Balezentis, T. (2022). The kernel-based comprehensive aggregation PROMETHEE (PROMETHEE-KerCA) method for multi-criteria decision making with application to policy modelling. *Journal of International Studies, 15*(1), 63-77. doi:10.14254/2071-8330.2022/15-1/4

Journal of International Studies

Scientific Papers

The kernel-based comprehensive aggregation PROMETHEE (PROMETHEE-KerCA) method for multi-criteria decision making with application to policy modelling

Tomas Balezentis
Faculty of Economics and Business Administration,
Vilnius University,
Lithuania
tomas.balezentis@evaf.vu.lt
ORCID 0000-0002-3906-1711

Abstract. As the economic and technological problems become more complex and require effective multi-criteria decision making (MCDM) tools for analysis thereof, there is a need for comprehensive MCDM techniques that would be capable to ensure robust optimization with minimum arbitrary assumptions. This paper proposes a new method for MCDM – the Kernel-based Comprehensive Aggregation PROMETHEE (PROMETHEE-KerCA). The proposed approach relies on the kernel density estimation which provides the bandwidths for scaling the differences in the performance of the alternatives. The kernel-based distances are aggregated to establish the performance measures thus following the principle of the outranking. Then, the measures of performance are aggregated in four different manners (additive, multiplicative, minimum and maximum values) to construct the comprehensive overall utility score. The proposed method does not require choosing the preference functions or parameters thereof. The empirical illustration is provided to show the feasibility of the proposed approach. The European Union Member States are ranked by the means of the KerCA method with regards to the objectives of the strategy Europe 2020. The isolated and pooled ranking allows comparing the progress of the countries compared with their initial situation and compared to the other countries in the sample.

Received: May, 2021
1st Revision: January, 2022
Accepted: March, 2022

DOI: 10.14254/2071-8330.2022/15-1/4

Keywords: multi-criteria decision making, kernel-based comprehensive aggregation, least squares cross validation, ranking, aggregation.

JEL Classification: C44, O2

1. INTRODUCTION

Multi-criteria decision making (MCDM) methods are important in identifying the compromise solutions for technological and economic problems. They allow aggregating multiple criteria expressed in different dimensions in order to compare the alternatives. The aggregation can be carried out following different theoretical assumptions (Mohammadi & Rezaei, 2020; Hashemkhani Zolfani et al., 2021; Janovská et al., 2021; Ahmed et al., 2020). The basic classification of the MCDM methods was offered by Belton & Stewart (2002) who identified the value measurement models, reference point models and outranking models. The property of compensation (Bouyssou, 1986) is also important in MCDM as it shows whether the utility of alternatives underperforming in terms of certain criteria can be increased due to outperforming others in some other criteria.

The value measurement models are based on the Multi-attribute Utility Theory (Keeney & Raiffa, 1976). Measures of this type seek to aggregate the normalized values of multiple criteria into single numbers which represent utility of a certain alternative. The aggregation is carried out for each alternative independently with regards to implicitly assumed point of origin (reference point) (Wierzbicka, 2020; Pietrzak, 2019; Bal-Domańska et al., 2020). The MCDM methods in this group include the Simple Additive Weighting (SAW) developed by MacCrimmon (1968) where the additive function is used to aggregate the weighted normalized values. The Analytic Hierarchical Process (AHP) developed by Saaty (1997) establishes the utility function via the pair-wise comparisons. The utility function can take multiplicative form. The Weighted Aggregated Sum Product (WASPAS) method was proposed by Zavadskas et al. (2012). The WASPAS method involves both additive and multiplicative utility functions thus generalizing the earlier approaches. The combined compromise solution (CoCoSo) method was proposed by Yazdani et al. (2019). It aggregates the additive and multiplicative utility functions during the two-step normalization.

The reference point approach relies on explicitly defined reference values (e.g., minima/maxima observed in the data or assumed theoretically, mean values). Therefore, the ideal or typical alternatives are used as benchmarks. The alternatives are then positioned with regards to the reference point(s) in order to derive the utility scores. The technique for order performance by similarity to ideal solution (TOPSIS) was introduced by Hwang & Yoon (1981). The TOPSIS identifies the ideal and negative-deal solutions and applies the Euclidean distance to aggregate the deviations across the criteria (Małkowska et al., 2021; Vavrek & Kovářová, 2021). The VIKOR method was proposed by Opricovic &Tzeng (2004). They suggested using the Manhattan and Chebyshev metric to aggregate the information with ideal alternatives used for linear normalization. Keshavarz Ghorabaee et al. (2016) presented the combinative distance-based assessment (CODAS) method. The CODAS relies on the Euclidean and Manhattan distances along with a threshold parameter. Zavadskas &Turskis (2010) proposed the additive ratio assessment (ARAS) method. The ARAS method introduces the ideal alternative in the decision matrix and then normalizes the resulting utility scores which is somewhat similar to the idea of the VIKOR and SAW. Keshavarz Ghorabaee et al. (2015) proposed the evaluation based on distance from average solution (EDAS) method, the EDAS method utilizes the mean values as the reference point and considers the positive and negative deviations from this point.

There have also been methods covering different types of MCDM techniques (Liao et al., 2020; Peng et al., 2021). Even though many of the aforementioned MCDM methods involve both additive or multiplicative utility functions and some kind of reference points, the Multi-Objective Optimization by Ratio analysis plus Full Multiplicative Form (MULTIMOORA) method (Brauers & Zavadskas, 2010) explicitly applied three different approaches. Specifically, it unifies additive and multiplicative value measurement along with the reference point and Chebyshev metric. Therefore, different combinations of the existing approaches can be applied for the MCDM.

The outranking methods are more complex in their computational approach. They allow for partial ranking which may be reasonable in case the differences among the alternatives under consideration are required to exceed a certain threshold. However, this is governed by the preference functions which require arbitrary choice of the functional form and parameters. Therefore, the decision process involves a substantial degree of subjectivity. The outranking approach involves the pair-wise comparisons of the alternatives under consideration. The ELECTRE (Benayoun et al., 1966) and PROMETHEE (Preference Ranking Organization Method for Enrichment Evaluations; Brans, 1982) are the two well-known outranking methods. These methods have been extended in a number of ways (Yu et al., 2018; Behzadian et al., 2010).

Therefore, there is a need for unifying the desirable properties of different MCDM methods in order to obtain robust results. In this paper, we seek to propose a novel MCDM method, namely the Kernel-based Comprehensive Aggregation (KerCA) method. We suggest exploiting the kernels and bandwidths in the pair-wise analysis of the alternatives. The kernels have been applied in decision making by, e.g., Albuquerque & Montenegro (2016) and Lin et al. (2020). The use of the kernels and Least Squares Cross-Validation allows for data-driven estimation of the bandwidths which can serve as the scaling measures similar to the preference function parameters in the traditional outranking methods. However, the KerCA method is devoid of the arbitrarily chosen preference functions. Specifically, we suggest applying the standard normal cumulative density function to establish the performance values of the alternatives. Then, four different aggregation methods are applied to derive the overall utility scores of the alternatives.

The paper proceeds as follows. Section 2 presents the key blocs for the multi-criteria ranking: the estimation of bandwidths and MCDM method KerCA. Section 3 presents the empirical case of the strategy Europe 2020 where the European Union Member States are ranked by the means of the KerCA. Finally, Section 4 concludes.

2. METHODS

The proposed MCDM approach relies on the kernel density estimation. Particularly, the bandwidth selection is the most important issue in this context. Therefore, we first discuss the preliminaries for data-driven bandwidth selection and then present the aggregation procedures for the MCDM.

2.1. Bandwidth estimation

Least Squares Cross-Validation (LSCV) is applied to compute the bandwidths (Silverman, 1986; Parmeter et al., 2017). The LCSV relies on the kernel estimators. The resulting bandwidths can be applied for the decision making problems, among multiple other possibilities.

Kernel is the function relating its arguments to the corresponding weights. The argument of the kernel function is the scaled distance between two data points. In this paper, we focus on the univariate kernel which considers single argument in the estimation.

Following Silverman (1986), the kernel satisfies the following properties:

$$\int K(t)\,dt = 1, \int tK(t)\,dt = 0, \int t^2 K(t)\,dt = k_2 \neq 0,$$

where $K(\cdot)$ is the kernel function differentiable to the required order. Obviously, the kernel function shares similar properties to the probability density function. The basic difference is that the kernel function is symmetric and centered at the point of origin.

There exist different types of the kernel functions, e.g., Epanechnikov, biweight, triangular, Gaussian or rectangular. The Gaussian kernel follows the shape of the Gaussian distribution. This kernel is not compact, i.e., its value only approaches zero asymptotically. The Gaussian kernel takes the following form (Silverman, 1986):

$$K(t) = \frac{1}{\sqrt{2\pi}} e^{-(1/2)t^2}, \quad (1)$$

where t is the point of evaluation.

For a univariate data series, $x_1, x_2, \ldots, x_m$, the kernel density at a certain point X is estimated as (Silvermna, 1986):

$$f(x) = \frac{1}{mh} \sum_{i=1}^{m} K\left(\frac{X - x_i}{h}\right), \quad (2)$$

where h is the bandwidth that governs the smoothness of the estimated density function. The choice of the bandwidth is important as it can mask the useful information if chosen improperly. Let us consider the mean integrated square error (MISE). The MISE can be defined as follows (Silverman, 1986):

$$MISE(f) = \int E\left(f(x) - f(x)\right)^2 dx = \int \left(E f(x) - f(x)\right)^2 dx + \int \operatorname{var}\left(f(x)\right) dx, \quad (3)$$

where the last two terms indicate the integrated bias and variance respectively. Therefore, the choice of the bandwidth should ensure that the MISE is minimized. In case a too small (resp. too large) bandwidth is chosen, the integrated variance (resp. bias) increases.

The optimal bandwidth can be estimated via the LSCV method. The essence of the LSCV is to minimize the integrated square error, $R\left(f\right) = \int f^2 - 2\int f f$, with respect to h. The kernel density is estimated by leaving each observation out. The procedure can be implemented by, e.g., the np package in R environment (Hayfield & Racine, 2008).

2.2. MCDM based on the kernel bandwidths

In the univariate space, one can easily rank the alternatives with regards to their performance as any two real numbers can be compared against each other once the direction of optimization is known. Therefore, we build our MCDM model on the univariate analysis repeated for each pair of alternatives for each criterion.

The MCDM problems involve multiple different dimensions (units of measurement). In order to account for these differences, the normalization techniques are applied. By using the kernel density approach, one may avoid the normalization and proceed directly with the original data set. Instead, a kind of standardization will be applied.

The proposed procedure for the KerCA method is outlined as follows:

Step 1. Decision matrix $\left[x_{ij}\right]_{m\times n}$ is established, where $i = 1, 2, \ldots, m$ is the index of alternatives and $j = 1, 2, \ldots, n$ is the index of criteria. Each criterion is attached with weight w_j such that $\sum_{j=1}^{n} w_j = 1$. Let there be two subsets of the criteria: subset *B* comprises the benefit criteria which should be maximized and subset *C* comprises cost criteria which should be minimized.

Step 2. For each criterion j, the kernel densities are estimated. The LSCV is applied. The resulting bandwidths are denoted by h_j, $j = 1, 2, \ldots, n$.

Step 3. For each criterion *j*, the alternatives are compared against each other taking into account the bandwidth h_j. The standard normal cumulative distribution function (cdf) is applied to summarize the results in a comparable manner. Thus:

$$p_{ij} = \frac{1}{m}\sum_{k=1}^{m}\Phi\left(\frac{x_{ij}-x_{kj}}{h_j}\right), j \in B,$$
$$p_{ij} = 1-\frac{1}{m}\sum_{k=1}^{m}\Phi\left(\frac{x_{ij}-x_{kj}}{h_j}\right), j \in C \tag{4}$$

where $\Phi(\cdot)$ is the standard normal cdf. Here, we suggest using the standard normal cdf for sake of brevity. However, one could exploit the cumulative kernel densities as measures of the distribution.

In this case, p_{ij} indicates the relative performance of alternative i against the other alternatives in terms of criterion *j*. note that the standard normal pdf is centered around 0.5. Therefore, values of p_{ij} greater than 0.5 indicate relatively good performance in a certain dimension of the assessment. The scaling is applied in Eq. 4 in order to ensure comparability across the criteria. Therefore, this step corresponds to the normalization stage.

This step is similar to the standardization when the values are compared to means and scaled by standard deviations. However, in this instance, we do not consider just a single observation (alternative) throughout the calculation. Rather, we present the pair-wise comparisons (distances) scaled by the bandwidth which based on the kernel density estimation. It could also be possible to avoid the cdf and use the kernel functions directly. In the case compact kernels are applied, one could set the thresholds of differences between any two arbitrarily chosen alternatives that should be ignored during the calculations. Also, the PROMETHEE method with the Gaussian preference function may be similar in this regard. However, the difference lies in the scaling as discussed above and the aggregation that proceeds as suggested below.

Step 4. The relative performance indicators are aggregated across the criteria for each alternative i to obtain the overall utility scores. We suggest using different instances of aggregation in this case. The two value measurement techniques are based on the additive and multiplicative aggregation, whereas the two measures based on the reference point approach focus on the minima and maxima for each alternative. Therefore, both the expected value and the upper and lower bounds of the performance values p_{ij} are taken into account.

The additive value measurement is facilitated as:

$$v_i^A = \sum_{j=1}^{n} w_j p_{ij}, i = 1,2,...,m. \tag{5}$$

The multiplicative value measurement proceeds as:

$$v_i^M = \prod_{j=1}^{n}\left(p_{ij}\right)^{w_j}, i = 1,2,...,m. \tag{6}$$

Reference point approach for the lower bounds of the performance values is implemented as follows:

$$r_i^{\min} = \min_{j=1,2,...,n} p_{ij}, i = 1,2,...,m. \tag{7}$$

Reference point approach for the upper bounds is applied as:

$$r_i^{\max} = \max_{j=1,2,\ldots,n} p_{ij}, i = 1,2,\ldots,m. \tag{8}$$

Step 5. The resulting measures need to be normalized prior to aggregation into the overall measure of utility. The utility scores are scaled by the observed maximum values and weighted during the aggregation:

$$\begin{aligned} u_i = \quad & \lambda_1 \frac{v_i^A}{\max_{i=1,2,\ldots,m} v_i^A} + \lambda_2 \frac{v_i^M}{\max_{i=1,2,\ldots,m} v_i^M} \\ & + \lambda_3 \frac{r_i^{\min}}{\max_{i=1,2,\ldots,m} r_i^{\min}} + \lambda_4 \frac{r_i^{\max}}{\max_{i=1,2,\ldots,m} r_i^{\max}} \end{aligned} \tag{9}$$

where $\lambda_l \in [0,1], l = 1,2,3,4$ governs the relative importance of the utility measures such that $\sum_l \lambda_l = 1$, $i = 1,2,\ldots,m,$. Indeed, the multiplicative value measure is more sensitive to extreme values (if compared to the additive one) and can affect the ranking of alternatives with such values of the criteria. The upper and lower bounds of the performance values are taken into account so that a non-compensatory mechanism would be involved in the analysis.

Step 6. The alternatives are ranked in descending order of u_i.

Note that the proposed method is characterized by the use of multiple different utility measures that are applied when constructing the overall utility measure. Indeed, there have been different methods proposed that involve additive, multiplicative or reference point approaches (Brauers & Zavadskas, 2010; Lakićević & Srđević, 2017). The KerCA method proposed in this paper offers even more complex structure of the utility function which may increase the discriminatory power.

3. EMPIRICAL ILLUSTRATION – STRATEGY EUROPE 2020

3.1. Data

We illustrate the proposed approach by considering the case of the European Union (EU) policies which seek to promote the competitiveness and sustainability of the EU economy. The Lisbon strategy in 2000 marked the initial attempt to propose an overarching strategy with corresponding structural indicators to track the progress of its implementation. The EU Member States then adopted the national plans to implement the strategic objectives. The Lisbon strategy focused on the period of 2000-2010. Currently, the EU is entering the ultimate stage of the strategy Europe 2020 which addresses multiple sectors in terms of cohesion, sustainability and growth (Soriano & Mulatero, 2010; Liobikiene & Butkus, 2017; Armstrong, 2012). The strategy focuses on the period of 2010-2020.

Rogge (2019) presented a benefit-of-the-doubt model along with the multiplicative index to assess the progress towards implementation of the objectives envisaged in the Europe 2020 strategy at the EU Member States level. The social, economic and environmental objectives were considered. The social objectives represent the inclusion in the labor market, education and poverty. The environmental objectives include the 20-20-20 goals for the renewable energy policy. The economic objectives include investments into research. Note that some indicators can be attributed to multiple groups (e.g., poverty rate identifies both social and economic objectives). We follow the latter study and consider the eight indicators representing the strategic goals of Europe 2020 (Table 1).

Table 1

Indicators for performance of the Europe 2020 strategy.

Notation	Criterion	Dimension	Type
		Employment	
C_1	Employment rate	% of population aged 20–64	*B*
		Research and Development	
C_2	Gross domestic expenditure on research and (experimental) development	% of GDP	*B*
		Climate change and energy	
C_3	Greenhouse gas emissions	% (compared to the 1990 level)	*C*
C_4	Share of renewable energy in gross final energy consumption	%	*B*
C_5	Gross primary energy consumption	% (compared to the 2005 level)	*C*
		Education	
C_6	Early leavers from education and training	% of population aged 18–24	*C*
C_7	Tertiary educational attainment	% of population aged 30–34	*B*
		Poverty and social exclusion	
C_8	Poverty and social exclusion	Number of persons who are at risk of poverty, severely materially deprived or living in households with very low work intensity (%)	*C*

Note: B and C indicate the sub-sets of the benefit and cost criteria respectively.
Source: own compilation

The indicator chosen for the analysis are relative ones and can be compared across the countries. However, certain indicators lack economic rationale. For instance, indicators related to energy use and greenhouse gas emissions are relative to the country-specific benchmarks (i.e., levels in the base periods). In this case, the economic growth remains ignored. However, we include these indicators in the analysis as they are used in the official documents and the Eurostat database (European Commission, 2020). The data for 2010-2018 are considered.

The distributions of the criterion values for 2010 and 2018 are provided in Fig. 1. The two distributions are compared by applying the non-parametric test by Li et al. (2009). This allows ascertaining whether the EU Member States achieved progress in terms of a certain indicator.

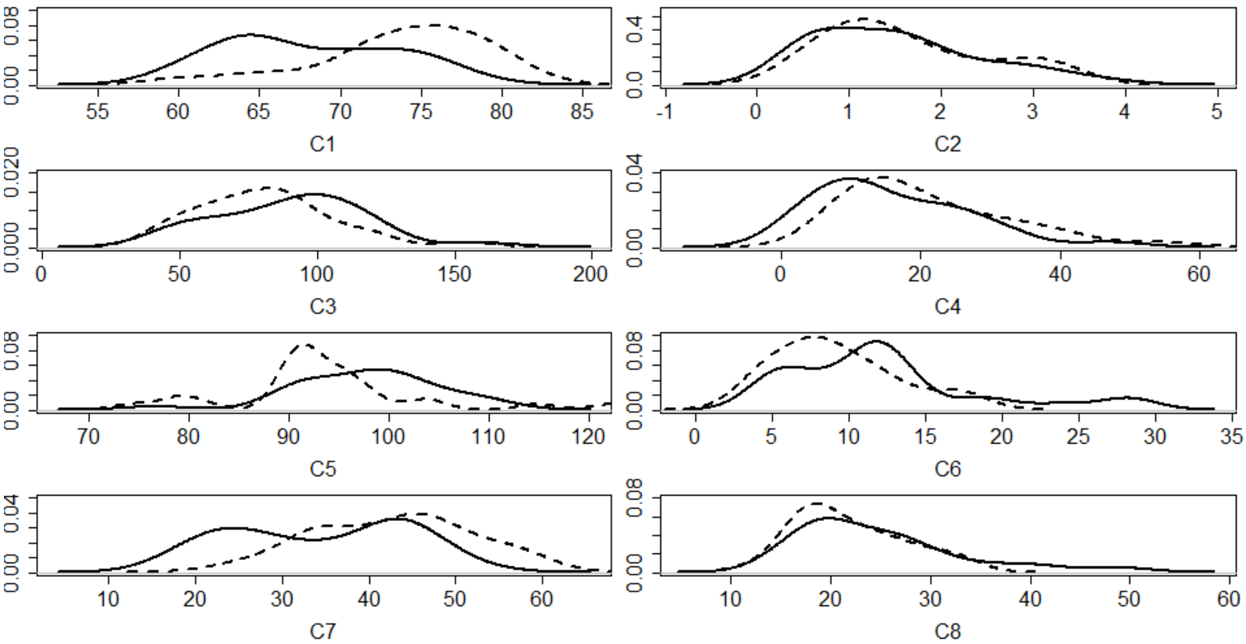

Figure 1. The kernel distributions of the criterion values for 2010 (solid line) and 2018 (dashed line)

Source: own calculation

For employment rate (C_1), the distribution shifted to the right during 2010-2018 which corresponds to an increasing employment rate. The distributions are significantly different ($p < 0.01$). There has been no progress in regards to expenditure on research and development (as measured in the percentage of the GDP, C_2). This is confirmed by both Fig. 1 and the statistical test ($p = 0.654$).

Greenhouse gas emission (C_3) shows a movement of the distribution towards left (i.e., a decline) during 2010-2018. However, this is not supported by the statistical test ($p = 0.178$) even though the associated p-value is marginally exceeding the 15% level of significance. The share of the renewable energy (C_4) shows a slight movement towards right which indicates an increase. The test does not indicate a significance difference ($p = 0.288$). The primary energy consumption (C_5) shows a decline in terms of the distribution movement. Furthermore, the test suggests a significant difference between the two distributions ($p = 0.012$).

The share of early leavers from education and training (C_6) declined significantly as suggested by the density plot moving towards left and the statistical test ($p = 0.148$) assuming the 15% level of significance. The share population completed the tertiary education (C_7) significantly went up during 2010-2018 ($p = 0.022$). Therefore, the education-related objectives were rather successfully implemented in the EU (in this case, we do not consider the country-specific targets but only the general trends). Note that C_7 shows bi-modal distribution which suggests that the use of the kernel density is more appropriate in this case than the parametric distributions.

The social cohesion has not been achieved as suggested by the stable distribution of the share of population facing poverty or social exclusion (C_8). The test fails to reject the null hypothesis of equality of distributions ($p = 0.684$). Therefore, the social dimension of the strategy requires more attention.

3.2. Data

We consider the case of the Europe 2020 strategy by embarking on the multi-criteria comparison of the EU Member States in terms of progress towards implementing objectives outlined in Table 1. As one can note, these indicators are measured in different dimensions and feature different directions of optimization. In the previous sub-section, we also showed that the actual trends of the changes in the criterion values do not necessarily correspond to the desirable ones. Therefore, we proceed with multi-criteria evaluation in order to identify the best- and worst-performing Member States.

The data for 2010 and 2018 are pooled into a single decision matrix. Then, the observations are ranked without considering the time periods they belong to. This allows not only to rank the countries during the two time periods, but also to meaningfully compare their progress in these two time periods. Thus, the changes in the relative performance of each country can be obtained. Besides, the overall utility scores are used to derive the isolated rankings for 2010 and 2018. This allows to track the changes in the relative positions across the two time periods. Therefore, we arrive at the three measures of performance: overall utility scores, pooled ranking and isolated ranking.

Note that indicators in Table 1 fall within the five thematic groups. Therefore, we assign these groups with equal importance (0.2). Then, the indicators are assigned with weights that add up to the group weight in the same manner. For instance, the three energy-related indicators are assigned with weights of $1/(5\times3)$ each. Even though more sophisticated approaches based upon expert assessment or data structure can be applied, we believe that taking into account all the thematic groups of indicators and assigning them with equal importance corresponds to the objectives of sustainable and inclusive growth stipulated in the strategy. We assign equal importance to the four utility measures in Eq. 9.

The pooled decision matrix is processed as outlined in Section 2.2. The bandwidths are obtained through the LSCV. The descriptive statistics, bandwidths and weights of the criteria are presented in Table 2.

Table 2

The descriptive statistics and parameters for the MCDM

Statistic	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
Min.	59.5	0.44	42.64	0.979	74.3	3.3	18.3	12.2
1st Q.	65.1	0.7825	67.07	10.405	91.03	6.95	31.88	18.23
Median	72.15	1.375	87.26	15.816	95.08	10.3	41.25	21.4
Mean	70.97	1.5688	85.94	18.499	95.63	10.7	38.57	23.26
3rd Q.	75.42	2.085	102.45	24.48	100.37	12.7	45.7	27.3
Max.	82.4	3.71	161.52	54.645	122.23	28.3	57.6	49.2
h_j	3.443	0.419	13.322	8.457	8.630	4.323	9.301	2.999
w_j	0.200	0.200	0.067	0.067	0.067	0.100	0.100	0.200

Source: own calculation

The results of the KerCA-based assessment of the EU Member Sates' performance are summarized in Table 3. The average values are provided for the measures based on the pooled data (i.e., utility scores and pooled ranks). Indeed, the averages for isolated ranks and their change have no meaning. The general trend is that the utility scores went up during 2010-2018 with the average growth of 32.7%. Correspondingly, the average decline in the ranks is -15.2 which indicates that the observations from 2018 dominate those from 2010.

Table 3

Multi-criteria ranking of the EU Member States based on the PROMETHEE-KerCA method

Member State	Utility score			Pooled rank			Isolated rank		
	2010	2018	Growth (%)	2010	2018	Change	2010	2018	Change
Austria	0.606	0.777	28.3	25	10	-15	7	9	2
Belgium	0.532	0.718	34.8	35	13	-22	13	12	-1
Bulgaria	0.390	0.520	33.1	51	39	-12	23	24	1
Cyprus	0.348	0.473	36.0	55	44	-11	27	27	0
Croatia	0.476	0.655	37.5	43	21	-22	17	16	-1
Czechia	0.585	0.842	43.9	27	5	-22	9	4	-5
Denmark	0.709	0.931	31.3	14	2	-12	2	2	0
Estonia	0.575	0.551	-4.1	29	32	3	11	20	9
Finland	0.704	0.874	24.2	15	4	-11	3	3	0
France	0.663	0.827	24.7	20	7	-13	5	6	1
Germany	0.671	0.834	24.2	19	6	-13	4	5	1
Greece	0.422	0.544	28.9	48	33	-15	20	21	1
Hungary	0.417	0.686	64.7	49	17	-32	21	14	-7
Ireland	0.520	0.628	20.8	37	24	-13	14	18	4
Italy	0.366	0.541	47.9	53	34	-19	25	22	-3
Latvia	0.465	0.645	38.7	45	23	-22	18	17	-1
Lithuania	0.580	0.770	32.8	28	11	-17	10	10	0
Luxembourg	0.567	0.692	22.0	31	16	-15	12	13	1
Malta	0.265	0.502	89.2	56	42	-14	28	26	-2
Netherlands	0.603	0.734	21.8	26	12	-14	8	11	3
Poland	0.454	0.526	15.9	46	36	-10	19	23	4
Portugal	0.384	0.572	49.0	52	30	-22	24	19	-5
Romania	0.411	0.512	24.7	50	41	-9	22	25	3
Slovakia	0.516	0.685	32.8	40	18	-22	16	15	-1
Slovenia	0.654	0.809	23.6	22	8	-14	6	7	1
Spain	0.360	0.447	24.0	54	47	-7	26	28	2
Sweden	0.882	1.000	13.4	3	1	-2	1	1	0
United Kingdom	0.520	0.791	52.0	38	9	-29	15	8	-7
Average	0.523	0.682	32.7	36.1	20.9	-15.2			

Note: observations for years 2010 and 2018 are ranked together for the pooled ranks and separately for the isolated ranks.

Source: own calculation

Comparison of the pooled ranks across the time periods for each country indicates how much they progressed in terms of the objectives of the Europe 2020 strategy. The only country showing decline is Estonia where the overall utility score declined by 4.1% and its results for 2018 are ranked below those for 2010 (pooled ranks of 32 and 29 respectively). Accordingly, the isolated ranks for Estonia show that it descended from the 11th place in 2010 down to 20th place in 2018.

Combining isolated and pooled ranking allows deriving conclusions regarding the progress of the countries compared with their initial situation in 2010 and compare to the other countries in the sample. Thus, even though the countries achieved progress with regards to their performance in 2010, the insolated ranking shows that some countries did not catch up with the overall sample growth. For instance, Poland

and Ireland show declines in the isolated ranking (descending by 4 places) even though their performance improved during 2010-2018 (utility scores went up by 16-20% and the pooled ranking suggests ascension).

Looking at the isolated ranking also allows one to identify the best- and worst-performing countries for each tine period. Obviously, Sweden, Denmark and Finland dominate the EU Member States in achieving the sustainable and inclusive growth. These results are valid for both 2010 and 2018. Therefore, the Nordic countries provide a useful case for further analysis. Still, even the lowest-ranking countries (e.g., Malta and Cyprus) managed to achieve progress as indicated by the pooled ranking.

As it was shown in Table 2 that the progress has been achieved in term s of the strategic objectives during 2010-2018. We further apply the non-parametric test (Li et al., 2009) to formally test the difference between the two distributions of the utility scores rendered by the KerCA (the pooled sample is used). The two distributions of the overall utility scores are provided in Fig. 2. As one can note, the 2018 distribution stochastically dominates the one for 2010. However, bimodality becomes more obvious in 2018 which indicates appearance of the two groups of countries. The statistical test confirms that the distributions for 2010 and 2018 are significantly different ($p < 0.01$).

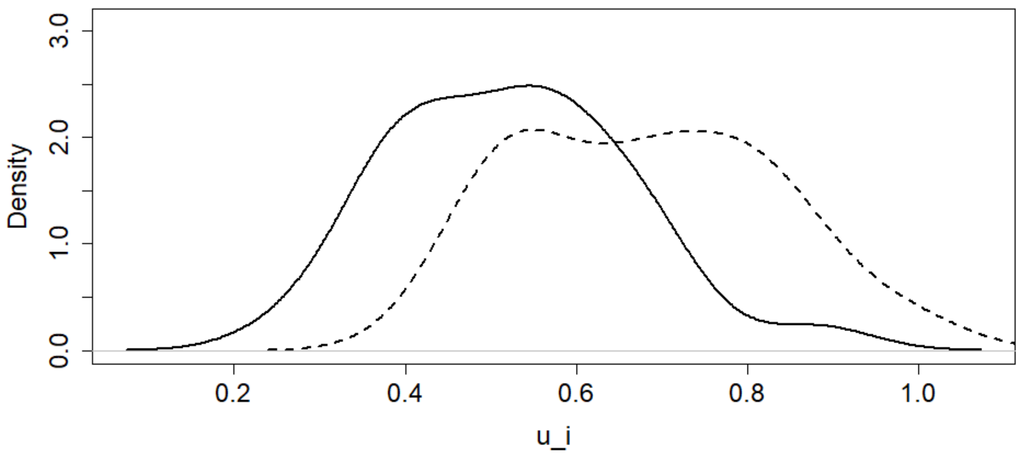

Figure 2. Kernel densities for the overall utility scores (u_i) for 2010 (solid line) and 2018 (dashed line)

Source: own calculation

The KerCA method involves four measures of utility as shown in Eq. 9. They are scaled with respect to the maximum values observed. We check if the ranking of the alternatives differs across the four measures involved in the construction of the overall utility score. Fig. 3 presents the graphical distribution. In this case, the points above or below the diagonal line indicate discordance in ranking across the two techniques. Table 4 presents the rank correlation coefficients for each pair of measures.

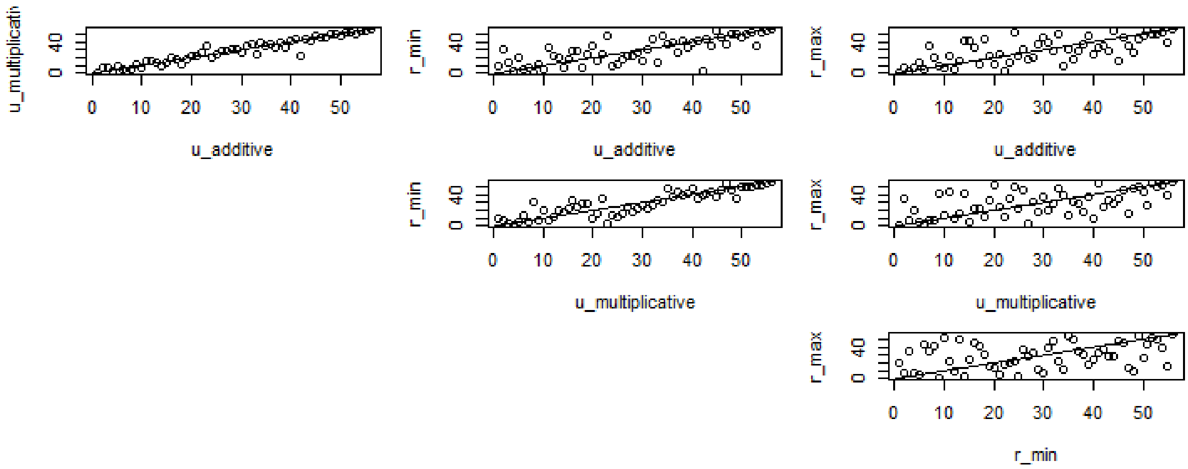

Figure 3. Ranking of the alternatives according to different utility measures used in the KerCA

Note: the normalized utility scores are used as shown in Eq. 9.

Source: own calculation

Table 4

Rank correlation between utility scores used in the KerCA

	multiplicative	min	max
additive	0.96	0.76	0.68
multiplicative		0.88	0.55
min			0.33

Note: the normalized utility scores are used as shown in Eq. 9.
Source: own calculation

The highest consistency in ranking is observed between the additive and multiplicative measures (R = 0.96). The data in Fig. 4 show that the discordance in ranking mostly occurs for the middle-ranked observations considering the latter pair of the utility measures. The coefficient of correlation declines for the other combinations of the utility measures and reaches the minimum value in the case of the reference point approach when rankings on the minimum and maximum preference values are compared for each alternative (R = 0.33). Therefore, the use of the multiple aggregation methods allows for a more comprehensive analysis and robust ranking. One can also adjust the weighting vector for the four measures of utility in Eq. 9 in order to take into account preferences towards the aggregation principles that may exist in the decision making process.

4. CONCLUSION

The paper proposed a novel MCDM method, namely the PROMETHEE Kernel-based Comprehensive Aggregation (KerCA). The new method allows for pair-wise comparison among the alternatives in the fashion of the outranking methods. The major difference is that the bandwidths obtained via the Leas Squares Cross Validation are used as the scaling variables rather than arbitrarily set parameters of the preference functions. The kernel-based distances are normalized by applying the standard normal cumulative distribution function. The resulting dimensionless numbers are aggregated over the alternatives to deliver the relative performance measures. These measures are then aggregated in the following ways:

- Additive value measurement,

- Multiplicative value measurement,
- Reference point approach for the minimum values, and
- Reference point approach for the maximum values.

Therefore, the proposed method combines the virtues of different types of the aggregation measures. The weighting can be applied to assign different importance to these four measures of utility when constructing the overall utility score.

The empirical illustration of the KerCA method dealt with the case of the Europe 2020 strategy. The European Union Member States were compared in terms of the strategic goals by applying the proposed method. The countries were ranked in a pooled manner (observations from two time periods) to identify their performance growth relative to the initial levels and an isolated manner (ranking for each time period separately) to compare the countries among themselves. The results also showed that the four aggregation techniques used in the construction of the overall efficiency score induced differences in the rankings. Thus, integration of these measures in the KerCA improves the robustness of the analysis.

The further research could aim to extend the proposed KerCA method with additional weight analysis tools. The Monte Carlo simulation could be adopted to check the stability of the ranking rendered by the KerCA. Finally, the averaging operators could be applied in the aggregation procedures underlying the KerCA.

REFERENCES

Ahmed, R. R., Romeika, G., Kauliene, R., Streimikis, J., & Dapkus, R. (2020). ES-QUAL model and customer satisfaction in online banking: evidence from multivariate analysis techniques. *Oeconomia Copernicana*, *11*(1), 59–93. https://doi.org/10.24136/oc.2020.003

Albuquerque, P. H. M., & Montenegro, M. R. (2016). PROMETHEE IV through kernel density estimation. *Communications in Statistics-Theory and Methods*, *45*(18), 5355-5362.

Armstrong, K. A. (2012). EU social policy and the governance architecture of Europe 2020. *Transfer: European Review of Labour and Research*, *18*(3), 285-300.

Bal-Domańska, B., Sobczak, E., & Stańczyk, E. (2020). A multivariate approach to the identification of initial smart specialisations of Polish voivodeships. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *15*(4), 785–810. https://doi.org/10.24136/eq.2020.034

Behzadian, M., Kazemzadeh, R. B., Albadvi, A., & Aghdasi, M. (2010). PROMETHEE: A comprehensive literature review on methodologies and applications. *European journal of Operational research*, *200*(1), 198-215.

Belton, V., & Stewart, T. (2002). *Multiple criteria decision analysis: an integrated approach*. Springer Science & Business Media.

Benayoun, R., Roy, B., & Sussman, B. (1966). ELECTRE: Une method pour guider le choix en présence de points de vue multiple, Note de travail 49. SEMA-METRA International, Direction Scientifique.

Bouyssou, D. (1986). Some remarks on the notion of compensation in MCDM. *European Journal of Operational Research*, *26*(1), 150-160.

Brans, J.P. (1982). Lingenierie de la decision. Elaboration dinstruments d'aide a la decision. Methode PROMETHEE. In: R. Nadeau, M. Landry (Eds.). Laide a la Decision: Nature, Instruments et Perspectives Davenir, Presses de Universite Laval, Quebec, Canada (1982), pp. 183-214.

Brauers, W. K. M., & Zavadskas, E. K. (2010). Project management by MULTIMOORA as an instrument for transition economies. *Technological and Economic Development of Economy*, *16*(1), 5-24.

Drumaux, A., & Joyce, P. (2020). New development: Implementing and evaluating government strategic plans—the Europe 2020 Strategy. *Public Money & Management*, *40*(4), 294-298.

European Commission. (2020). Eurostat - Europe 2020 indicators. https://ec.europa.eu/eurostat/web/europe-2020-indicators/europe-2020-strategy/main-tables

Hashemkhani Zolfani, S., Ebadi Torkayesh, A., Ecer, F., Turskis, Z., & Šaparauskas, J. (2021). International market selection: a MABA based EDAS analysis framework. *Oeconomia Copernicana*, *12*(1), 99–124.

Hwang, C. L., & Yoon, K. (1981). Methods for multiple attribute decision making. In *Multiple attribute decision making* (pp. 58-191). Springer, Berlin, Heidelberg.

Janovská, K., Vozňáková, I., Besta, P., & Šafránek, M. (2021). Ecological and economic multicriteria optimization of operating alternative propulsion vehicles within the city of Ostrava in the Czech Republic. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *16*(4), 907–943. https://doi.org/10.24136/eq.2021.034

Keeney, R.L., & Raiffa, H. (1976). *Decisions with multiple objectives: preferences and value tradeoffs*. Wiley, New York

Keshavarz Ghorabaee, M., Zavadskas, E. K., Olfat, L., & Turskis, Z. (2015). Multi-criteria inventory classification using a new method of evaluation based on distance from average solution (EDAS). *Informatica*, *26*(3), 435-451.

Keshavarz Ghorabaee, M., Zavadskas, E. K., Turskis, Z., & Antucheviciene, J. (2016). A New Combinative Distance-based Assessment (CODAS) Method for Multi-criteria Decision-making. *Economic Computation & Economic Cybernetics Studies & Research*, *50*(3), 25-44.

Lakićević, M. D., & Srđević, B. M. (2017). Multiplicative version of Promethee method in assesment of parks in Novi Sad. *Zbornik Matice srpske za prirodne nauke*, (132), 79-86.

Li, Q., Maasoumi, E., & Racine, J. S. (2009). A nonparametric test for equality of distributions with mixed categorical and continuous data. *Journal of Econometrics*, *148*(2), 186-200.

Liao, H.C., Xue, J.F., Nilashi, M., Wu, X.L., Antucheviciene, J. (2020). Partner Selection for Automobile Manufacturing Enterprises with a Qrung Orthopair Fuzzy Double Normalization-based Multi-aggregation Method", *Transformations in Business & Economics*, 19 (50A),.338-368.

Lin, M., Xu, W., Lin, Z., & Chen, R. (2020). Determine OWA operator weights using kernel density estimation. *Economic Research-Ekonomska Istraživanja*, *33*(1), 1441-1464.

Liobikienė, G., & Butkus, M. (2017). The European Union possibilities to achieve targets of Europe 2020 and Paris agreement climate policy. *Renewable Energy*, *106*, 298-309.

MacCrimmon, K. R. (1968). *Decisionmaking among multiple-attribute alternatives: a survey and consolidated approach*. Rand Corp Santa Monica Ca.

Małkowska, A., Urbaniec, M., & Kosała, M. (2021). The impact of digital transformation on European countries: insights from a comparative analysis. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *16*(2), 325–355. https://doi.org/10.24136/eq.2021.012.

Mohammadi, M., & Rezaei, J. (2020). Ensemble ranking: Aggregation of rankings produced by different multi-criteria decision-making methods. *Omega*, 102254.

Opricovic, S., & Tzeng, G. H. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. *European journal of operational research*, *156*(2), 445-455.

Parmeter, C. F., Wang, H. J., & Kumbhakar, S. C. (2017). Nonparametric estimation of the determinants of inefficiency. *Journal of Productivity Analysis*, *47*(3), 205-221.

Peng, X.Y., Luo, L., Liao, H.C., Zavadskas, E.K., Al-Barakati, A. (2021. A novel Decision-making Method for Resilient Supplier Selection during COVID-19 Pandemic Outbreak based on Hesitant Fuzzy Linguistic Preference Relations, *Transformations in Business & Economics*, 20 (54), 238-258.

Pietrzak, M. B. (2019). Modifiable Areal Unit Problem: the issue of determining the relationship between microparameters and a macroparameter. *Oeconomia Copernicana*, *10*(3), 393–417. https://doi.org/10.24136/oc.2019.019

Rogge, N. (2019). EU countries' progress towards 'Europe 2020 strategy targets'. *Journal of Policy Modeling*, *41*(2), 255-272.

Saaty, T.L. 1980. *The analytic hierarchy process: Planning, priority setting, resource allocation.* McGraw-Hill, New York.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis* (Vol. 26). CRC press.

Soriano, F. H., & Mulatero, F. (2010). Knowledge policy in the EU: From the Lisbon strategy to Europe 2020. *Journal of the Knowledge Economy*, *1*(4), 289-302.

Vavrek, R., & Kovářová, E. (2021). Assessment of the social exclusion at the regional level using multi-criteria approach: evidence from the Czech Republic. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *16*(1), 75–102. https://doi.org/10.24136/eq.2021.003

Wierzbicka, W. (2020). Socio-economic potential of cities belonging to the Polish National Cittaslow Network. *Oeconomia Copernicana*, *11*(1), 203–224. https://doi.org/10.24136/oc.2020.009

Yazdani, M., Zarate, P., Zavadskas, E. K., & Turskis, Z. (2019). A Combined Compromise Solution (CoCoSo) method for multi-criteria decision-making problems. *Management Decision*, 57(9), 2501-2519.

Yu, X., Zhang, S., Liao, X., & Qi, X. (2018). ELECTRE methods in prioritized MCDM environment. *Information Sciences*, *424*, 301-316.

Zavadskas, E. K., & Turskis, Z. (2010). A new additive ratio assessment (ARAS) method in multicriteria decision-making. *Technological and economic development of economy*, *16*(2), 159-172.

Zavadskas, E. K., Turskis, Z., Antucheviciene, J., & Zakarevicius, A. (2012). Optimization of weighted aggregated sum product assessment. *Elektronika ir elektrotechnika*, *122*(6), 3-6.