Romanian Journal of Ecology & Environmental Chemistry, 4(1), 2022

https://doi.org/10.21698/rjeec.2022.101

The multicriteria analysis, a useful tool in classifying counties from a socio-economic point of view and investments made in water management: a case study on Romania

ANA MARIA MIHAELA IORDACHE[1], GABRIELA GEANINA VASILE[2*],
LUOANA FLORENTINA PASCU[2]

[1]*Romanian - American University, No.1B Expozitiei Bld., district 1, 012101, Bucharest, Romania*
[2]*National Research and Development Institute for Industrial Ecology ECOIND, 57-73 Drumul Podu Dambovitei Street, district 6, 060652, Bucharest, Romania*
**Corresponding author: gabriela.vasile@incdecoind.ro*

Received:	Accepted:	Published:
28.03.2022	17.06.2022	30.06.2022

Abstract
The article presents a classification methodology of regional development in Romania (41 counties and Bucharest) using a set of data registered in 2019 by the National Institute of Statistics. Several indicators cover information related to the population (the migration, permanent residence, and education), local investments, private turnover, and investments in water supply and sewerage systems, including waste management. After applying the statistical tests, the separation of the counties was done in three classes, the delimitation indicating a high percentage of 59.5% of regions that have low values of the selected indicators. These regions require major investments that will lead to rising living standards, and the emergence of new jobs, driving down the average age of the population. 21.4% of the regions present an average level of the analyzed indicators while only 19% registered the highest values of the parameters, indicating urban agglomerations and a high level of industrial development.

Keywords: *water, SAS Enterprise Guide, classification, discriminant, analysis*

INTRODUCTION

From ancient times, humankind has been concerned with ensuring the resources necessary for survival. Water, the essential element that sustains life on Earth, has been, is, and will be intensively studied both in terms of quality and quantity too because, as it is well known, no resource is unlimited. Although we are in the 21st century, the drinking water supply of the population continues to be a global problem. The lack of adequate resources for the raw water to be used for potable is either due to pollution with organic compounds [1] or toxic metals [2, 3], either due to insufficient amounts (arid disadvantaged in terms of resources) [4, 5].

Due to the high toxicity and stability in the environment, the presence of persistent organic compounds (POPs) in water resources is a real problem in many areas (Turkey, China, Mexico). Compounds such as α-HCH, β-HCH, γ-HCH, ∑-HCH, Heptachlor, Aldrin, p,p′-DDE, p,p′-DDT, ∑-DDT, and ∑-OCP have been detected in various sources of drinking water [6].

Industrial and agricultural activities are the main sources of contamination of both surface and groundwater. Frequent cases of metal contamination caused by mining activities or improper storage of the resulting waste are frequently reported worldwide [7, 8]. In addition, decentralized agriculture has led in some regions to the contamination of groundwater with different compounds (nitrates, phosphates, etc.), depending on the composition of the used fertilizers and the amount applied to the soil [9-11].

Another source of groundwater contamination is zootechnical activities, in which improperly stored animal manure can lead to contamination of aquifers on the extended area adjacent to the respective farm with toxic chemical compounds, pathogenic microorganisms (such as coliforms (i.e.

Escherichia coli), fecal enterococci), and improper biological indicators (such as a large number of heterotrophic species) [9, 11]. In this context, especially in rural areas, it is necessary to use a source of treated drinking water, which is distributed through a water supply system, to avoid to use for drinking purposes underground sources extracted from private wells. If the laboratory results indicate that the groundwater is unfit for human consumption, the solution is to implement a water supply system through which drinking water is provided that meets the conditions imposed by Water Framework Directive [12]. Rehabilitation of existing pipelines, creation of new distribution networks, treatment plants, storage tanks, water intake boreholes, surveillance systems, and data acquisition to control network processes and data collection are activities that involve major local investments.

To protect the quality of surface water, it is necessary to treat wastewater before discharging it into natural receptors. In this sense, the wastewater collection and treatment system requires the extension of the sewerage networks, the modernization of the existing ones, as well as the construction of wastewater pumping and treatment plants. Domestic and industrial wastewater treatment requires specific wastewater treatment plants provided with different treatment stages (mechanical, chemical, biological). Modern technological solutions have been developed, applicable to small communities as well as to low-income countries that do not have the necessary financial funds for the construction and operation of large treatment plants [13].

In recent years, major investments have been made using locally funded projects and European funds in certain regions of Romania (EU Cohesion Fund, Operational Program "Environment" "Extension and modernization of water and wastewater systems" [14, 15]) regarding the water supply for human consumption, respectively the construction of sewerage systems connected to domestic and industrial wastewater treatment plants. However, there are still disadvantaged areas, especially rural zones, where local investments have been reduced. The gaps that are found between rural and urban areas are due to both different economic development and the aging population in rural areas, young people migrating from the village to the city, or even moving to another country. From the data provided by the National Institute of Statistics of Romania, on January 1, 2005, Romania had a population of 21.38 million inhabitants, while on January 1, 2020, it decreased to 19.32 million inhabitants, which indicates a migration of 2 million inhabitants in other countries within 15 years [16].

In this context, this paper aims to analyze the development of different counties (regions) in Romania, using statistical methods of data processing provided by the National Institute of Statistics of Romania, the result being to highlight ways to improve the quality of life in disadvantaged and less developed areas of the country.

MATERIALS AND METHODS

This section includes the proposed methodology for the studied problem, indicators, and data set selected for the analysis.

The proposed methodology for the studied problem

The study contains a multicriteria analysis applied to a set of statistical data regarding the use of the public water supply system, sewerage systems, as well as urban wastewater treatment plants for activities carried out in Romanian counties. Thus, the study aims to detect gaps in regional development and propose solutions to improve activities aimed at environmental protection. The multicriteria analysis also took into account economic factors, such as the level of development of the private sector and investments in environmental protection, as well as social factors related to education and population migration. The applied methodology consisted in identifying the indicators, ranking the counties by applying the least-squares method, discriminant analysis, transfer functions (scoring), and finally, solutions were proposed so that counties from lower classes can access the higher ones.

Discriminant analysis is mainly used in the field of banking - credit scoring, where depending on the characteristics of the applicant, the loan will be granted or not, concerning the value of a score that allows estimating the risk of default. Another common field of application of discriminant analysis is that of consumer behavior [17], where an individual's behavior can be calculated probabilistically to a particular product or service, depending on the state of the explanatory variables that define a particular attitude. This method can also be used for other economic research such as bankruptcy risk analysis [18], but also in other fields such as agriculture [17], physics [19], engineering [20], medicine [21, 22], biology [22], genetics [22], ecology [17, 23], etc.

Defining indicators and data set

For the statistical analysis, nine indicators were selected, their codifications and complete description being presented in Table 1. The data necessary for the multicriteria analysis were taken from the Tempo database of the National Institute of Statistics of Romania (NIS) for 2019 [24].

Table 1. Encoding and description of the indicators used in the application

Encoding indicator	Description of the indicator
I1	The population connected to the sewerage systems in the county
I2	The population connected to urban wastewater treatment systems in the county
I3	The population connected to the public drinking water supply system in the county
I4	Departures with residence (including international migration) from the county
I5	Settling of residence (including international migration) in the county
I6	The total population in the county framed in a form of education
I7	The total number of private entrepreneurs in the county
I8	The turnover obtained by the local units in the county
I9	The gross investments in drinking water distribution systems, sanitation, waste management, depollution activities including wastewater treatment systems in the county

Indicator I1, representing the population connected in each county to the sewerage systems, quantifies in absolute value the number of households that discharge domestic and wastewater into the urban or communal sewerage system. The local sewerage system can be connected to a wastewater treatment plant or not, and in the last case, the collection of domestic and wastewater is directed, the discharge being made in a natural emissary (inland and coastal surface waters). Thus, the indicator I2 used in the multicriteria analysis represents the number of households in absolute value connected to the urban and wastewater treatment plants.

Indicator I3 represents the number of households in the county that were connected to the public drinking water distribution network, the network operated by units specialized in capturing, treating, and distributing water intended for human consumption.

Indicators I4 and I5 refer to the migration of the population, namely the total number of persons who move their residence from the county, respectively the total number of persons who establish their residence in the respective county.

Indicator I6 quantifies the total number of people enrolled in a form of education such as nurseries, pre-university, university, or post-university education system within the county. The introduction of this indicator in the multicriteria analysis aims to identify the need to introduce educational measures, and disciplines related to environmental protection, which can lead to long-term pollution reduction.

Indicators I7, I8, and I9 are economic indicators that target the total number of private economic agents that carry out activities in the county (I7), the turnover of the active local units in the county representing the revenues from sales of goods and services (I8). The last economic indicator is the component of investments in all fixed assets that make up the drinking water distribution and wastewater treatment systems, sanitation, and waste management activities (I9). This last component includes new and existing tangible fixed assets, regardless of whether they are purchased from third parties, acquired under a financial leasing contract, or are produced for their use (including the capitalized production of tangible fixed assets, respectively), intended to be used for more than one year, including non-produced tangible goods, such as land.

RESULTS AND DISCUSSION

The proposed methodology involves defining the analysis indicators, recording their values, and classifying the counties using SAS Enterprise Guide software. Subsequently, after applying the cluster analysis and the discriminant analysis for each class [25], the scoring functions are identified and the initial classification is improved, establishing the final ranking of the counties by classes.

Classification of counties using the least-squares method

The database for which the multicriteria analysis was performed, data set with the nominal values of the nine indicators taken from the Tempo database of the National Institute of Statistics in Romania for 2019 [24], is presented in Table 2. The capital Bucharest is included in the study, having a large population of over 2 million inhabitants.

Table 2. The initial data of the indicators registered for each county in Romania

County	ID	I1	I2	I3	I4	I5	I6	I7	I8	I9
Alba	OB1	196390	184962	264093	6263	5358	55046	6580	29819	39
Arad	OB2	162354	161512	308942	8101	8503	69664	8618	40052	83
Arges	OB3	296936	295843	419914	12417	10503	100484	7640	66442	61
Bacau	OB4	268982	256348	343055	14245	12562	100102	7715	23683	33
Bihor	OB5	268945	244091	368585	10355	10768	107613	11591	43673	56
Bistrita-Nasaud	OB6	129120	126410	188630	4964	4288	48455	3672	14287	146
Botosani	OB7	117803	110640	144175	8495	11317	64370	5040	8575	134
Braila	OB8	210552	210552	300927	5789	4058	43629	4332	11569	21
Brasov	OB9	468334	463450	545813	10875	13003	109637	8593	61722	112
Buzau	OB10	196682	196682	306048	9173	7118	63904	4894	20770	49
Calarasi	OB11	80100	79590	149896	6310	4742	40490	3554	10057	41
Caras-Severin	OB12	147301	119864	174312	6294	4742	39072	3052	7778	42
Cluj	OB13	515889	515524	658399	13539	18368	171640	14165	75380	121
Constanta	OB14	459639	459639	580256	14736	13966	132298	8305	64104	153
Covasna	OB15	114983	109544	116724	2876	2396	33287	3803	7580	145
Dambovita	OB16	134852	134852	348978	9969	9030	73942	6848	19733	36
Dolj	OB17	294317	294317	325468	11887	11073	116174	7021	44017	42
Galati	OB18	330088	324996	408323	11046	14074	92665	5139	27284	119
Giurgiu	OB19	71427	71427	110863	5053	4380	33786	1724	8772	16
Gorj	OB20	102896	89892	200682	8107	6159	53585	3727	10973	9
Harghita	OB21	162322	161942	200648	3863	3465	51562	6886	13141	13
Hunedoara	OB22	312061	310028	331787	8592	6630	58993	5559	16674	43
Ialomita	OB23	90801	90801	161215	5649	4305	37063	2641	9315	28
Iasi	OB24	330773	330773	472706	21761	33316	186319	11665	39550	62
Ilfov	OB25	237813	199033	251751	10107	23856	53106	5169	102987	106
Maramures	OB26	189384	188327	266641	7121	6446	76704	8507	22880	45
Mehedinti	OB27	126631	121802	190646	6265	4849	38306	3330	5709	19
Bucharest	OB28	1781008	1781008	1780585	53221	73165	463700	25599	485022	441
Mures	OB29	283918	277658	341901	9866	9158	97014	9797	40449	40
Neamt	OB30	161675	161675	252669	10292	9407	70290	6861	15737	9
Olt	OB31	119508	119468	179326	8838	6773	59597	3492	19170	103
Prahova	OB32	334015	321682	600234	13784	11866	111343	10324	69127	70
Salaj	OB33	99474	99474	146211	4178	3455	36397	4724	11114	4
Satu Mare	OB34	162773	162773	227468	6118	5367	53641	4839	18208	17
Sibiu	OB35	322004	316137	329623	8301	9896	80674	7576	44311	52
Suceava	OB36	201869	195311	236412	12273	14620	123448	8013	26263	29
Teleorman	OB37	87746	87746	117328	8160	5484	42915	3363	9032	6
Timis	OB38	457552	443998	618955	18277	22281	140748	10440	80008	117
Tulcea	OB39	87666	53914	150788	5069	3651	29722	1861	9213	9
Valcea	OB40	150078	142836	241416	8250	6630	51915	4877	15617	145
Vaslui	OB41	122847	122706	148829	11821	20173	63154	4162	7757	53
Vrancea	OB42	125416	125075	216922	7696	6499	49735	3780	9685	154

For the first classification of the counties, SAS Enterprise Guide software was used and the descriptive indicators mean, standard deviation, skewness, kurtosis, and bimodality were calculated (Fig. 1). The value of the Skewness indicator provides information about the level of asymmetry of the data series and their proximity to the normal distribution. Statistical analysis indicated that all data series are close to a normal distribution, so no correction is required on the initial data [26].
The Kurtosis parameter provides information about the vaulting data series. If the Kurtosis parameter contains zero values in excess, then the distribution of the studied series is of mesokurtic type, and if its values are negative in excess, the distribution becomes platykurtic.
In the studied case, the values of this indicator were high, ranging between 9.7 and 26.6, which indicates that the data series have a leptokurtic distribution, therefore there are either larger queues or there will be outliers. The first conclusion to be taken is that the data series under study have a super-Gaussian distribution [27].
From the covariance matrix point of view (Fig. 1), it is observed that indicators I3, I6, and I7 do not bring additional information, with maximum information being extracted from six indicators, as seen in the Cumulative column. However, these data will be kept in the analysis to obtain an overview of the socio-economic situation in each county.

Variable	Mean	Std Dev	Skewness	Kurtosis	Bimodality
I1	250355	268680	4.7321	26.6105	0.7839
I2	244388	269608	4.7403	26.6651	0.7850
I4	10238.0	7808.1	4.2939	22.9505	0.7424
I5	11135.7	11706.0	3.9739	19.5246	0.7379
I9	71.9762	75.0045	3.0668	13.6344	0.6169
I8	39696.2	74365.4	5.4938	33.1127	0.8579
I7	6654.2	4150.3	2.5066	9.7252	0.5621
I6	83956.9	70916.0	3.9932	20.2610	0.7213
I3	326861	271808	3.9568	20.1122	0.7135

Eigenvalues of the Covariance Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	2.26584E11	2.24486E11	0.9872	0.9872
2	2098271248	1697339622	0.0091	0.9963
3	400931627	6806552.99	0.0017	0.9980
4	394125074	356005106	0.0017	0.9998
5	38119968.2	25412307.5	0.0002	0.9999
6	12707660.7	10460880.4	0.0001	1.0000
7	2246780.3	1459175.57	0.0000	1.0000
8	787604.728	785922.588	0.0000	1.0000
9	1682.13958		0.0000	1.0000

Fig. 1. Indicators calculated for the initial description of data sets

The dendrogram from Fig. 2 results from the computing of the least-squares method on data sets and it is noticed that OB28, namely the city of Bucharest, will be removed from the analysis, being an outlier. For the application of the discriminant analysis, the city of Bucharest will be added manually to the group with counties whose indicators values are as similar as possible to its.
If the graph was cut closer to the zero value of the average distance between clusters parameter, the counties will be divided into five classes. In this case, even if this division reflects better the economic reality, from a statistical point of view the distance between the classes is quite small, these being too close to each other. If the graph is cut closer to the value 1 of the parameter's average distance between clusters, there will be two classes of the division of the counties and even if a big difference is highlighted between them, this fact does not reflect in a satisfactory level the real economic state.
Therefore, the graph was cut closer to the value of 0.5, resulting three distinct classes for counties dividing:

- class 1: Alba, Arad, Bistrita-Nasaud, Botosani, Braila, Buzau, Calarasi, Caras-Severin, Covasna, Dambovita, Giurgiu, Gorj, Harghita, Ialomita, Ilfov, Maramures, Mehedinti, Neamt, Olt, Salaj, Satu Mare, Suceava, Teleorman, Tulcea, Valcea, Vaslui, Vrancea;
- class 2: Arges, Bacau, Bihor, Dolj, Galati, Hunedoara, Iasi, Mures, Prahova, Sibiu;
- class 3: Brasov, Cluj, Constanta, Timis.

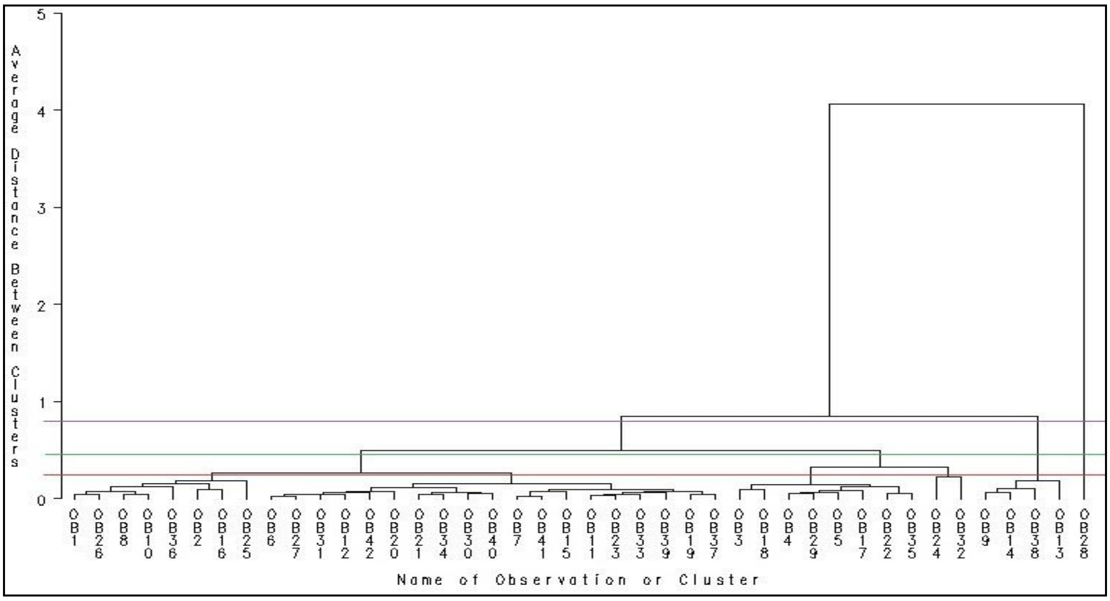

Fig. 2. Dendrogram resulting from the use of SAS Enterprise Guide related to the counties from Romania

Classification of counties using discriminant analysis

Discriminant analysis is a combination of Multivariate Analysis of Variance (MANOVA), multiple regression, and factor analysis, aimed at the main objective of identifying the functions that allow determining the affiliation of a county to one class or another, thus minimizing the classification error.

The main stages of the discriminant analysis algorithm consist of the following steps:

- calculation of the descriptive indicators specific to the MANOVA analysis of the data series;
- identifying coefficients for transfer functions (discrimination);
- calculating a score of the maximum probability regarding the affiliation in the respective class for each object from the data sample;
- comparison between the initially made classification, by applying the least-squares method, and that resulting from the application of the discriminant analysis algorithm.

The statistical tests specific to the MANOVA analysis that were applied to the data series are Wilk`s Lambda, Pillai`s Trace, Hotelling - Lawley Trace, and Roy`s Greatest Root (Table 3). These tests provide information on the possibility of accepting or rejecting the null hypothesis, i.e. the objects are not in the class for which the indicator was calculated. For each class determined in the previous analysis, the discriminant analysis was performed, and a first step in the procedure, implemented in the SAS Enterprise Guide, was the calculation of the MANOVA tests.

In practice, Wilk`s test is mainly used, thus, if its values are close to zero, then the null hypothesis is rejected. It is important to stress that the null hypothesis assumes that the independent variables (indicators used in the application) do not affect the dependent variable (probability of belonging to a certain class). The non-null hypothesis contains the opposite of the null hypothesis, i.e. the results of the discriminant analysis achieve their purpose so that the indicators (independent variables) influence the belonging to the respective class. The other calculated tests (Pillai`s Trace, Hotelling - Lawley Trace, and Roy`s Greatest Root) are used additionally in the case the Wilk`s Lambda test does not provide conclusive information, but in the end, all of them lead to the same conclusions. For example, Roy`s test is used only if the variables are affected by collinearity [28].

Table 3. The MANOVA statistical test values for each class

Test name	Value for Class 1	Value for Class 2	Value for Class 3
Wilk`s Lambda	0.3202	0.6754	0.5115
Pillai`s Trace	0.6798	0.3246	0.4885
Hotelling - Lawley Trace	2.1234	0.4807	0.9549
Roy`s Greatest Root	2.1234	0.4807	0.9549

Table 3 presents the values of the first MANOVA statistical test for each class, indicating that the hypothesis that the objects are not in the class in which the discriminant analysis is performed is ignored.

Among the purposes of the discriminant analysis are the construction of a discriminant space and a rule for the distribution of individuals, a rule that can be used in the future. The discriminant space is the result of the graphical representation of the discrimination functions and a discrimination function is a linear combination of explanatory variables (I1-I9). To determine the discrimination functions for each class were mainly used equations (1) and (2).

$$Sk1_i = \sum_{j=1}^{9} Ij_i * Ck1_j + ak1, \forall i = \overline{1,42}, k = 1,2,3 \quad (1)$$

$$Sk0_i = \sum_{j=1}^{9} Ij_i * Ck0_j + ak0, \forall i = \overline{1,42}, k = 1,2,3 \quad (2)$$

where:

$Sk1_i$ represents the probability (score) of county i to be in class k;

$Sk0_i$ represents the probability (score) of county i not to be in class k;

Ij_i represents the value of the indicator Ij for county i, j = 1 ... 9;

$Ck1_j$ represents the coefficients of the function which calculates the probability of belonging to a county in class k (k = 1, 2, or 3) results from the SAS Enterprise Guide software;

$Ck0_j$ represents the function coefficients that calculate the probability that a county does not belong to class k (k = 1, 2, or 3) results from SAS Enterprise Guide software;

ak1 and ak0 are constants related to linear functions that calculate probabilities;

i represents the number of the county.

Based on the application of general equations (1) and (2) of the scoring model for class 1, the relations (3) and (4) resulted. These relationships helped calculated the belonging probabilities of a county in class 1.

$$S10_i = 0.46 * 10^{-3} * I1 - 0.38 * 10^{-3} * I2 - 2{,}62 * e^{-6} * I3 + 2.79 * 10^{-3} * I4 - 0.67 * 10^{-3} * I5 - 0.21 * 10^{-3} * I6 + 2.93 * 10^{-3} * I7 - 0.4 * 10^{-3} * I8 + 0.025 * I9 - 22.03, \forall i = \overline{1,42} \quad (3)$$

$$S11_i = 0.28 * 10^{-3} * I1 - 0.23 * 10^{-3} * I2 - 0.1 * 10^{-3} * I3 + 2.47 * 10^{-3} * I4 - 0.57 * 10^{-3} * I5 - 0.2 * 10^{-3} * I6 + 2.35 * 10^{-3} * I7 - 0.22 * 10^{-3} * I8 + 0.034 * I9 - 8.87, \forall i = \overline{1,42} \quad (4)$$

Based on the discrimination functions, the probabilities of belonging of each observation (county) to class 1 (Fig. 3a) were calculated. Thus, in class 1 were added two counties: Arges, with a probability of 99.0%, and Dolj, with a probability of 92.4%.

The application of the discriminant analysis for the affiliation of each county to class 2 was made with equations (5) and (6).

$$S20_i = 0.21 * 10^{-3} * I1 - 0.17 * 10^{-3} * I2 - 0.02 * 10^{-3} * I3 + 2.55 * 10^{-3} * I4 - 0.6 * 10^{-3} * I5 - 0.21 * 10^{-3} * I6 + 2.28 * 10^{-3} * I7 - 0.15 * 10^{-3} * I8 + 0.034 * I9 - 7.59, \forall i = \overline{1,42} \quad (5)$$

$$S21_i = 0.3 * 10^{-3} * I1 - 0.25 * 10^{-3} * I2 - 0.03 * 10^{-3} * I3 + 3.2 * 10^{-3} * I4 - 0.78 * 10^{-3} * I5 - 0.24 * 10^{-3} * I6 + 2.89 * 10^{-3} * I7 - 0.22 * 10^{-3} * I8 + 0.023 * I9 - 13.25, \forall i = \overline{1,42} \quad (6)$$

The verification of the affiliation hypothesis from the application of the discriminant analysis for class 2 led to obtaining results that indicate nine counties with a probability of belonging of over 80% (Bacau, Bihor, Caras-Severin, Cluj, Bucharest, Mures, Neamt, Suceava, Timis) (Fig. 3b). The final decision to separate the counties into classes will be taken after applying the discriminant analysis for class 3, choosing, by comparison, the class with the highest probability of belonging, corroborated with the probability of validating the results if there is a situation in which a county has the same probability, framing in two classes simultaneously.

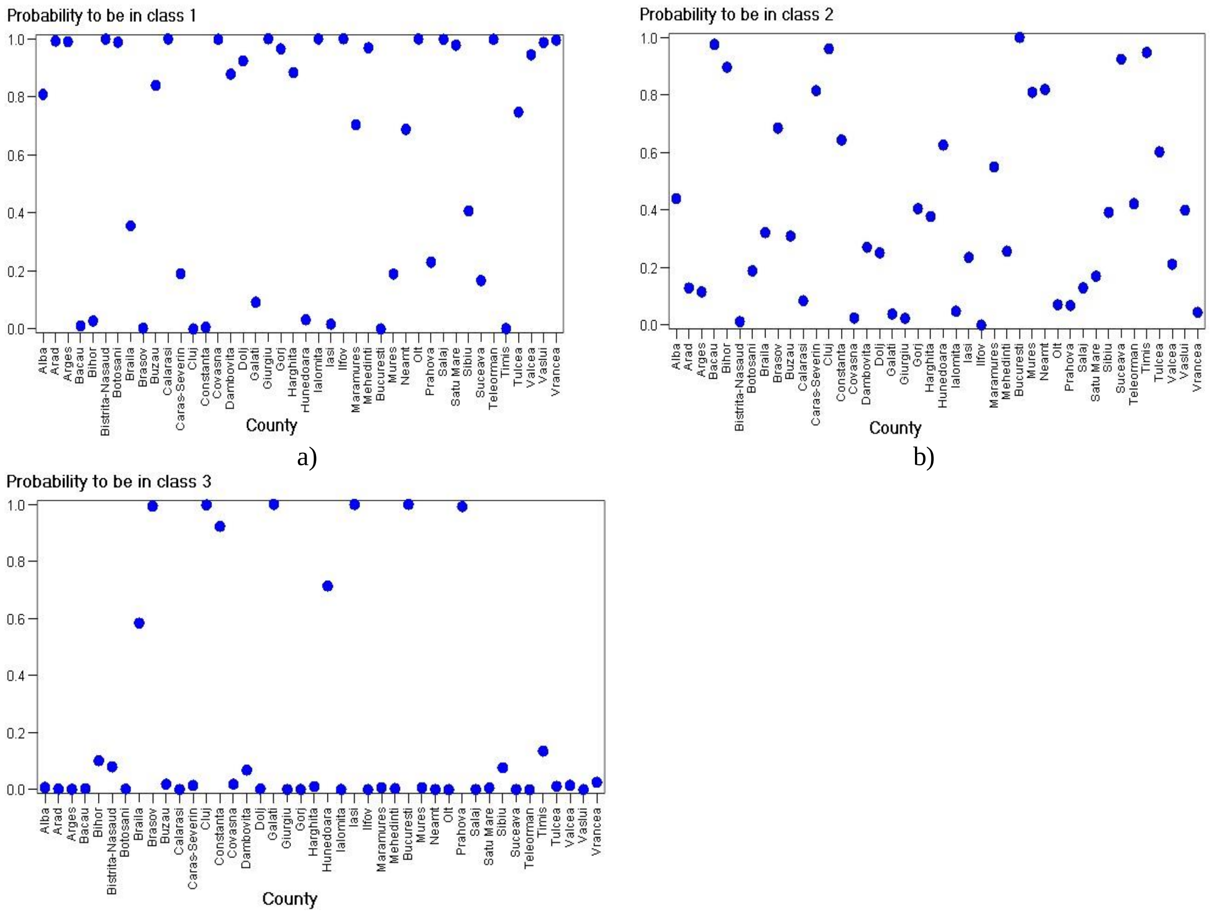

Fig. 3. The probability of the counties belonging to class 1 (a), 2 (b), or 3 (c)

The same algorithm was applied for the last class, the functions according to which the score of belonging to a county in class 3 was calculated are represented by equations (7) and (8).

$$S30_i = 0.18 * 10^{-3} * I1 - 0.16 * 10^{-3} * I2 - 8.33 * e^{-6} * I3 + 2.0 * 10^{-3} * I4 - 0.45 * 10^{-3} * I5 - 0.19 * 10^{-3} * I6 + 1.94 * 10^{-3} * I7 - 0.14 * 10^{-3} * I8 + 0.041 * I9 - 7.26, \forall i = \overline{1,42} \quad (7)$$

$$S31_i = 0.26 * 10^{-3} * I1 - 0.22 * 10^{-3} * I2 + 0.04 * 10^{-3} * I3 + 0.54 * 10^{-3} * I4 + 0.03 * 10^{-3} * I5 - 0.12 * 10^{-3} * I6 + 1.31 * 10^{-3} * I7 - 0.29 * 10^{-3} * I8 + 0.056 * I9 - 16.63, \forall i = \overline{1,42} \quad (8)$$

For belonging to class 3, the counties are very clearly delimited (Fig. 3c), seven of them having a probability higher than 80% (Brasov, Cluj, Constanta, Galati, Iasi, Prahova, and Bucharest). The counties of Braila and Hunedoara have a probability of being in this class higher than 50%, but the decision in their case will be taken after checking the cross-validation results (Table 4).

Therefore, the third class will contain the following counties: Braila (58.4%), Brasov (99.444%), Cluj (99.824%), Constanta (92.304%), Galati (100%), Hunedoara (71.4%), Iasi (100%), Prahova (99.2%), and Bucharest (100%).

Bucharest has the maximum probability of belonging to two classes (Table 4): class 2 and class 3. The decision regarding its classification in one of the two groups is taken based on cross-validation results, thus, for class 2, the probability is 21.8%, which is much higher than the 5.22%, probability for class 3. In conclusion, Bucharest will be framed in class 2.

After an overall analysis of the obtained results and taking into account both the maximum probability of belonging of a county to a class and the probability of validation of the results (Table 4), the final classification of the counties is the following:

- class 1: Alba, Arad, Arges, Bistrita-Nasaud, Botosani, Buzau, Calarasi, Covasna, Dambovita, Dolj, Giurgiu, Gorj, Harghita, Ialomita, Ilfov, Mehedinti, Olt, Salaj, Satu Mare, Sibiu, Teleorman, Tulcea, Valcea, Vaslui, Vrancea;
-class 2: Bacau, Bihor, Caras-Severin, Mures, Neamt, Suceava, Timis, Maramures, Bucuresti;
-class 3: Braila, Brasov, Cluj, Constanta, Galati, Hunedoara, Iasi, Prahova.

Table 4. Resubstitution and cross-validation results for each county, %

County	Resubstitution Results			Cross-Validation Results		
	Class 1	**Class 2**	**Class 3**	**Class 1**	**Class 2**	**Class 3**
Alba	80.9	44.0	0.76	84.8	40.7	0.62
Arad	99.3	12.9	0.22	99.4	11.2	0.18
Arges	99.0	11.5	0.11	92.6	27.0	0.10
Bacau	1.04	97.6	0.27	0.52	98.3	0.22
Bihor	2.74	89.6	10.1	0.87	95.7	2.73
Bistrita-Nasaud	99.9	1.18	7.95	99.9	1.68	3.97
Botosani	99.8	18.9	0.18	99.2	15.8	0.16
Braila	35.5	32.2	58.4	60.5	26.6	36.1
Brasov	0.21	68.5	99.4	0.16	52.6	99.8
Buzau	84.0	31.0	1.91	87.1	28.7	1.56
Calarasi	99.9	8.42	0.01	99.9	8.07	0.01
Caras-Severin	19.1	81.5	1.50	56.6	64.4	0.87
Cluj	0.00	96.1	99.8	0.01	70.6	100
Constanta	0.56	64.4	92.3	0.39	51.7	98.3
Covasna	99.9	2.42	1.91	99.9	3.07	1.04
Dambovita	87.8	27.1	6.83	94.4	21.0	3.19
Dolj	92.4	25.1	0.22	59.9	54.4	0.19
Galati	9.24	3.82	100	3.37	30.1	99.2
Giurgiu	100	2.33	0.01	100	2.75	0.02
Gorj	96.5	40.5	0.07	97.5	35.1	0.07
Harghita	88.4	37.8	1.07	93.3	30.9	0.75
Hunedoara	3.15	62.5	71.4	1.03	87.5	21.5
Ialomita	99.9	4.84	0.07	99.9	4.86	0.07
Iasi	1.66	23.5	100	0.3	92.5	42.2
Ilfov	100	0.00	0.00	100	0.94	0.01
Maramures	0.43	55.0	0.69	80.0	48.3	0.54
Mehedinti	96.9	25.6	0.37	97.4	24.0	0.31
Bucharest	0.00	100	100	0.00	21.8	5.22
Mures	18.9	80.9	0.60	9.90	86.5	0.46
Neamt	68.8	81.9	0.02	80.8	73.0	0.02
Olt	99.9	7.06	0.01	99.9	6.87	0.01
Prahova	23.0	6.78	99.2	3.30	62.3	66.1
Salaj	99.8	12.9	0.03	99.8	11.9	0.03
Satu Mare	97.8	17	0.58	98.2	15.7	0.48
Sibiu	40.7	39.2	7.68	22.1	57.9	4.85
Suceava	16.7	92.4	0.03	67.4	72.2	0.04
Teleorman	99.9	42.2	0.00	99.9	34.1	0.00
Timis	0.12	94.79	13.5	0.10	86.2	85.5
Tulcea	74.8	60.2	1.15	94.7	33.8	0.54
Valcea	94.4	21.1	1.51	96.9	17.3	0.94
Vaslui	98.1	40.0	0.01	99.3	23.5	0.03
Vrancea	99.5	4.43	2.58	99.6	4.74	1.47

The classification of the counties into groups was performed according to the indicators under study. Each class was characterized, by comparison with the others, according to the aggregate indicators of sum and arithmetic mean (Table 5), and scenarios for the transition of a county from one class to another were elaborated.
Table 5 presents the aggregate indicators of the sum and arithmetic mean at the level of each class, taking into account the division made after applying the discriminant analysis. For the indicators I1-I6, the expression was achieved as a percentage reporting the data in absolute value to the total population of all counties from a group, in 2019, while indicators I7, I8, and I9, being predominantly economic indicators, the data were expressed in absolute value.

Table 5. The aggregate indicators of the sum and arithmetic mean of each class

	I1	I2	I3	I4	I5	I6	I7	I8	I9
Total class 1	43.2%	41.7%	62.0%	2.14%	2.16%	16.0%	119821	556336	1510
Average class 1	1.73%	1.67%	2.48%	0.08%	0.08%	0.64%	4793	22253	60.4
Total class 2	62.5%	61.0%	72.9%	2.36%	2.71%	20.3%	91575	745493	812
Average class 2	6.95%	6.78%	8.09%	0.26%	0.31%	2.25%	10175	82833	90.2
Total class 3	75.9%	75.2%	99.9%	2.56%	2.95%	23.2%	68082	365410	701
Average class 3	9.48%	9.40%	12.5%	0.32%	0.36%	2.90%	8510	45676	87.6

Class 1 has 25 counties and is characterized by low values of indicators compared to other classes, as follows: the number of people connected to sewerage systems, water treatment systems, and public water supply system is small; the ratio between the number of people who left and the number of people who have been established is unitary (the real movement of the population tends to zero). An important aspect is the indicator that quantifies the school-age population which has low values, that indicates a fairly low birth rate. Not being a local workforce, the number of private economic agents is reduced both from this point of view, as well as due to the limited economic development perspectives, and the turnover is diminished. At the level of this class, the gross investments in depollution, waste management, sewerage, and drinking water activities (indicator I9) are high in absolute value due to a large number of counties in this class, but on average, they are the lowest compared with classes 2 and 3. To improve the indicators, it is recommended that these counties become attractive from a business environment point of view, respecting the norms of the European Union regarding greening and the increase of the living standard.
Class 2 has nine counties and is characterized by average values of indicators that quantify the connection of the population to sewerage systems, sewage treatment systems, and the public water supply system. The departure-establishment ratio of the population from the counties is slightly super unitary, which means that the number of persons who have established their residence in these counties, on average, is slightly higher than those who have left. Compared to the first class, the indicators I6 (total school population) and I7 (number of private entrepreneurs) have higher values than in the first group, and the immediate consequence is that the turnover tends to have high values, so this class is preferred by the business sector. Going by the cascade, if the number of companies increases, automatically and the turnover per county will increase too. Investments in actions related to ecology have an average level, one reason would be the fact that from an industrial point of view the pollution level was kept under control either by previous investments in this sector or by the small number of economic agents that have no polluting activities.
Class 3 consists of eight counties, characterized in turn by urban agglomerations and industrialization. In addition, there are districts with touristic sights, such as Black Sea resorts (Constanta County) and Prahova Valley resorts (Prahova and Brasov Counties). Due to the increased population, on average, the values of indicators I2 and I3 are high. From the population movement point of view, the tendency is to have more settlements with residences in these counties than departures (I5> I4). A consequence of the increased population is that the school population has high values. In addition, entrepreneurs will prefer this class to expand their business due to the offered facilities, and especially due to the possibility of hiring locals, so they will be able to

minimize their staff costs. From investments in ecological activities point of view, these have average values.
In the following, it will be analyzed depending on indicators values how a county can slide from one class to another. Three scenarios were proposed (Table 6) for passing, for example, Arges County from class 1 (where it was initially classified) to class 2, the approach being made in three directions: social, economic, and ecological.

Table 6. Scenarios for passing Arges from class 1 to class 2

	The initial values of the indicators	The new values of the indicators
Scenario 1	I4=2.14%, I5=1.81%, I6=17.33%	I4=0.1%, I5>10%, I6>35%
Scenario 2	I7=7640, I8=66442	I7>20000, I8>100000
Scenario 3	I1=51.2%, I2=51.02%, I3=72.42%, I9=61	I1>75%, I2>74%, I3>90%, I9>900

Scenario 1 brings to the fore the change of indicators I4, I5, and I6, approaching the problem from a social point of view. Thus, to achieve the transition from class 2 to class 3, it was proposed to modify the indicators to reduce departures of the population from the county or departures to be significantly small compared to the total population (I4 = 0.1%). Therefore, it is important to adopt policies at the county level to encourage the settlements, and to increase the number correlated with population residence (I5 higher than 10%). The percentage of newly established people in Arges should be as high as possible to the existing population, and the indicator that quantifies the inclusion of the school population in a form of education should increase at least to 35%.
Scenario 2 is considering the change of economic indicators, I7 and I8, massive investments being able to be made in the economy and, implicitly, increase substantially the turnover.
Scenario 3 refers to the ecological aspect, thus it was proposed to increase the percentages of the population connected to the sewerage, water purification, and public water supply systems. Currently, the increase in these percentages cannot take place due to the low-level investments occurring in activities specific to indicator I9.
The three proposed scenarios do not exclude each other and, somewhat, derive from each other. Thus, attracting new inhabitants to the county also entails economic development by increasing the number of economic agents who will bring their businesses closer to the local labor force and other resources. Since the number of economic agents will increase, the need to make investments in activities related to drinking water supply, sewerage systems, and waste management activities as well as in wastewater treatment plants would also increase, thus modifying the specific indicators.

CONCLUSIONS

A multidisciplinary overall analysis was performed on the counties from Romania, choosing indicators regarding the population's drinking water supply, access to sewerage and wastewater treatment plants, waste collection and investments in the ecological field, population fluctuations regarding establishing the residence in each county, and the level of economic activity quantified by the turnover, as well as the total number of entrepreneurs in the county. After applying the discriminant analysis, the counties were divided into three distinct classes, relatively small differences being between counties within a single class. For the transition from one class to another, several scenarios were proposed, by choosing some thresholds, depending on the policy adopted by the decision-making forums at the county level. Thus, policies can be adopted either in a single plan (social, economic, or environmental), but most of the time, involvement of all three directions is needed.
The model proposed in the article has as a first novelty the choice of indicators, therefore there are indicators related to population migration, economic indicators, and indicators on pollution and its control (number of water treatment plants, sewerage systems, etc.). Another novelty is the performance of an overall multidisciplinary analysis at the level of each county in Romania, combining three different directions: social, economic, and environment. The combination of supervised recognition techniques of forms such as cluster analysis and discriminant analysis offers

both a robust vision of the existing reality and directions for its improvement. Thus, by proposing scenarios, strategies can be developed in one direction or another, depending on socio-economic or environmental needs at a given time.

REFERENCES

[1] NICULAE, A., VASILE, G.G., CRUCERU, L., ENE, C., Rev. Chimie, **69**, no. 1, 2019, p. 6, https://doi.org/10.37358/RC.18.1.6034.
[2] SOMA, G., ABHAY. K.S., Environ. Earth Sci., **72**, 2019, p. 1, https://doi.org/10.1007/s12665-019-8200-9.
[3] GHEORGHE, S., STOICA, C., VASILE, G.G., NITA-LAZAR, M., STANESCU, E., LUCACIU, I.E., ″Metals toxic effect in aquatic ecosystems: modulators of water quality″. In *Water Quality*, edited by H. Tutu, 416 p. London, United Kingdom: InTechOpen, 2017, https://doi.org/10.5772/62562.
[4] YU, Y., PI, Y., YU, X., TA, Z., SUN, L., DISSE, M., ZENG, F., CHEN, X., YU, R., J. Arid Land, **11**, no. 1, 2019, p. 1, https://doi.org/10.1007/s40333-018-0073-3.
[5] SHARAFATMANDRAD, M., MASHIZI, A.K., Water Resour. Manag., **35**, 2021, p. 63, https://doi.org/10.1007/s11269-020-02706-1.
[6] VASSEGHIAN, Y., HOSSEINZADEH, S., KHATAEE, A., DRAGOI, E.N., Sci. Total Environ., **796**, 2021, p.1, https://doi.org/10.1016/j.scitotenv.2021.149000.
[7] VERMA, C., HUSSAIN, A., MADAN, S., KUMAR, V., Appl. Water Sci., **11**, no. 58, 2021, p. 1, https://doi.org/10.1007/s13201-021-01390-9.
[8] CUCIUREANU, C., KIM, L., LEHR, C.B., ENE, C., Rev. Chimie, **68**, no. 8, 2017, p. 1695, https://doi.org/10.37358/RC.17.8.5746.
[9] CELLONE, F., CAROL, E., PUGLIESE, I., CORDOBA, J., BUTLER, L., LAMARCHE, L., Environ. Earth Sci., **79**, 2020, p.1, https://doi.org/10.1007/s12665-020-09005-3.
[10] STANESCU, B., KIM, L., BATRINESCU, G., CUCIUREANU, A., Proceedings of the International Multidisciplinary Scientific GeoConference-SGEM, Geoconference on Science and Technologies in Geology, Exploration and Mining, Albena, Bulgaria, July 2014, vol. 2, p. 1051, https://doi.org/10.5593/sgem2014/b12/s2.134.
[11] SASAKOVA, N., GREGOVA, G., TAKACOVA, D., MOJZISOVA, J., PAPAJOVA, I., VENGLOVSKY, J., SZABOOVA, T., KOVACOVA, S., Front. Sustain. Food Syst., **2**, 2018, p. 1, https://doi.org/10.3389/fsufs.2018.00042.
[12] DIRECTIVE (EU) 2020/2184 of the European Parliament and of the Council of 16 December 2020 on the quality of water intended for human consumption (recast). Accessed April 15, 2021. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32020L2184.
[13] TOMPKINS, D., BUMBAC, C., CLIFFORD, E., DUSSAUSSOIS, J.B., HANNON, L., SALVADO, V., SCHELLENBERG, T., Water, **11**, no. 12, 2019, p. 1, https://doi.org/10.3390/w11122461.
[14] European Commission (2021a). Investment in water and sewerage system to improve public health and environment in western region of Romania. Accessed May 10, 2021. https://ec.europa.eu/regional_policy/ro/projects/romania/investment-in-water-and-sewerage-system-to-improve-public-health-and-environment-in-western-region-of-romania.
[15] European Commission (2021b). Water supply and wastewater treatment systems modernized in north-east Romania. Accessed May 10, 2021. https://ec.europa.eu/regional_policy/ro/projects/romania/water-supply-and-waste-water-treatment-systems-modernised-in-north-east-romania.
[16] NATIONAL INSTITUTE OF STATISTICS (2021a). Accessed February 10, 2021. https://insse.ro/cms/ro.
[17] MAZUREK-KUSIAK, A., SAWICKI, B., KOBYLKA, A., Sustainability, **13**, no. 14, 2021, p. 1, https://doi.org/10.3390/su13148005.
[18] KESKIN, A.I., DINCER, B., DINCER, C., Sustainability, **12**, no. 6, 2020, p. 1, https://doi.org/10.3390/su12062346.

[19] DANIEL, T., CASENAVE, F., AKKARI, N., RYCKELYNCK, D., Math. Comput. Appl., **26**, no. 1, 2021, p.1, https://doi.org/10.3390/mca26010017.
[20] SUBHAN, F., SALEEM, S., BARI, H., KHAN, W.Z., HAKAK, S., AHMAD, S., EL-SHERBEENY, A.M., Sustainability, **12**, no. 24, 2020, p. 1, https://doi.org/10.3390/su122410627.
[21] RIVERA-KEMPIS, C., VALERA, L., SASTRE-CASTILLO, M.A., Sustainability, **13**, no. 15, 2021, p.1, https://doi.org/10.3390/su13158252.
[22] LEE, S.H., KAO, G.D., FEIGENBERG, S.J., DORSEY, J.F., FRICK, M.A., JEAN-BAPTISTE, S., UCHE, C.Z., CENGEL, K.A., LEVIN, W.P., BERMAN, A.T., AGGARWAL, C., FAN, Y., XIAO, Y., Int. J. Radiat. Oncol. Biol. Phys., **110**, no. 5, 2021, p. 1451, https://doi.org/10.1016/j.ijrobp.2021.02.030.
[23] LIU, G., MA, F., LIU, G., ZHAO, H., GUO, J., CAO, J., Sustainability, **11**, no. 12, 2019, p. 1, https://doi.org/10.3390/su11123345.
[24] NATIONAL INSTITUTE OF STATISTICS (2021b). *TEMPO Online Database.* Accessed February 10, 2021. http://statistici.insse.ro:8077/tempo-online/#/pages/tables/insse-table.
[25] SAS Support (2021). *Discriminant Analysis*. Accessed February 10, 2021. https://support.sas.com/rnd/app/stat/procedures/DiscriminantAnalysis.html.
[26] FREEMAN, J.B., DALE, R., Behav. Res. Methods, **45**, 2013, p. 83, https://doi.org/10.3758/s13428-012-0225-x.
[27] BENVENISTE, A., GOURSAT, M., RUGET, G., IEEE Trans. Automat. Contr., **25**, no. 3, 1980, p. 385, https://doi.org/10.1109/TAC.1980.1102343.
[28] RENCHER, A. (2002). *Methods of Multivariate Analysis*, 2nd. ed.: Brigham Young University, John Wiley & Sons Inc.

Citation: Iordache, A.A.M, Vasile, G.G., Pascu, L.F., The multicriteria analysis, a useful tool in classifying counties from a socio-economic point of view and investments made in water management: a case study on Romania, *Rom. J. Ecol. Environ. Chem*., **2022**, 4, no.1, pp. 9-21.