Selection of the Optimal Way of Linear Ordering of Objects: Case of Sustainable Development in EU Countries

Agnieszka Sompolska-Rzechuła[1] | *West Pomeranian University of Technology Szczecin, Szczecin, Poland*

Abstract

The aim of the article was to assess selected methods of linear ordering of objects and choosing the optimal method. The measures based on different properties of the synthetic variable were selected for evaluation. The selection of the optimal linear ordering procedure is the last step in creating a synthetic variable and is often not included in the research. The analysis was based on data from the EUROSTAT database (2017) countries. The level of socio-economic development in the context of sustainable development for 28 European Union was adopted as the ordering countries. The paper proposes a comparison of results in various methods, e.g. due to the way of normalization of diagnostic features or type of methods (based on a pattern object or a non-pattern object). Out of all the selection methods for this study, the TOPSIS methods based on zero unitarization proved to be the optimal.

Keywords	*JEL code*
Linear ordering of objects, selection of method of linear ordering, level of socio-economic development, sustainable development	*C38, Q01, O57*

INTRODUCTION

The methods of ordering objects make possible to determine the order of objects depending on the degree of intensity of specific features. Linear ordering methods are included among the ordering methods. Linear ordering is based on a feature called synthetic, but objects are multidimensional. A synthetic feature aggregates partial information contained in simple features that make up the evaluation criterion (Wysocki, 2010). The first proposal for a synthetic variable was presented by Hellwig (1968).

Among others, Hartigan (1975), Pluta (1977), Hwang and Yoon (1981), Anderson (1984), Seber (1984), Morrison (1990), Grabiński (1992), Chen (2000), Kukuła (2000) had a significant share in the development of these methods.

The idea of linear ordering of multidimensional objects is based on the concept of ordering binary relations (reflexive, non-symmetrical, transitive and coherent). The axioms of this relation show

---

[1] West Pomeranian University of Technology Szczecin, Department of Mathematics Applications in Economy, Janickiego 31, 71-270 Szczecin, Poland. E-mail: asompolska@zut.edu.pl.

that it is possible to state which of any two objects of the set is the first (better) and which is the second (worse), as well as whether they are identical (Bąk, 2015). The subject of linear ordering can be objects such as countries, enterprises, products or people. With many methods of linear ordering of objects available, it is not always clear which procedure to choose. The quality assessment of linear ordering procedures is the last step in creating a synthetic variable. In the literature on the subject, one can find mainly works whose final result is the ranking of objects without assessing the optimality of the results obtained.

The issue of choosing the optimal method of linear object ordering was taken up in the works Bąk (2015, 2018), and Sompolska-Rzechuła (2020). In the latter, an evaluation of selected methods of linear ordering was proposed, adopting various methods of normalizing diagnostic features and one method of aggregating variables. However, this work is a continuation and extension of research on the selection of final results obtained with the methods of linear ordering of objects, because both the methods based on a pattern object and a non-pattern object were used. Moreover, the measures of correctness of the linear ordering procedures from each group of measures presented in Table 1 were used and some modifications to the measures of correctness of the linear ordering procedures have been proposed. Both articles used data from the EUROSTAT database (2017) for 28 European Union countries. The information is pertaining to the level of socio-economic development in the context of sustainable development.

The level of socio-economic development was adopted as the criterion for ordering countries, which was presented with the use of indicators reflecting the concept of sustainable development. Sustainable development is implemented in three dimensions: economic, social and ecological and is based on the pursuit of the best economic result while respecting the natural environment and social development. It is therefore a social and economic development that ensures that the needs of modern society are met without hindering future generations from meeting their needs. The 3 × P abbreviation is often presented, *planet*, *people*, with *profit* at the very end. This order suggests an emphasis put above all on preserving Earth's resources, not threatening the environment, and the profit is seen only at the very end (Latoszek, 2016). Therefore, the essence of the concept of sustainable and permanent development is to meet the needs of present generation without reducing the chances of future generation to meet them. This definition was included in the 1987 report of the World Commission on the Environment and Development entitled: *Our Common Future* (United Nations General Assembly, 1987). Although there are many definitions, the most commonly used definition of sustainable development is the one proposed by the Brundtland Commission: "Sustainable development is development that meets the needs of the present without compromising the ability of future generation to meet their own needs" (Cerin, 2006). The concept of sustainable development was designed in the 1980s and is one of the most important contemporary concepts for economic development.

1 RESEARCH METHODOLOGY

Linear ordering methods are used to evaluate multi-feature objects (e.g. countries) allowing them to be ranked, according to a specific general criterion, from "best" to "worst". This criterion is treated as a property of the examined objects and is a complex phenomenon. Socio-economic research very often examined phenomena that are not directly measured. Sets of diagnostic features are used then, measured on various measuring scales. Linear ordering of objects is obtained on the basis of a feature called aggregate or synthetic, which is created by aggregating the initial features describing the tested objects.

The synthetic feature creation procedure is a multi-step process and includes (Wysocki, 2010; Sompolska-Rzechuła, 2020):

1) gathering preliminary information about potential diagnostic variables;
2) selecting of diagnostic variables;
3) determining the type of variables: stimulant, destimulant or nominant;
4) normalising of diagnostic variables using the selected normalizing method;

5) assigning weights to standardized features;
6) calculating values of the aggregation of features, i.e. creating a synthetic variable;
7) construction of the linear ordering of facilities due to the level of the complex phenomenon in question;
8) quality assessment of rankings using partial quality assessment criteria and aggregate measures were calculated.

The first step in creating a synthetic feature is to establish a set of diagnostic features. There are two approaches to this issue – non-statistical (theoretical) and statistical (Wysocki, 2010). The substantive approach is based on the qualitative assessment of the studied phenomenon, taking into account economic knowledge and theory. The statistical approach is designed to limit the set of diagnostic features and exclude those features that do not fully characterize the examined objects in terms of the adopted criterion. Analysis of variability and correlations between features is often used. Another condition that diagnostic features should meet is the lack of correlation between features. Therefore, from the set of potential features, features strongly correlated with others should be eliminated, because they are a carrier of similar information. In the literature on the subject, you can find many methods used in the selection of features. These include methods based on, e.g., the analysis of correlation coefficients between features. One of these methods is the procedure proposed by Hellwig and known as the parametric method of feature selection. A detailed description of this method can be found, e.g. in the works by (Wysocki, 2010, pp. 146–147), and (Sompolska-Rzechuła, 2018, pp. 74–76). The algorithm of the parametric Hellwig method is as follows:

1. Calculating the correlation matrix **R** of *k* variables.
2. Determining the threshold value of the correlation coefficient (r^*) e.g. based on the formula:

$$r^* = \min_i \max_j |r_{ij}| \quad i, j = 1, \ldots, k, \tag{1}$$

where: r_{ij} – Pearson linear correlation coefficients between features, k – number of features. The threshold value of the coefficient can also be taken arbitrarily, often as $r^* = 0.5$.

3. Calculating the sum of the absolute values of the correlation coefficients for each column of the matrix **R**:

$$R_j = \sum_{i=1}^{k} |r_{ij}|. \tag{2}$$

4. Determining the column number (m), for which the sum R_j is the largest:

$$R_m = \max_j \{R_j\}. \tag{3}$$

5. Classification of variables: the variable with the number (k) is the central variable, and the variables for which $|r_{ik}| < r^*$ are satellite variables (they form a cluster of highly correlated variables, thus causing information redundancy).
6. Removing rows and columns from the matrix **R** corresponding to satellite variables and the column corresponding to the central variable.
7. The procedure is repeated until the set of variables is exhausted.
8. Variables that are not in any cluster are isolated variables (they form one-element clusters).
9. Central and isolated variables are included in the analysis (satellite variables are discarded).

The set of diagnostic features is the basis for further analysis, in which the nature of the features should be determined, i.e. stimulants, destimulants and nominants should be distinguished.

After recognizing the nature of the features, they must be transformed; most often destimulants are converted into stimulants by means of difference or quotient transformations.

The concept of stimulants and destimulants was introduced by Hellwig (1968).

The stimulant means a feature which higher value indicates a better condition of the object in a given context. Thus, the maximum value of the stimulant is considered the most favourable, and the minimum – the least favourable for the examined objects. While the destimulant is a feature which lower values mean a better situation of the object in a given respect. Therefore, the maximum value of destimulants is considered the least favourable, and the lowest – the most favourable for the examined objects. While the neutral variable is characterized by the existence of an optimal value (for this reason sometimes called a nominal value), below which such a variable has the character of a stimulant (and therefore larger values are more favourable), and above the destimulant (which means that after exceeding the optimal value, a further increase in the value of the feature becomes unfavourable) or the other way round. Neutral variables are often overlooked in empirical studies due to difficulties in establishing nominal values. If it is difficult to determine the nature of the characteristics, specific substantive criteria or correlation analysis should be used. It is also possible to evaluate the nature of the features after determining the value of the synthetic feature, then stimulants should be positively correlated with the synthetic feature and destimulants – negatively.

The next stage of building the synthetic feature is the normalization of features. It leads to deprivation of physical units of measurement results and unification of orders of magnitude. The literature contains many proposals for these methods and discussions on the criteria for their selection. The rest of the work will present those normalizing formulas that relate to stimulus traits.

The following standardizing formulas have been used in this work (Kukuła, 2000):

- zero unitarization:

$$z_{ij} = \frac{x_{ij} - \min_l x_{lj}}{\max_l x_{lj} - \min_l x_{lj}} \quad (\max_l x_{lj} \neq \min_l x_{lj}), \tag{4}$$

- quotient transformation:

$$z_{ij} = \frac{x_{ij}}{\overline{x}_j} \quad (\overline{x}_j \neq 0), \tag{5}$$

where: z_{ij} – standardized value of the j-th feature ($j = 1, 2, \ldots, k$) for the i-th object ($i = 1, 2, \ldots, n$), n – number of object.

In the zero unitarization method, a constant pattern point is assumed – the range of the normalized variable. The use of this method makes the range of the normalized feature constant and amounts to one. The normalized feature assumes values in the range [0,1]. Moreover, this method makes possible to normalize the features taking positive, negative and zero values.

In the next step of creating a linear ordering of objects, the values of the synthetic feature are determined. There are many methods of constructing a synthetic development measure that can be divided into non-pattern and pattern (Grabiński, 1992). The main difference between the pattern and non-pattern methods lies in the fact that in pattern methods the basis of analyses takes on the form of a concept of a development pattern, which is understood as a certain artificially constructed object, characterized by some optimal properties expressed in properly defined values of diagnostic features.

A non-pattern object methods rely on the operation of averaging the values of standard features (Wysocki, 2010):

$$\mu_i = \frac{1}{k} \sum_{j=1}^{k} z_{ij}\,, \tag{6}$$

where: μ_i – the value of the synthetic feature for the i-th object.

The idea of pattern methods of aggregation of features is to determine the distance of individual objects from a certain pattern object. Among the pattern methods, the method proposed by Hellwig (1968) deserves attention. It is based on standardized values of diagnostic features $X_1, X_2, \ldots, X_k$ which are treated as equally important. The Euclidean distances of each object are calculated from the pattern according to the formula:

$$d_i = \sqrt{\sum_{i=1}^{k} (z_{ij} - z_{0k})^2} \; (i = 1, 2, \ldots, n), \tag{7}$$

where: $z_{0k} = \max\limits_i \{z_{ik}\}$ – standardized value of the k-th feature for the pattern object, n – number of objects. In this paper zero unitarization method was used to standardize variables.

Based on the d_i value, the relative taxonomic measure of development is constructed, defined as (Nowak, 1990):

$$\mu_i = 1 - \frac{d_i}{d_0} \; (i = 1, 2, \ldots, n), \tag{8}$$

where: $d_0 = \bar{d} + 2 \cdot s_d$,

wherein: $\bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_i$ and $s_d = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (d_i - d)^2}$.

The synthetic Hellwig development measure almost always takes values from [0,l]. The smaller the difference of the μ_i measure from one, the less the level of object development differs from the level of object development recognized as the standard. The synthetic measure of development is a resultant of all the features characterizing the examined objects, it allows to determine the "average" level of the value of the features achieved at some time.

The linear ordering method using the pattern and non-pattern is the TOPSIS method (*Technique for Order Preference by Similarity to an Ideal Solution*; Hwang et Yoon, 1981). It consists in calculating the Euclidean distances of each assessed object from both the pattern and non-pattern of the development, which distinguishes it from the Hellwig method, which only takes into account the distance from the development pattern (Wysocki, 2010). The coordinates of the model units are set – development pattern and non-pattern. The values of the pattern (A^+) and non-pattern of development (A^-) are defined as (Wysocki, 2010):

$$A^+ = \left(\max_i(z_{i1}), \max_i(z_{i2}), \ldots, \max_i(z_{ik})\right) = (z_1^+, z_2^+, \ldots, z_k^+), \tag{9}$$

$$A^- = \left(\min_i(z_{i1}), \min_i(z_{i2}), \ldots, \min_i(z_{ik}),\right) = (z_1^-, z_2^-, \ldots, z_k^-). \tag{10}$$

If zero unitarization is used as the normative formula, it is:

$$z^+ = \underbrace{(1, 1, \ldots, 1)}_{k} \quad z^- = \underbrace{(0, 0, \ldots, 0)}_{k}. \tag{11}$$

Calculating the Euclidean distances of each object from the pattern and non-pattern is made according to the formulas:

$$d_i^+ = \sqrt{\sum_{j=1}^{k} (z_{ij} - z_j^+)^2}, \; d_i^- = \sqrt{\sum_{j=1}^{k} (z_{ij} - z_j^-)^2}, \; i = 1, 2, \ldots, n. \tag{12}$$

While the value of the synthetic feature is determined as follows (Hwang and Yoon, 1981):

$$\mu_i = \frac{d_i^-}{d_i^+ + d_i^-}, \tag{13}$$

wherein: $0 \leq \mu_i \leq 1$, $i = 1, 2, \ldots, n$.

The smaller the distance of a given object from the development pattern, and thus greater than the development non-pattern, the closer the value of the synthetic feature.

The final stage in the construction of the synthetic variable is the assessment of the correctness of the procedure of linear ordering of objects. This stage is often overlooked in the analysis of the linear ordering of objects. Perhaps this is due to the lack of publicly available computer software capable of performing this type of analysis. Assessing the optimality of the procedure of linear ordering of objects is a time consuming and quite complicated issue.

In the final stage, measures are used to characterize the effectiveness of individual methods for determining synthetic variables. These measures can be divided into five groups, each of which includes measures related to different properties of synthetic variables (Grabiński et al., 1989; Bąk, 2018; Sompolska-Rzechuła, 2020):

1) mapping compatibility ($m_1 - m_3$),
2) linear correlation of the synthetic variable with diagnostic variables ($m_4 - m_5$),
3) rank correlation of the synthetic variable with diagnostic variables ($m_6 - m_8$),
4) variability and concentration of the synthetic variable ($m_9 - m_{10}$),
5) taxonomic distance of the synthetic variable from the original variable ($m_{11} - m_{12}$).

The optimality measures for linear ordering procedures are shown in Table 1.

Table 1 The optimality measures for linear ordering procedures

<table>
<tr><th>Group of measures</th><th>Measure</th><th>Comments</th></tr>
<tr><td>(1)</td><td>$m_1 = \frac{\sum_{i=1}^{n-1}\sum_{j>i}^{n}(d_{ij} - \tilde{d}_{ij})^2}{\sum_{i=1}^{n-1}\sum_{j>i}^{n} d_{ij}^2}$
$m_2 = \frac{2}{n(n-1)}\sum_{i=1}^{n-1}\sum_{j>i}^{n}\left(\frac{d_{ij} - \tilde{d}_{ij}}{d_{ij}}\right)^2$
$m_3 = \frac{\sum_{i=1}^{n-1}\sum_{j>i}^{n}(d_{ij} - \tilde{d}_{ij})^2 / \tilde{d}_{ij}}{\sum_{i=1}^{n-1}\sum_{j>i}^{n} d_{ij}}$</td><td>$\tilde{d}_{ij}$ – average distance between the <i>i</i>-th and <i>j</i>-th object in the <i>k</i>-dimensional space of diagnostic variables,
d_{ij} – distance between the <i>i</i>-th and <i>j</i>-th object in the one-dimensional space of the synthetic variable,
<i>n</i> – number of objects</td></tr>
<tr><td>(2)</td><td>$m_4 = 1 - \frac{1}{k}\sum_{j=1}^{k} r_{qx_j}$
$m_5 = \frac{1}{k}\sum_{j=1}^{k} l_j$</td><td>$r_{qx_j}$ – correlation coefficient between the <i>j</i>-th diagnostic variable and synthetic variable,
<i>k</i> – number of features and
<table>
<tr><td></td><td>l_j</td><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>r_{qx_j}</td><td>max</td><td>1.0</td><td>0.5</td><td>0.0</td><td>–0.5</td></tr>
<tr><td></td><td>min</td><td>0.5</td><td>0.0</td><td>–0.5</td><td>–1.0</td></tr>
</table></td></tr>
<tr><td>(3)</td><td>$m_6 = 1 - \frac{1}{k}\sum_{j=1}^{k} \rho_{qx_j}$
$m_7 = \frac{1}{k}\sum_{j=1}^{k} l_j$
$m_8 = \frac{2}{ml}\sum_{i=1}^{n}\sum_{j=1}^{k} \lvert x_{ij}^{\cdot} - q_i^{\cdot} \rvert$</td><td>$\rho_{qx_j}$ – Spearman's rank correlation coefficient between the <i>j</i>-th diagnostic variable and the synthetic variable,
$x_{ij}^{\cdot}$ – rank of the <i>i</i>-th object due to the j-th primary variable,
$q_i^{\cdot}$ – rank of the <i>i</i>-th object due to synthetic variable,
$l = n^2$ for <i>n</i> even and
$l = n^2 - 1$ for odd <i>n</i></td></tr>
</table>

Table 1 (continuation)

Group of measures	Measure	Comments
(4)	$m_9 = \frac{s_\mu}{\bar{\mu}}$ $m_{10} = \frac{s_\Delta}{\bar{\Delta}}$	$\bar{\mu}$ and s_μ – mean and standard deviation of the synthetic variable, $\bar{\Delta}$, s_Δ – mean and standard deviation for: $\Delta_i = \tilde{\mu}_i - \tilde{\mu}_{i-1}$ ($i = 1, \ldots, n-1$) $\tilde{\mu}_i$ – ordered non-descending values of the synthetic variable
(5)	$m_{11} = \frac{1}{nk}\sum_{i=1}^{n}\sum_{j=1}^{k}\lvert x_{ij}' - \mu_i' \rvert$ $m_{11} = \left[\frac{1}{nk}\sum_{i=1}^{n}\sum_{j=1}^{k}\lvert x_{ij}' - \mu_i' \rvert^2\right]^{1/2}$	x_{ij}' – standardized value of *j*-th primary variable for *i*-th object, μ_i' – standardized value of the synthetic variable for *i*-th object

Source: Own elaboration based on Grabiński et al. (1989)

In this paper, some modifications were made that relate to the determination of the value $\tilde{d}_{ij}$ as distance between the i-th and j-th object in the k-dimensional space of diagnostic variables, compared to the information contained in (Grabiński et al., 1989, pp. 122–123). The calculations include the values of diagnostic variables after normalization (zero unitarization). The distance $\tilde{d}_{ij}$ was determined as the average distance in the k-dimensional space. The introduced modifications made it possible to obtain comparable values of $\tilde{d}_{ij}$ and d_{ij} and partial measures m_l. In this paper was made some modification of measure m_9 too. Grabiński et al. (1989) introduced a "minus" sign for the measure m_9. In this paper, the value of the measure m_9 was calculated without the "minus" sign. It was assumed that the lower value of the measure m_9 indicates a lower diversity of objects in terms of the phenomenon under consideration.

The following measures from individual groups were used in this study: m_1, m_4, m_6, m_9 and m_{11}.

The aggregation of partial measures was performed according to the following formula (Bąk, 2015):

$$M_q = \sqrt{\sum_{l=1}^{g} m_l^2}\,, \tag{14}$$

where: M_q – aggregate measure, m_l – partial measure ($l = 1, \ldots, g$), g – number of partial measures.

2 RESULTS AND DISCUSSION

To achieve the goal, data from the EUROSTAT database (2017) for 28 European Union countries was used (Sompolska-Rzechuła, 2020):

X_1 – live births per 1 000 population,
X_2 – deaths per 1 000 population,
X_3 – infant deaths rate per 1 000 population,
X_4 – natural increase per 1 000 population,
X_5 – age dependency (population aged 0–14 and 65 and more per 100 persons aged 15–64),
X_6 – activity rate in %,
X_7 – employment rate in %,
X_8 – unemployment rate in %,
X_9 – at-risk-of poverty rate in %,
X_{10} – severely materially deprived people in %,
X_{11} – GDP per capita in thous euro,
X_{12} – investment rate in %,
X_{13} – industrial production (2015 = 100),
X_{14} – obtaining primary energy per 1 000 inhabitants from renewable energy sources (in tone),
X_{15} – final energy consumption per capita (in thous. kgoe),

X_{16} – share of high-tech exports in total exports in %,

X_{17} – net current account balance in % of GDP.

The parametric Hellwig method was used to eliminate strongly correlated features and the final set of diagnostic features was obtained, taking into account: X_8, X_{10}, X_{12}, X_{14}, X_{16}, X_{17}.

Table 2 presents the values of the basic descriptive parameters of the features finally adopted for the study.

Table 2 Summary statistics

Statistics	**Variables**					
	X_8	X_{10}	X_{12}	X_{14}	X_{16}	X_{17}
Mean	8.65	8.93	20.01	0.53	12.26	2.16
Median	7.70	5.65	19.85	0.37	10.50	1.85
Minimum	4.00	0.70	11.40	0.04	3.80	–5.30
Maximum	23.60	31.90	29.30	1.92	24.20	8.40
Standard deviation	4.48	7.44	3.45	0.47	6.18	3.74
Variation coefficient (%)	51.82	83.35	17.22	87.91	50.43	172.76
Skewness coefficient	1.94	1.56	0.19	1.85	0.46	0.07

Source: Own elaboration based on Eurostat (2017)

All features are characterized by strong or very strong volatility, in addition, X_8, X_{10} and X_{14} are characterized by strong right-sided asymmetry.

In 2016, the lowest unemployment rate was recorded in the Czech Republic and the highest – in Greece. In many countries (Belgium, Cyprus, Finland, Spain, France, Croatia, Ireland, Italy, Lithuania, Latvia, Portugal, Slovenia and Slovakia), the unemployment rate was higher than 7.70%, i.e. the median. However, in countries such as: Austria, Bulgaria, Germany, Denmark, Estonia, Hungary, Luxembourg, Malta, Netherlands, Poland, Romania, Sweden and United Kingdom the unemployment rate in 2016 did not exceed the median value.

The highest level of the deeper material deprivation rate was recorded in Bulgaria (31.9%), and the lowest – in Sweden (0.7%). In addition, the value of this indicator higher than the median was observed in Cyprus, Spain, Ireland, Hungary, Greece, Croatia, Italy, Lithuania, Latvia, Poland, Portugal, Romania and Slovakia. In countries such as: the Czech Republic, Germany, Denmark, Estonia, Finland, France, Luxembourg, Malta, Netherlands, Sweden, Slovenia and United Kingdom level of the deeper material deprivation rate was lower as median (5.7%).

The greatest diversity of countries is due to the current account balance as a percentage of GDP. Some countries, such as Belgium Cyprus, Greece, Finland, France, Lithuania, Poland, Romania, Slovakia and the United Kingdom recorded a negative balance in 2016. In Cyprus, the balance was the lowest and amounted to –5.3%. The highest positive balance of 8.4% was recorded in the Netherlands.

The average investment rate for the 28 EU countries amounted to 20.01% in 2016. In eleven countries, the investment rate above the average value was observed, and in Ireland its level was the highest and amounted to 29.30%. However, the lowest value (11.40%) occurred in Greece.

Malta (24.2%) has the highest percentage of high technology exports in total exports, while Portugal has the lowest – 3.8%. In addition, countries such as Austria, Cyprus, the Czech Republic, Germany, Estonia, France, Hungary, Luxembourg, the Netherlands, Sweden and the United Kingdom recorded a percentage of high technology exports in total exports above the average.

The non-pattern and pattern methods were used for comparison of the results of the linear ordering of European Union countries by socio-economic situation in 2016. In case of non-pattern methods, they were based on the zero unitarization to standardize variables and quotient transformation with an arithmetic mean. While in the analysis using standard methods at the stage of feature standardization, for the Hellwig and TOPSIS methods – the zero unitarization was used. The features: X_8 and X_{10} were recognized as destimulants and they were transformed into stimulants by means of quotient transformation as the inverse of the feature's value.

Table 3 presents the results of the linear ordering of EU countries by socio-economic situation in 2016.

Table 3 Results of the linear ordering of EU countries by socio-economic situation in 2016

Country	**Method**			
	a non-pattern object based on		**a pattern object**	
	zero unitarization (1)	**quotient transformation (2)**	**Hellwig (3)**	**TOPSIS (4)**
Austria (AT)	7	7	2	7
Belgium (BE)	17	20	16	17
Bulgaria (BG)	20	15	23	19
Croatia (HR)	21	17	22	21
Cyprus (CY)	26	28	27	22
Czech Republic (CZ)	6	13	7	6
Denmark (DK)	10	5	9	10
Estonia (EE)	8	8	3	8
Finland (FI)	11	9	11	11
France (FR)	13	16	13	13
Germany (DE)	3	3	4	3
Greece (EL)	28	27	28	28
Hungary (HU)	12	12	12	12
Ireland (IE)	4	10	10	4
Italy (IT)	24	18	24	26
Latvia (LV)	15	14	14	16
Lithuania (LT)	23	23	21	24
Luxembourg (LU)	9	6	8	9
Malta (MT)	2	4	6	2
Netherlands (NL)	5	2	5	5
Poland (PL)	19	21	18	20
Portugal (PT)	27	22	26	27
Romania (RO)	18	26	19	18
Slovakia (SK)	22	24	20	23
Slovenia (SI)	14	11	15	15
Spain (ES)	25	19	25	25
Sweden (SE)	1	1	1	1
United Kingdom (UK)	16	25	17	14

Source: Own elaboration based on Eurostat (2017)

Table 4 presents the evaluation of order compliance with selected methods measured by the Kendall rank correlation coefficient (1948, p. 82).

Table 4 Values of the Kendall rank correlation coefficients according to individual methods

Method	(1)	(2)	(3)	(4)
(1)	1.000	0.688	0.868	0.958
(2)	0.688	1.000	0.683	0.656
(3)	0.868	0.683	1.000	0.825
(4)	0.958	0.656	0.825	1.000

Source: Own elaboration based on Table 3

Assessment of order compliance with selected methods, measured by Kendall rank correlation coefficient, indicates the existence of significant links between country positions. The strongest correlation was observed between orders made using the non-pattern with zero unitarization and TOPSIS methods, which were obtained on the basis of zero unitarization and between the non-pattern method with Hellwig method. While the weakest relationship occurs between the results according to the following methods: non-pattern using the quotient transformation with the arithmetic mean and the pattern Hellwig method and TOPSIS based on zero unitarization.

When analysing the position occupied by individual countries, it can be seen that some countries took the same or similar position in individual orders, e.g. Austria, Estonia, France, Germany or Hungary. While in case of Bulgaria, Denmark or Ireland one can notice differences in the positions occupied in the obtained orders. The question arises, the results of which ordering should be adopted as optimal? In response to this question, help is provided by partial measures of the optimality of linear ordering procedures and the aggregate measure determined on their basis, the values of which for individual procedures are presented in Table 5.

Table 5 Values of the aggregate measure of optimality of procedures for linear ordering of EU countries in terms of socio-economic situation

Measure	Method			
	a non-pattern object based on zero unitarization (1)	a non-pattern object based on quotient transformation (2)	Hellwig based on zero unitarization (3)	TOPSIS method with zero unitarization (4)
M_q	0.860	1.101	0.925	0.729
m_1	0.331	0.157	0.238	0.204
m_4	0.463	0.641	0.411	0.405
m_6	0.435	0.467	0.578	0.431
m_9	0.379	0.514	0.500	0.304
m_{11}	0.288	0.542	0.212	0.220

Source: Own elaboration based on Eurostat (2017)

The optimality assessment of the linear ordering procedures can be performed using the following methods by comparing the results obtained with:

- all methods,
- a pattern object or a non-pattern object,

- methods based on the same way of standardizing features.

In assessing the optimality of linear ordering procedures, the criteria characterized in the chapter devoted to the research method were taken into account, i.e. mapping compatibility, linear and rank correlation of the synthetic variable with diagnostic variables, and variability of the synthetic variable and taxonomic distance of the synthetic variable from the original variable.

Taking into account the results of all methods, the most correct way of linear ordering of objects is to order EU countries obtained using the TOPSIS method based on zero unitarization (4). Also the results of this method are "better" compared to the results obtained according to Hellwig pattern method based on zero unitarization (3). The ordering obtained according to non-pattern method based on zero unitarization (1) gave more correct results compared to the ordering using non-pattern method based on quotient transformation (2).

Of the methods based on zero unitarization, the results obtained using the TOPSIS method (4) and then method non-pattern (1) were the most correct.

Taking into account the information obtained on the basis of the results included in Table 5, including the above-mentioned criteria, the results obtained using the TOPSIS method based on zero unitarization as the normative formula were considered the most correct synthetic variable (4).

In this order, Sweden came first, followed by Malta and Germany. The last place was occupied by Greece, Italy and Portugal occupied only slightly better places.

Sweden obtained its first place due to the favourable values of many features adopted in the study. The unemployment rate was 6.6% (only in the Czech Republic a lower value was observed – 4%). The in-depth deprivation rate was lowermost – 0.7%. In addition, the value referring to obtaining primary energy from renewable energy sources (thousand tonnes) per 1 000 inhabitants was at a high level. In 2016, Germany recorded a positive balance of the current balance of payments account in % GDP.

Greece occupied the last place in the linear ordering of countries, with the highest unemployment rate (23.6%) and the lowest investment rate (11.4%), a high deep deprivation rate of 22.2%, which is almost two and a half times higher than the average for all countries. Primary energy extraction from renewable energy sources (thousand tonnes) per 1 000 inhabitants was in Greece at one of the lower levels (18.95) and constituted only 16% of the highest value of this indicator, which concerned Austria. The share of exports of high technology products in total exports was also very low, at 4.6% with a maximum value of 24.2% for Malta. A negative value was recorded for the current account balance of payments in % GDP of – 0.6%.

CONCLUSION

The paper presents the results of a comparative analysis of four methods of linear ordering and two methods of variable standardization (quotient transformation and zero unitarization). The study was conducted on the basis of data on 28 European Union countries due to the level of socio-economic development in the context of sustainable development. The paper proposes a comparison of the results obtained in various methods concerning, e.g. the method of normalizing diagnostic features, the type of methods (based on a pattern object or a non-pattern object) and including all procedures. The assess of selected methods of linear ordering and selection of the optimal method was carried out on the basis of measures of correctness of the procedures of linear ordering of objects. Measures based on different properties of the synthetic variable were used in this paper. Some modifications of the measures were proposed regarding the measures of the mapping compliance and variability of the synthetic variable. The least correct procedure of linear ordering of objects among the methods selected for the study was non-pattern method based on quotient transformation. Comparing the results of two non-pattern methods based on different methods of standardization, the results obtained with the use of zero unitarization proved to be better. Taking into account all methods, the TOPSIS method with zero unitarization proved to be the most correct.

References

ANDERSON, T. W. *An Introduction to Multivariate Statistical Analysis.* 2nd Ed. New York: Wiley and Sons, 1984.

BĄK, A. Application of multivariate analysis methods comparative assessment of the state of the environment in the Dolnośląskie Voivodeship. *The Polish Statistician*, 2018, 1(680), pp. 7–20 (in Polish: Zastosowanie metod wielowymiarowej analizy porównawczej do oceny stanu środowiska w województwie dolnośląskim, *Wiadomości Statystyczne*).

BĄK, A. Problem of Choosing the Optimal Linear Ordering Procedure in the Pllord Package. *Research Papers of Wrocław University of Economics*, 2015, 384, pp. 33–41 (in Polish: Zagadnienie wyboru optymalnej procedury porządkowania liniowego w pakiecie pllord, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*). DOI: 10.15611/pn.2015.384.03

CERIN, P. Bringing Economic Opportunity into Line With Environmental Influence: a Discussion on the Coase theorem and the Porter and van der Linde hypothesis. *Ecological Economics*, 2006, pp. 209–225.

CHEN, C.-T. Extensions of the TOPSIS for Group Decision-Making under Fuzzy Environment. *Fuzzy Sets and Systems*, 2000, 114, pp. 1–9.

EUROSTAT. *Database – Eurostat* [online]. Eurostat, 2017. [cit. 10.9.2019]. <https://ec.europa.eu/eurostat/data/database>.

GRABIŃSKI, T. *Taxonometry Methods.* Cracow: Cracow University of Economics, 1992 (in Polish: *Metody taksonometrii*, Wydawnictwo Akademii Ekonomicznej).

GRABIŃSKI, T., WYDYMUS, S., ZELIAŚ, A. *Numerical Taxonomy Methods in Modeling Socio-Economic Phenomena.* Warsaw: PWN, 1989 (in Polish: *Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych*).

HARTIGAN, J. A. *Clustering Algorithms.* New York: Wiley and Sons, 1975.

HELLWIG, Z. Application of the Taxonomic Method to the Typological Division of Countries by Level of Their Development as well as Resources and Structure of Qualified Staff. *Statistical Review*, 1968, 4, pp. 307–327 (in Polish: Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr, *Przegląd Statystyczny*).

HWANG, C. L. AND YOON, K. *Muliple Attribute Decision Making. Methods and Applications.* Berlin: Springer, 1981.

KENDALL, M. G. *Rank Correlation Methods.* Londyn: Charles Griffin & Company Limited, 1948.

KUKUŁA, K. *Zero Unitarization Method.* Warsaw: Scientific Publisher PWN, 2000 (in Polish: *Metoda unitaryzacji zerowanej*, Wydawnictwo Naukowe PWN).

LATOSZEK, E. The Concept of Sustainable Development in UN Theory and Practice. In: LATOSZEK, E., PROCZEK, M., KRUKOWSKA, M. (eds.) *Sustainable Development and Global Public Goods in the Theory and Practice of International Organizations*, Warsaw: Warsaw School of Economics Press, 2016. (in Polish: Koncepcja zrównoważonego rozwoju w teorii Ii praktyce ONZ, *Zrównoważony rozwój a globalne dobra publiczne w teorii i praktyce organizacji międzynarodowych*).

MORRISON, D. F. *Multidimensional Statistical Analysis.* Warsaw: PWN, 1990 (in Polish: *Wielowymiarowa analiza statystyczna*).

NOWAK, E. *Taxonomic methods in the classification of socio-economic objects.* Warsaw: PWE, 1990 (in Polish: *Metody Taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*).

PLUTA, W. *Multidimensional Comparative Analysis in Economic Research.* Warsaw: PWE, 1977 (in Polish: *Wielowymiarowa analiza porównawcza w badaniach ekonomicznych*).

SEBER, G. A. F. *Multivariate Observations.* New York: Wiley and Sons, 1984.

SOMPOLSKA-RZECHUŁA, A. Selection of the method of linear ordering using the example of assessing the level of socio-economic development of European Union countries. *Research Papers of Wrocław University of Economics*, 2020, 64(7), pp. 118–129 (in Polish: *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*).

SOMPOLSKA–RZECHUŁA, A. *The Measurement and Assessment of the Quality of Life in the Region.* Szczecin: University Publishing House of the West Pomeranian University of Technology in Szczecin, 2018 (in Polish: *Pomiar i ocena jakości życia w ujęciu regionalnym*, Wydawnictwo Uczelniane Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie).

UNITED NATIONS GENERAL ASSEMBLY. *Report of the World Commission on Environment and Development: Our Common Future* [online]. Oslo, Norway: United Nations General Assembly, Development and International Cooperation Environment, 1987. [cit. 26.10.2019]. <https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf>.

WYSOCKI, F. *The Methods of Taxonomy for Recognition of Economic Types in Agriculture and Rural Areas*. Poznan: University Publisher Poznan University of Life Sciences, 2010 (in Polish: *Metody taksonomiczne w rozpoznawaniu typów ekonomicznych rolnictwa i obszarów wiejskich*, Wydawnictwo Uniwersytetu Przyrodniczego w Poznaniu).