, M. N. (2012) Applications of principal component analysis integrated with GIS. Procedia Environmental Sciences, 14, 247-256. DOI: https://doi.org/10.1016/j.proenv.2012.03.024

Pike, A., Rodríguez-Pose, A., Tomaney, J. (2016) Local and regional development (2nd Edition). London. Routledge. DOI: https://doi.org/10.4324/9781315767673

Pîrvu, R., Bădîrcea, R., Manta, A., Lupăncescu, M. (2018) The effects of the cohesion policy on the sustainable development of the development regions in Romania. Sustainability, 10 (7), 2577. DOI: https://doi.org/10.3390/su10072577

Popa, I., Popescu, D.I. (2013) The importance of innovative clusters' proliferation for sustainable economic growth of Romania. In: Proceedings of the 7th International Management Conference—New Management for the New Economy, Bucharest, Romania, 7–8 November, pp. 583–595.

Rotaru, A. S., Pop, I. D., Vâtcă, A., Cioban, A. (2012) Usefulness of Principal Components Analysis in Agriculture. Bulletin of the University of Agricultural Sciences & Veterinary Medicine Cluj-Napoca. Horticulture, 69 (2).

Sabău-Popa, C. D., Simut, R., Droj, L., Bențe, C. C. (2020) Analyzing Financial Health of the SMES Listed in the AERO Market of Bucharest Stock Exchange Using Principal Component Analysis. Sustainability, 12 (9), 3726. DOI: https://doi.org/10.3390/su12093726

Saporta, G., Stefanescu, V. (1996) Analiza datelor si informatica: cu aplicatii la studii de piata si sondaje de opinie, ISBN 973-9198-24-4. Editura Economica.

Sedlacek, S.; Gaube, V. (2010) Regions on their way to sustainability: The role of institutions in fostering sustainable development at the regional level. Environment, Development and Sustainability 12, 117–134. DOI: https://doi.org/10.1007/s10668-008-9184-x.

Snedecor, G. W., Cochran, W. G. (1989) Statistical Methods, eight edition. Iowa state University Press, Ames, Iowa, 1191 (2).

The Government of Romania (2017) Decision No. 229/2017. Available at: https://www.lege-online.ro/lrDECIZIE-229%20-2017-(190860)-(1).html [Accessed 28 June 2023].

Wikipedia. Available at: https://ro.m.wikipedia.org/ [Accessed 12 July 2023].

XLSTAT (2024) By LUMIVERO. Available at: https://www.xlstat.com/en/products-solutions/feature/principal-componentanalysis [Accessed 12 July 2023].

Xu, W. (2021) An Empirical Analysis of Regional Economic Development Level in Fujian Province Based on PCA. In: 6th International Conference on Social Sciences and Economic Development (ICSSED 2021), Atlantis Press, pp. 5-10. DOI: https://doi.org/10.2991/assehr.k.210407.002

Zaman, G., Vasile, V.; Panait, M., Croitoru, C., Enescu, G. (2011) Impactul investitiilor straine directe (ISD) asupra exporturilor si dezvoltării durabile în România. Rom. J. Econ., 33, 1–60. DOI: https://doi.org/10.3390/su12145799