2-Tuple unbalanced linguistic multiple-criteria group decision-making using prospect theory data envelopment analysis

Imran Khan1 · Anjana Gupta2 · Aparna Mehra3

Accepted: 13 March 2022 / Published online: 4 April 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
The linguistic terms used to describe the qualitative data in problems are symmetrical about the central linguistic word. With the increasing complexity and a large number of DMs with different domain expertise involved in decision-making, the symmetric linguistic term set seems constricted. This paper examines the multiple-criteria group decision-making problems in which the decision-makers use the unbalanced 2-tuple linguistic term set to describe the alternative–criteria assessments matrices. We adopt a data envelopment analysis (DEA) approach and develop a linear programming model to evaluate alternatives criteria weights for each decision-maker. The non-rational factor of risk in criteria gets modeled using the value function from the prospect theory. The values of prospect gain and prospect loss, respectively, on cost and benefit parameters are computed and applied to formulate a DEA model to assess the weights of each criterion on each alternative. Eventually, the cross-efficiency scores and their entropy values provide a ranking of the alternatives. A numerical example is presented to illustrate the methodology. A comparative analysis of the proposed method with fuzzy ordinal priority approach and fuzzy TOPSIS is performed for the illustrative example. Finally, an application example of ranking the faculties on their end-semester feedback from their class students in an academic institute is presented.

Keywords Multiple-criteria group decision-making · 2-Tuple unbalanced linguistic term · Data envelopment analysis · Prospect value function · Cross-efficiency

1 Introduction
Over the last few decades, multi-criteria group decision-making (MCGDM) problems are widely studied and applied in a variety of areas, including supply chain management, medical diagnosis, alternative selection, technology evaluation, investment decision, to quote a few. For instance, refer to Dursun and Karsak (2013), Ehrrott et al. (2004), Liu et al. (2015), Yue (2013), and Yayla. A MCGDM method involves three essential entities: decision-makers, criteria, and alternatives, and, in essence, requires each decision-maker (DM) to provide the alternative–criteria matrices, which then are aggregated applying weights of the decision-makers (DMs) and weights of each criterion to arrive at a ranking of alternatives finally. The usual MCGDM modeling assumes that the entries in all alternative–criteria decision matrices to be precisely available. Therefore, it is inadequate to handle situations that can be described through precise quantitative values. The problem data are vague and may involve qualitative expressions. Researchers have expanded the scope of MCGDM to include vagueness in the model using a variety of tools. For instance, MCGDM problems are widely studied in a fuzzy environment using fuzzy sets (Zhang and Liu 2011), intuitionistic fuzzy sets (Efe 2016), type-2 fuzzy sets (Wang et al. 2017), hesitant fuzzy sets (Farhadinia and Xu 2020) and hesitant dual fuzzy sets (Zhang and Liu 2011), Pythagorean fuzzy sets (Fei and Deng 2020), and many others. The literature is very elaborate to cite all but a few.
The frameworks based on fuzzy sets or their generalized variants typically employ membership function scale representation. All computing is performed on this scale using fuzzy computing rules or some other defined mathematical operators in the context of research. These methods are not apt when the uncertainty in problems arises due to vagueness in the meanings of the qualitative assessments by DMs. Herrera and Martínez (2000) introduced the 2-tuple linguistic model to unfold the vague information and introduced the idea of computing with words (CWW). The language of CWW is simple to follow; it adds flexibility by allowing expressing information in natural language and provides meaning to the mathematical operations on such information. The 2-tuple linguistic model (Herrera and Martínez 2000) improves upon the basic linguistic model by including an additional parameter to bring in the translation process in computing. In other words, the 2-tuple linguistic model possesses a continuous scale, resulting in no loss of information while computing with linguistic terms. Typically, the linguistic term set is balanced, meaning that the words in set are uniform and symmetrically spread across the mid-term.

We believe that every DM may have a different perspective on assessing each alternative on different criteria, and the DM can express it in natural language. Moreover, the DM may weight favoring or positive viewpoints differently than the opposing viewpoint on objects. This paper aims to study the MCGDM problems where the DMs use the non-uniform, unsymmetrical 2-tuple linguistic term on alternatives–criteria valuation. If the information is unevenly distributed and possesses higher granularity on either side of the central term, (Liu et al. 2004; Torra 2000) comes with its challenges. Such a linguistic term set is called an unbalanced linguistic term set (Herrera et al. 2008b). The uneven spread could be due to a higher cardinality of terms on one side of the central term than the other, or we have an equal number of terms on both sides of the mid-term, but the spacing between them is unequal.

Recent years witnessed significant contributions in developing procedures analogous to the CWW in the unbalance linguistic term set. Herrera and fellow researchers (Herrera et al. 2008a; Dong et al. 2016) utilized the notion of linguistic hierarchy to allocate semantics to each linguistic term via a parametric membership function. Wang and Hao (2007) unified the frameworks of Lawry (2001) and 2-tuple linguistic model. Zou et al. (2012) proposed a linguistic aggregation operator elicited from the 2-tuple linguistic model and linguistic hierarchy. Dong and Cooper (2016) presented a new interpretation of the 2-tuple linguistic model by putting forward the notion of numerical scale. Zhang et al. (2020) studied large-scale group decision-making problems in which the decision-makers provide their multi-granular information using the hesitant fuzzy linguistic term sets. The authors also proposed two algorithms, one for transforming an unbalanced hesitant fuzzy linguistic term set into a balanced linguistic distribution, and the other transforms the balanced linguistic distribution to the unbalanced linguistic distribution assessment. More recently, Zhang et al. (2021b) put forward a simplified linguistic model for expressing multi-granular unbalanced linguistic term and developed two optimization models that generate adjustment advice for DMs in reaching consensus in group decision-making. Yu et al. (2021) build a consensus reaching model in a multi-criteria group decision-making involving multi-granular hesitant fuzzy linguistic term sets. Their model employs the trapezoidal fuzzy numbers representation of the multi-granular hesitant fuzzy linguistic term set. Malhotra and Gupta (2020a) proposed a novel idea to represent unbalanced linguistic information via a multiplicative linguistic label set. They developed a 2-tuple unbalanced linguistic term set applying the concept of minimum distance measure. Their multiplicative labeling model offers two advantages. One, it enables presenting a formal definition of unbalanced linguistic information. Secondly, it is computationally simple and less cumbersome than the linguistic hierarchy approach (Herrera and Herrera-Viedma 1997) and numerical scale concept (Dong and Cooper 2016).

In the aggregation process, the weights of decision-makers on each criterion and the weights of each alternative criterion are vital. The decision-makers assessments may not carry equal weight; based on the positional hierarchy, responsibility, professional expertise, and familiarity with the relevant problems, some experts may play more decisive roles than the others in a group. As a result, determining the weights of the decision-makers in a group is critical. Furthermore, a decision maker may not be an expert on each criterion to assess an alternative accurately. If not always, the participating decision-makers often employ quantitative descriptors to communicate their assessments of alternatives on criteria. In many studies, the weights of the DMs are given in advance. In contrast, in certain other studies, the decision-makers weights get determined by describe procedures. Geng et al. (2017) employed the data envelopment analysis (DEA) approach to solve MCGDM problems integrating 2-tuple linguistic DEMATEL for capturing influence relationships among criteria. Wan et al. (2016) studied the MCGDM problem with intuitionistic fuzzy numbers and determined DMs weights using TOPSIS. Dong and Cooper (2016) used the Markov chain method to determine DMs weights. Ju (2014) applied the similarity degree in the 2-tuple linguistic evaluation and defined the 2-tuple linguistic positive, negative, and left negative ideal solutions and proposed optimization models to determine the weights.

Against this brief backdrop, the foremost objective of the present study is to provide a ranking of the alternatives in an MCGDM problem when the decision-makers use the 2-tuple unbalanced linguistic term set to give the alternative–criteria
matrices entries. We use the DEA-based linear programming model to evaluate each DM weight on each alternative criteria pair. The DEA model provides a more reliable optimization framework yielding the global optimum weights and completely removes any subjective judgments in the process. These weights are applied to find the aggregate alternative–criteria matrix. Besides, the criteria are put in binary baskets of benefit and cost. Behavioral studies showed that the pain of loss (cost) of a value is greater in magnitude than the pleasure of gaining (benefit) the same value. We envision capturing the impact of criteria on the alternatives by integrating the cumulative prospect theory (CPT) value function-based DEA model that applies the prospect gain and loss values on the aggregated matrix’s cost and benefits criteria and computes criterion weight on each alternative. These weights help to compute the cross-efficiency scores and their entropy values to rank the alternatives finally.

DEA is a nonparametric technique to measure the relative efficiency of homogeneous decision-making units (DMU) that use multiple inputs to produce multiple outputs. In the context of MCGDM, the alternatives act as DMUs. Since pioneering work of Farrell (1957) and Charnes et al. (1978), DEA has come a long way in research and applications in a variety of domains. Each DMU self-evaluates its best relative efficiency using its most favorable input and output weights from the DEA model. On the other hand, the cross-efficiency, proposed by Sexton et al. (1986) and elaborated by Doyle and Green (1994), assessed the efficiency of each DMU not only through self-evaluation only but also takes into account the peer evaluation. The peer evaluation evaluates each DMU with the optimal weights of other DMUs. The cross-efficiency score is the average of self-evaluation efficiency and peer-evaluation efficiencies of a DMU.

Soleimani-Damaneh and Zarepisheh (2009) applied Shannon (1948) entropy to aggregate different efficiency results. Wu et al. (2012) introduced the entropy value of cross-efficiency score and developed a distance entropy function between cross-efficiency score and CCR efficiency score to generate relative weights for cross-efficiency aggregation. Xie et al. (2014) used Shannon’s entropy to derive the degree of importance of each DMU and merged the calculated efficiency scores and the degrees of importance to help discriminate between traditional DEA models. A modified weight-restricted DEA model for the derivation of non-zero optimal weights proposed by Qi and Guo (2014) using Shannon’s entropy to aggregate those weights to be the common weights. Wang et al. (2016) used the DEA entropy model to find the cross-efficiency intervals with imprecise inputs and outputs. The DMUs are ranked based on the distance to ideal CWW for it. We then evaluate the weights of each DM on each alternatives–criteria pair using a non-parametric DEA linear program model. Subsequently, we applied the CPT power function to accommodate the risk-averse perspective in computing the criteria weights. These weights finally led to comprehensive decision-making of ranking the alternatives.

We organize this paper with the following sections. Section 2 presents preliminary concepts, mainly the unbalanced linguistic term set and the prospect theory-based DEA model computing the cross-efficiency of alternatives. Section 3 explains the DEA model for an MCGDM problem with alternative–criteria matrices entries from the unbalanced linguistic term set. Section 4 describes the step-wise procedure to solve the underline MCGDM problem and ranking of the alternatives. Section 5 presents a numerical example to illustrate the proposed methodology. A comparative analysis of the proposed scheme with fuzzy ordinal priority approach and fuzzy TOPSIS is also included. Section 6 presents a case study to show the availability of the proposed model. Section 7 concludes the paper.

2 Preliminaries

In this section, we present two key ideas that we have applied in developing our scheme for MCGDM problems. We keep our descriptions brief for the two up to the point which are vital in the context of our study.

2.1 Unbalanced Linguistic term set

The 2-tuple linguistic model adds a new parameter called symbolic translation to the basic fuzzy linguistic model. The model is represented by a pair \((l, \alpha)\), where \(l \in L\) is a linguistic term from linguistic terms set \(L = \{l_0, l_1, \ldots, l_k\}\) and \(\alpha \in [-0.5, 0.5]\) is a numerical value for the symbolic translation. The 2-tuple linguistic model assists in explaining the process of CWW by imparting continuity to the linguistic domain. This model is also referred by continuous 2-tuple linguistic model. Martínez and Herrera (2012) and Malhotra and Gupta (2020b) presented a comprehensive review of the 2-tuple linguistic model and its applications to the decision-making problems.

The multiplicative unbalanced 2-tuple linguistic term (MU2LT) set is introduced by Malhotra and Gupta (2020a).

**Definition 1** Let \(LS = \{l_{a^{-}}, \ l_{a^{-}(−1)}, \ldots, l_{a0}, \ldots, l_{a^+}\}\) be a multiplicative unbalanced linguistic term (MULT) set with cardinality \(f + g + 1\), where \(f, g\) are positive integers, and \(a\) is a real number, \(a > 1\). Let \(\beta \in [a^{-f}, a^{+f}]\) be a numerical value representing the result of a symbolic aggregation. Then the unbalanced 2-tuple linguistic value that expresses...
the equivalent information to β is explained by the function \( \Omega : [a^{-f}, a^g] \rightarrow LS \times [0.5, -0.5] \) defined by

\[
\Omega(\beta) = (t_\lambda, \gamma),
\]

where

\[
\begin{align*}
\lambda &= a^{\alpha_L}, \quad \gamma \in \{ -f, -1, \ldots, 0, \ldots, g \}, \\
\sigma_L &= \{ -f, \ldots, 0, \ldots, g \}, \\
\sigma_R &= \{ -f, \ldots, 0, \ldots, g \}, \\
\end{align*}
\]

and the following procedure (Malhota and Gupta 2020a) is used to compute \( \sigma_L, \sigma_R, d_L, \) and \( d_R. \)

**Algorithm 1** Computing \( \lambda \) and \( \gamma \)

**Input:** \( LS = \{ t_{a^{-f_1}}, t_{a^{-f_2}}, \ldots, t_{a^{g_1}}, t_{a^{g_2}}, \ldots, t_{a^g} \} \) and \( \beta \in [a^{-f}, a^g] \);

**Begin:** \( \sigma_L = \arg \min_\sigma ( \beta - a^\sigma : \beta - a^\sigma \geq 0, \sigma \in \{ -f, \ldots, 0, \ldots, g \} ); \)

\( \sigma_R = \arg \max_\sigma ( \beta - a^\sigma : \beta - a^\sigma \leq 0, \sigma \in \{ -f, \ldots, 0, \ldots, g \} ); \)

**Compute:** \( d_L = \beta - a^{\sigma_L}, d_R = \beta - a^{\sigma_R}. \)

**Define:** \( d = d_L \) then \( \lambda = a^{\sigma_L}; \gamma = d_L + d_R. \)

else \( \lambda = a^{\sigma_R}; \gamma = d_L + d_R. \)

end if

**Definition 2** Let \( LS = \{ t_{a^{-f}}, t_{a^{-f_2}}, \ldots, t_{a^{g}}, t_{a^{g_2}}, \ldots, t_{a^g} \} \) be linguistic term set. The function \( \Omega^{-1} : LS \times [-0.5, 0.5] \rightarrow [a^{-f}, a^g] \) transforms the unbalanced 2-tuple linguistic variable \( t_\gamma, \gamma \) into its equivalent numerical value \( \beta \in [a^{-f}, a^g] \) by

\[
\Omega^{-1}(t_\lambda, \gamma) = \gamma (d_L + d_R) + \lambda.
\]

**Definition 3** Let \( ULS = \{ (t_{a^{\lambda_1}}, \gamma_{\lambda_1}), (t_{a^{\lambda_2}}, \gamma_{\lambda_2}), \ldots, (t_{a^{\lambda_n}}, \gamma_{\lambda_n}) \} \) be a set of \( n \) multiplicative unbalanced 2-tuples, where \( \lambda_i \in \{ -f, -1, \ldots, 0, \ldots, g \}, i = 1, \ldots, n, \) and \( t_{a^{\lambda_i}} \in LS \) and \( \gamma_{\lambda_i} \in [-0.5, 0.5]. \) Let \( W = (w_1, \ldots, w_n) \) be a weight vector, \( w_j \in (0, 1) \) and \( \sum_{i=1}^n w_i = 1. \) The multiplicative unbalanced 2-tuple weighted average operator (UWA) is defined as follows:

\[
UWA \left( (t_{a^{\lambda_1}}, \gamma_{\lambda_1}), (t_{a^{\lambda_2}}, \gamma_{\lambda_2}), \ldots, (t_{a^{\lambda_n}}, \gamma_{\lambda_n}) \right) = \Omega \left( \sum_{i=1}^n w_i \Omega^{-1}(t_{a^{\lambda_i}}, \gamma_{\lambda_i}) \right). \]

2.2 Prospect theory based cross-efficiency evaluation

Kahneman and Tversky (1979) put forward the prospect theory (PT) to model human behavior toward the preference structures or risk in the choices in an uncertain environment. The following three are the guiding principles behind their approach.

1. Reference dependence. A DM generally perceives outcomes as gains or losses relative to a reference point. The prospect value curve of a DM by the reference point is into the gain and loss domains.

2. Loss aversion. Decision-makers have different risk appetites toward gains and losses. In a decision process, a DM is typically more sensitive to losses than to absolute commensurate gains. For this purpose, the prospect value curve is steeper in the loss domain than the gain domain.

3. Diminishing sensitivity. A DM possesses a risk-averse tendency for gains and a risk-seeking tendency for losses. Consequently, the prospect value curve is concave in the gain domain and convex in the loss domain. The marginal values of both gains and losses decreases with an increase in their sizes.

One of the most common prospect value functions \( v(\Delta z) \) representing the subjective attitude of a decision-maker is the power function, defined as follows:

\[
v(\Delta z) = \begin{cases} 
(\Delta z)^\alpha & \Delta z \geq 0 \\
-\theta (\Delta z)^\phi & \Delta z < 0.
\end{cases}
\]

Here, \( \Delta z = z - z_o \), is the deviation of \( z \) from the reference point \( z_o \). It exhibits gain above the reference point and losses below the reference point through its non-negative and negative values. The parameters \( \alpha \in (0, 1) \) and \( \phi \in (0, 1) \) denote the bump degree in the gain domain and diminishing sensitivity in the loss domain of the power function, respectively. The greater values of \( \alpha \) and \( \phi \) indicate the greater risk appetite of the DM. The parameter \( \theta > 0 \) is a loss-aversion coefficient; \( \theta > 1 \) indicates that the decision-maker is more worrisome for losses than gains. Liu et al. (2014) suggested to take \( \alpha = \phi = 0.85, \) and \( \theta = 4.1. \) An experimental validation by Tversky and Kahneman (1992) suggested to take \( \alpha = 0.89, \phi = 0.92, \) and \( \theta = 2.25. \)

Tversky and Kahneman (1992) added a minor extensions to the prospect theory to present cumulative prospect theory (CPT) that employs cumulative rather than separable decision weights to the value function. The CPT has attracted attention in many diverse decision-making areas. For example, Harel et al. (2018) compared single-period utility of wealth functions with the value functions. Wang et al. (2020) introduced CPT into three way decision-making
Let \( A = \{A_1, A_2, \ldots, A_m\} \) be a set of \( m \) alternatives, \( m \geq 2 \), and \( C = \{C_1, C_2, \ldots, C_n\} \) be the set of predetermined \( n \)-criteria. Let there be \( d \) decision-makers participating in a MCGDM problem, denoted by the set \( D = \{DM_1, DM_2, \ldots, DM_d\}, \ d \geq 2 \). Each DM expresses assessment of each alternative on each criterion in a linguistic description belonging to the set \( LS = \{i_{df}, i_{d(f-1)}, \ldots, i_{d0}, \ldots, i_{a0}\} \).

Let the alternative–criteria matrix of the \( k \)th DM be \( \tilde{\mathbf{B}}^{(k)} = \begin{bmatrix} \tilde{b}_{11}^{(k)} & \ldots & \tilde{b}_{1c}^{(k)} \\ \vdots & \ddots & \vdots \\ \tilde{b}_{m1}^{(k)} & \ldots & \tilde{b}_{mc}^{(k)} \end{bmatrix} \), where \( \tilde{b}_{ij}^{(k)} = t_{ij}^{(k)}, \ a_{ij} \in \{-f, \ldots, 0, \ldots, g\}, \) and \( i = 1, \ldots, m, \ j = 1, \ldots, n, \ k = 1, \ldots, d \). The 2-tuple linguistic representation of each linguistic term in the \( d \) matrices is \( \gamma_{ij}^{(k)} \), with \( \gamma_{ij}^{(k)} = 0 \).

Without loss of generality, we assume that in each matrix \( \tilde{\mathbf{B}}^{(k)} \), the first \( \ell \)-columns corresponding to the cost criteria (lower is better) and the remaining \( n - \ell = s \) columns are beneficial criteria (higher is better).

\[
\begin{align*}
\tilde{\mathbf{B}}^{(k)} &= \begin{bmatrix} A_1 & \ldots & C_\ell & C_{\ell+1} & \ldots & C_n \\
\tilde{b}_{11}^{(k)} & \ldots & \tilde{b}_{1\ell}^{(k)} & \tilde{b}_{1\ell+1}^{(k)} & \ldots & \tilde{b}_{1n}^{(k)} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\tilde{b}_{m1}^{(k)} & \ldots & \tilde{b}_{m\ell}^{(k)} & \tilde{b}_{m\ell+1}^{(k)} & \ldots & \tilde{b}_{mn}^{(k)} 
\end{bmatrix},
\end{align*}
\]

where \( \tilde{b}_{ij}^{(k)} = (t_{a_{ij}}^{(k)}, \gamma_{ij}^{(k)} = 0), \ i = 1, \ldots, m, \ j = 1, \ldots, n, \ k = 1, \ldots, d \).

The distinction in the criteria makes DEA a natural to implement in MCGDM problem. The cost criteria data values are considered as input data, while the beneficial criteria data values are the output data in the DEA model. The \( m \) alternatives act as the \( m \) DMUs. The DEA model computes the weights \( w_{ij}^{(k)} \) of \( k \)th DM for the \( i \)th alternative on the \( j \)th criterion, \( i = 1, \ldots, m, \ j = 1, \ldots, n, \ k = 1, \ldots, d \). We shall be using the index \( p \) to denote the cost criteria and index \( q \) for the beneficial criteria; \( p = 1, \ldots, \ell; \ q = \ell + 1, \ldots, \ell + s \).

We use the output oriented DEA model by Charnes et al. (1978) for efficiency evaluation of a DMU \( o \) as follows:

\[
(M1) \quad \max \sum_{q=1}^{s} u_{iq}^{(k)} \Omega^{-1}(v_{a_{iq}}^{(k)}, 0) \\
\text{subject to} \\
\sum_{p=1}^{\ell} v_{ip}^{(k)} \Omega^{-1}(v_{a_{ip}}^{(k)}, 0) = 1 \\
\sum_{q=\ell+1}^{\ell+s} u_{iq}^{(k)} \Omega^{-1}(u_{a_{iq}}^{(k)}, 0) - \sum_{p=1}^{\ell} v_{ip}^{(k)} \Omega^{-1}(v_{a_{ip}}^{(k)}, 0) \leq 0, \\
\quad i = 1, \ldots, m, \\
u_{iq}^{(k)}, v_{ip}^{(k)} \geq 0, \\
\quad i = 1, \ldots, m, \ p = 1, \ldots, \ell, \ q = \ell + 1, \ldots, \ell + s.
\]

Model (M1) is to be solved for each DMU \( o, \ o = 1, \ldots, m, \) and for each DM \( k = 1, \ldots, d \). The optimal weights of \( \ell \) inputs \( v_{ip}^{(k)} \) and \( s \) outputs \( u_{iq}^{(k)} \) are obtained for \( DMU_o \). The optimal weight vector is \( w_{ij}^{(k)} = (u_{ip}^{(k)}, v_{iq}^{(k)}, p = 1, \ldots, \ell, \ q = \ell + 1, \ldots, \ell + s, \ i = 1, \ldots, m, \ j = \ldots, d) \).
1,..., n = \ell + s, k = 1,..., d. We normalize the optimal weight vector over the index k and obtain the aggregated matrix \( B_{agg} = [\bar{b}_{ij}]_{m \times n}, \) with

\[
\bar{b}_{ij} = \Omega \left( \sum_{k=1}^{d} w_{ik}^{(k)} \Omega^{-1} \left( f_{ij}^{(k)}, 0 \right) \right),
\]

\( i = 1, \ldots, m, \quad j = 1, \ldots, n. \) (1)

Using the worst and the best reference points in DEA, Liu et al. (2019) proposed a cross-efficiency model based on the prospect theory. The worst DMU uses the maximum inputs to yield the least outputs, while the best DMU does exactly the opposite by producing maximum outputs using the least inputs. The gain and loss of any DMU are measured relative to values above the worst DMU and below the best DMU, respectively.

In context of the MCGDM, the \( i \)th DMU input and output data values in \( B_{agg} \) matrix are, respectively, \( \bar{b}_{ip}, p = 1, \ldots, \ell, \) and \( \bar{b}_{iq}, q = \ell + 1, \ldots, \ell + s. \)

**Definition 4** If the reference point is the worst DMU, then the prospect gain value with respect to \( p \)th input and \( q \)th output of DMU \( j, i = 1, \ldots, m, \) is defined as

\[
(G_{ip}^+)_l = (\bar{b}_p - \bar{b}_{ip})^\omega, \quad p = 1, \ldots, \ell,
\]

\[
(G_{iq}^+)_o = (\bar{b}_{iq} - \bar{b}_q^\omega)^\omega, \quad q = \ell + 1, \ldots, \ell + s,
\]

where \( \bar{b}_p^\omega = \max_{1 \leq i \leq m} (\bar{b}_{ip}) \) and \( \bar{b}_q^\omega = \min_{1 \leq i \leq m} (\bar{b}_{iq}) \) are, respectively, the \( p \)th input and \( q \)th output of the worst DMU.

**Definition 5** If the reference point is the best DMU, then the prospect loss values with respect to \( p \)th input and \( q \)th output of DMU \( i, j = 1, \ldots, m, \) are defined as

\[
(L_{ip}^-)_l = -\theta (\bar{b}_{ip} - \bar{b}_p^-)^\phi, \quad p = 1, \ldots, \ell,
\]

\[
(L_{iq}^-)_o = -\theta (\bar{b}_q^+ - \bar{b}_{iq})^\phi, \quad q = \ell + 1, \ldots, \ell + s,
\]

where \( \bar{b}_p^- = \min_{1 \leq i \leq m} (\bar{b}_{ip}) \) and \( \bar{b}_q^+ = \max_{1 \leq i \leq m} (\bar{b}_{iq}) \) are, respectively, the \( p \)th input and \( q \)th output of the best DMU.

The DM always prefer to allocate weights to inputs and outputs so to maximize gain and minimize loss for DMU \( j, i = 1, \ldots, m. \) To accomplish this task, Liu et al. (2019) proposed the following cross-efficiency model for DMU \( j, i = 1, \ldots, m. \)

\[
(CE) \quad \max \eta \left( \sum_{q=\ell+1}^{\ell+s} u_{iq}(G_{iq}^+) + \sum_{p=1}^{\ell} v_{ip}(G_{ip}^+) \right) + (1 - \eta) \left( \sum_{q=\ell+1}^{\ell+s} u_{iq}(L_{iq}^-) + \sum_{p=1}^{\ell} v_{ip}(L_{ip}^-) \right)
\]

subject to

\[
\sum_{p=1}^{\ell+s} v_{ip}^\eta \bar{b}_{ip} = 1,
\]

\[
\sum_{q=\ell+1}^{\ell+s} u_{iq}^\eta \bar{b}_{iq} = 1,
\]

\[
\sum_{q=\ell+1}^{\ell+s} u_{iq}^\eta \bar{b}_{iq} - \sum_{p=1}^{\ell} v_{ip}^\eta \bar{b}_{ip} \leq 0, \quad i = 1, \ldots, m,
\]

\[
u_{iq}, v_{ip} \geq 0, \quad p = 1, \ldots, \ell, \quad q = \ell + 1, \ldots, \ell + s.
\]

Here, \( \eta \in [0, 1] \) depicts the trade-off parameter between the prospect gain and prospect loss for the decision maker.

**4 Proposed methodology**

Taking motivation from the cited works, in this section, we put forward an algorithm to solve MCGDM problem with entries coming from the unbalanced 2-tuple linguistic term set.

**Algorithm 2 Ranking the alternatives**

Step 1: Prepare a linguistic term set \( LS \) for assessing the criteria. Convert the linguistic assessment of criteria into the corresponding 2-tuple linguistic term. If the linguistic information on criteria is partly supplied in the form of \( \ell_{gb} \) with \( \beta \in [a^{-}, a^{+}] \), then construct the 2–tuple linguistic assessment using the map \( \Omega \).

Step 2: Collect the alternate–criteria matrices \( \tilde{B}^{(k)} = [\tilde{b}_{ij}^{(k)}]_{m \times n}, \) \( k = 1, \ldots, d. \) from the decision-makers with entries from the set \( LS \).

Step 3: Convert the matrix \( \tilde{B}^{(k)} \) into the corresponding unbalanced 2-tuple linguistic decision matrix \( \tilde{b}_{ij}^{(k)} = (t_{ij}^{(k)}, g_{ij}^{(k)}), \) \( k = 1, \ldots, d. \)

Step 4: Solve the model (M1) described in Section 3, and determine the optimal weights \( w_{ij}^{(k)} \) \( i = 1, \ldots, m, \quad j = 1, \ldots, n, \) with respect to \( l \)th alternative and \( j \)th criterion for the \( k \)th DM, \( k = 1, \ldots, d. \)

Step 5: Compute the aggregated matrix \( \tilde{B}_{agg} = [\tilde{b}_{ij}]_{m \times n} \) using relation (I) listed in Section 3.

Step 6: Compute the prospect gain and prospect loss values by choosing appropriate values of the parameters \( \alpha, \phi \) and \( \theta \) in Definitions 4 and 5, respectively.

Step 7: Solve the model (CE), explained in Section 3, for each alternative \( A_i, i = 1, \ldots, m, \) to get the optimal weights \( \tilde{w}_{ij}^{(k)} \), \( q = 1, \ldots, s, \) and \( \tilde{v}_{ij}^{(k)} \), \( p = 1, \ldots, \ell. \)

Step 8: Apply Algorithm 3 to compute the aggregated cross-efficiency scores \( E_{ij}^{agg}, i' = 1, \ldots, m. \)

Step 9: Rank the alternatives in the ascending order of \( E_{ij}^{agg} \), that means, the least value of \( E_{ij}^{agg} \) results is the highest rank of the alternative \( A_i. \)
Algorithm 3 Aggregated cross-efficiency scores
Step 1: Compute the cross-efficiency $E_{ii'}$ of DMU$_i$, using the optimal weights $w^*_q$ and $v^*_p$ by (CE) model, described in Section 3, for DMU$_i$.

$$
E_{ii'} = \sum_{q=1}^{t} w^*_q b_{iq} / \sum_{p=1}^{t} v^*_p b_{ip}, \quad i, i' = 1, \ldots, m.
$$
Step 2: Compute Shannon’s entropy of $E_{ii'}$ defined by

$$
e_{ii'} = - \left( \frac{E_{ii'} \sum_{i'=1}^{m} E_{ii'}}{m} \right) \ln \left( \frac{E_{ii'} \sum_{i'=1}^{m} E_{ii'}}{m} \right), \quad i, i' = 1, \ldots, m.
$$
Step 3: Compute the mean and standard deviation of entropy for DMU$_i$, $i = 1, \ldots, m$, as

$$
\bar{e}_i = \frac{1}{m-1} \sum_{i'=1}^{m-1} e_{ii'}, \quad \pi_i = \sqrt{\frac{1}{m-1} \sum_{i'=1}^{m-1} (e_{ii'} - \bar{e}_i)^2}.
$$
Step 4: Compute the variation coefficients $\delta_i$ and normalize it to find the cross-efficiency weights of alternative $A_i$ as $\tilde{\xi}_i$.

$$
\delta_i = \frac{\pi_i}{e_i}, \quad \tilde{\xi}_i = \frac{\delta_i}{\sum_{i=1}^{m} \delta_i}, \quad i = 1, \ldots, m.
$$
Step 5: The aggregated cross-efficiency of DMU$_i$ is defined by $E_{i}^{agg} = \sum_{i'=1}^{m} \tilde{\xi}_i E_{ii'}, \quad i' = 1, \ldots, m$.

5 Illustration

In this section, an example is presented to showcase the proposed MCGDM methodology. The example is set in the backdrop of medical devices selection. Medical devices are expensive and incur a high maintenance cost. How to select the best medical device is a key problem for a hospital. Suppose a hospital H is ready to order a set of sensor devices to improve the technological level. Medical experts have shortlisted eight sensor devices $\{A_i, i = 1, \ldots, 8\}$. Seven criteria are identified for assessment: maintenance support difficulty ($C_1$), purchase cost ($C_2$), degree of environmental interference ($C_3$), stability ($C_4$), sensitivity ($C_5$), linear range ($C_6$), and the degree of intelligence ($C_7$). Among these, $C_1$, $C_2$, and $C_3$ are cost criteria thus considered as inputs, while $C_4$, $C_5$, $C_6$, and $C_7$ are benefit criteria taken as outputs in DEA model ($M1$). The assessment of eight alternatives on seven criteria is carried out by four decision-makers ($D_k$, $k = 1, \ldots, 4$) from the linguistic term set having a semantic representation as follows:

$$
LS = \{t_2^{(-4)}: 	ext{extremely poor (EP)}, \ t_2^{(-3)}: \text{very poor (VP)}, \ t_2^{(-2)}: \text{poor (P)}, \ t_2^{(-1)}: \text{slightly poor (SP)}, \ t_2^0: \text{medium (M)}, \ t_2^{(1)}: \text{slightly good (SG)}, \ t_2^{(2)}: \text{good (G)}, \ t_2^{(3)}: \text{very good (VG)}, \ t_2^{(3)}: \text{extremely good (EG)}\}.
$$

Here, one can view medium as average, slightly poor as better than poor but below average, slightly good as above average but not good.

Table 1 exhibits the alternative–criteria assessment matrices by the four decision-makers. For instance, the four decision makers have rated alternative $A_1$ on criterion $C_1$ by SG, SP, M, P, respectively, forming the entry in cell (1, 1) in Table 1. Thereafter, we apply Step 1 of the Algorithm 2 to construct the 2—tuple linguistic assessment using the map $\Omega$.

Applying Step 2 of Algorithm 2, Table 2 reports the optimal weights for alternative–criteria of the decision-makers on solving model ($M1$) (defined in Sect. 3).

|    | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | $C_6$ | $C_7$ |
|----|-------|-------|-------|-------|-------|-------|-------|
| $A_1$ | SG, SP | G, SP | M, M | G, SP | SG, SP | M, M | G, G |
|      | M, P  | SG, SP | P, SG | SP, P | G     | SP   | P, SP |
| $A_2$ | SP, VP| SG, SP | G, SP | G, SG | G, M  | M, VP | SG   |
|      | VP, M | SG, M | VP, P | P, M  | VP, P | SP   | VP   |
| $A_3$ | VG, EG| VG, EG | G, VG | G, G  | EG, VG | VG, VG| G, EG |
|      | SP, M | G, VG | G, G  | G, VG | M, MG | EG   | G    |
| $A_4$ | G, SP | VG, P | SG, G | G, VP | G, VG | VG, G| SG, EG|
|      | EP, SG| VP, M | M, EP | SP, VP| P, G  | SG   | SP   |
| $A_5$ | SG, SG| G, M  | VG, M | SG, SG| M, SP | P, SP | EG, VG|
|      | M, SP | P, SG | SP, M | SP, M | G     | VP   | G    |
| $A_6$ | M, G | G, MG | EG, VG| G, VG | SG, VG| G, SG | SP   |
|      | SG, SP| G, VG | G, SP | SP, P | EG    | G    | M, EG|
| $A_7$ | SG, SG| SG, G | G, SG | G, SG | G, VG | SG, VG| M, G |
|      | G, SG | M, EG | M, VG | G, VG | SG, VG| G, SG | SP   |
| $A_8$ | SP, M | SG, MG| G, P  | M, M  | EG, SP| P, G  | P, SP |
|      | M, M  | M, P  | SP, SP| M, P  | SG, M | VP, SG| M, M |
Following Step 3 of Algorithm 2, the optimal weights are given in Table 2, and applying 1, we obtain the aggregated matrix $B_{agg}$ as follows:

$$
\begin{pmatrix}
0.651 & 2.237 & 0.775 & 0.719 & 2.23 & 1 & 4 \\
0.267 & 2 & 0.346 & 0.533 & 0.802 & 1 & 0.125 \\
0.603 & 4.377 & 5.019 & 7.504 & 3.909 & 10.854 & 5.631 \\
1.196 & 2.96 & 1.350 & 0 & 5.859 & 4 & 1 \\
1.017 & 2.556 & 1.222 & 0.965 & 4 & 4 & 3.23 \\
1.363 & 3.025 & 8 & 5.136 & 13.559 & 8.615 & 4 \\
2.563 & 2.789 & 1.955 & 3.52 & 2 & 5.143 & 4 \\
1 & 0.756 & 2.536 & 1 & 7.433 & 0 & 0.896 \\
\end{pmatrix}
$$

We take $\eta = 0.5$, $\alpha = 0.89$, $\phi = 0.92$ and $\theta = 2.25$, and solve $(CE)$ model for each of the alternatives. The values of the parameters are taken from Kahneman and Tversky (1979), while in the absence of any information on the trade-off between loss and gain, we set $\eta = 0.5$.

Table 3 exhibits the optimal weights from $(CE)$ model and Table 4 provides the cross-efficiency values $E_{ii'}$, $i, i' = 1, \ldots, 8$.

We next compute the entropy values of the cross-efficiencies, and list the same in Table 5.

Table 6 presents the cross-efficiency weights $\xi$, $i = 1, \ldots, 8$, of alternatives, the aggregated cross-efficiency scores, and ranks of alternatives.

Keeping the parameters $\alpha = 0.89$, $\phi = 0.92$ and $\theta = 2.25$, at the same levels, we vary the trade-off parameters $\eta \in (0, 1)$, to see the change it brought in the ranks of the alternatives. The results are summarized in Table 7.
Table 3 Optimal solution of (CE) model, defined in Sect. 3, for each alternatives when the parameters are $\alpha = 0.89$, $\phi = 0.92$, $\theta = 2.25$, $\eta = 0.5$

| C1  | C2  | C3  | C4  | C5  | C6  | C7  |
|-----|-----|-----|-----|-----|-----|-----|
| A1  | 1.403 | 0.112 | 0 | 0 | 0 | 0.25 |
| A2  | 0.94 | 2.166 | 0.341 | 0 | 0.818 | 0 |
| A3  | 1.659 | 0 | 0 | 0 | 0.092 | 0 |
| A4  | 0.49 | 0.307 | 0 | 0.171 | 0 | 0 |
| A5  | 0.722 | 0.218 | 0.1473 | 0.061 | 0.052 |
| A6  | 0.734 | 0 | 0 | 0.345 | 0 | 0 |
| A7  | 0.003 | 0.507 | 0.0849 | 0 | 0.136 | 0 |
| A8  | 0 | 1.322 | 0 | 0 | 0.135 | 0 |

Table 4 The cross-efficiency $E_{i,i'}$, $i, i' = 1, \ldots, 8$, for the alternatives following Step 1 of the Algorithm 3 explained in Sect. 4

| A1  | A2  | A3  | A4  | A5  | A6  | A7  | A8  |
|-----|-----|-----|-----|-----|-----|-----|-----|
| A1  | 0.076 | 1 | 0.137 | 0.516 | 0.356 | 0.262 | 0.133 |
| A2  | 0.464 | 1 | 0.81 | 1 | 0.473 | 0.814 | 0.053 |
| A3  | 0.085 | 0.208 | 1 | 0.186 | 0.218 | 0.351 | 0.111 | 0 |
| A4  | 0.684 | 0.578 | 0.363 | 1 | 0.782 | 0.741 | 0.184 | 1 |
| A5  | 0.936 | 0.691 | 0.999 | 0.999 | 1 | 0.358 | 0.895 |
| A6  | 0.344 | 0.302 | 0.652 | 0.493 | 0.395 | 1 | 0.079 | 0.747 |
| A7  | 0.493 | 0.998 | 0.827 | 0.786 | 0.999 | 0.396 | 1 | 0.066 |
| A8  | 0.101 | 0.041 | 0.091 | 0.201 | 0.16 | 0.456 | 0.073 | 1 |

Table 5 Entropy $e_{i,i'}$, $i, i' = 1, \ldots, 8$, of the cross-efficiency values of alternatives obtained on applying Step 2 of Algorithm 3 described in Sect. 4

| A1  | A2  | A3  | A4  | A5  | A6  | A7  | A8  |
|-----|-----|-----|-----|-----|-----|-----|-----|
| A1  | 0.358 | 0.083 | 0.358 | 0.127 | 0.283 | 0.233 | 0.195 | 0.125 |
| A2  | 0.206 | 0.307 | 0.307 | 0.279 | 0.307 | 0.208 | 0.280 | 0.044 |
| A3  | 0.128 | 0.225 | 0.357 | 0.211 | 0.232 | 0.295 | 0.153 | 0 |
| A4  | 0.263 | 0.241 | 0.183 | 0.314 | 0.282 | 0.274 | 0.116 | 0.314 |
| A5  | 0.271 | 0.231 | 0.280 | 0.280 | 0.280 | 0.280 | 0.154 | 0.265 |
| A6  | 0.211 | 0.195 | 0.295 | 0.258 | 0.228 | 0.346 | 0.077 | 0.313 |
| A7  | 0.215 | 0.308 | 0.283 | 0.276 | 0.308 | 0.188 | 0.308 | 0.053 |
| A8  | 0.145 | 0.076 | 0.135 | 0.223 | 0.194 | 0.330 | 0.116 | 0.355 |

Table 6 Cross efficiency weights and aggregated cross-efficiency obtained on applying Step 4 and Step 5 of Algorithm 3 described in Sect. 4; the final ranking of alternatives follows from Step 9 of Algorithm 2

| $\xi_i$ | $E_{i,i'}^{agg}$ | Rank |
|---------|-----------------|------|
| A1  | 0.157 | 0.457 | 5 |
| A2  | 0.121 | 0.421 | 7 |
| A3  | 0.1767 | 0.732 | 1 |
| A4  | 0.0894 | 0.481 | 4 |
| A5  | 0.056 | 0.558 | 2 |
| A6  | 0.114 | 0.535 | 3 |
| A7  | 0.119 | 0.337 | 8 |
| A8  | 0.166 | 0.426 | 6 |

Table 7 Ranks of alternatives for different values of $\eta$ in (0, 1) in the (CE) model

| $\eta \rightarrow$ | A1  | A2  | A3  | A4  | A5  | A6  | A7  | A8  |
|----------------|-----|-----|-----|-----|-----|-----|-----|-----|
| Alternatives ↓ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
| A1  | 1 | 2 | 1 | 3 | 5 | 2 | 3 | 1 |
| A2  | 8 | 7 | 8 | 7 | 7 | 8 | 8 | 8 |
| A3  | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |
| A4  | 7 | 8 | 7 | 7 | 4 | 8 | 7 | 7 |
| A5  | 5 | 6 | 5 | 2 | 4 | 5 | 4 | 4 |
| A6  | 3 | 2 | 3 | 2 | 3 | 2 | 2 | 3 |
| A7  | 6 | 5 | 6 | 8 | 6 | 6 | 6 | 7 |
| A8  | 4 | 4 | 4 | 6 | 5 | 5 | 4 | 5 |

The different values of the parameter $\eta$ capturing different perspectives toward gain and loss yield different rankings of alternatives that cannot be ignored. We apply Kendall’s coefficient of concordance test to test the statistical significance of rank change for the pessimistic and optimistic situations. Kendall’s coefficient of concordance is a measure of agreement among several quantitative variables that assess a set of a finite number of objects of interest. The test statistics $W$ is given by

$$W = \frac{12 \sum_{i=1}^{m} (R_i - \bar{R})^2}{\tau^2 (m^3 - m)} ,$$

where $R_i$ is the sum of ranks for the $i$th object judged by $\tau$ assessing variables, $\bar{R}$ is the mean of all $R_i$ values, $\tau$ is the number of judgments on $m$ objects. The random variable $\chi^2 = \tau (m - 1) W$ is asymptotically chi-square distributed with $m - 1$ degrees of freedom.

We first consider the pessimistic case when $\eta = 0.1, 0.2, 0.3, 0.4, 0.5, \text{so } \tau = 5$ (judges) and $m = 8$ alternatives ($A_1, \ldots, A_8$). The null hypothesis and the alternative hypothesis are set as follows:

$H^{0}_{\text{pess}}$: there is no concordance in rankings of alternatives in pessimistic case;

$H^{0}_{\text{pess}}$: there is a concordance in rankings of alternatives in pessimistic case,

The test statistics $W = 0.9714$, and $\chi^2 = 34$ with $p$ value $1.76 \times 10^{-4}$. The null hypothesis gets rejected. On the similar lines, in the optimistic case when $\eta = 0.5, 0.6, 0.7, 0.8, 0.9$, the test statistics $W = 0.958$, and $\chi^2 = 33.53$ with the $p$ value $7.8 \times 10^{-5}$. The null hypothesis of no concordance in rankings of alternatives in optimistic case also gets rejected. In both the cases $W$ value is close to 1, indicating the rankings are more toward in agreement.
Table 8 Transformation of the linguistic variables to TFN

| Linguistic variable | TFN            |
|---------------------|---------------|
| Very poor (VP)      | (0.9, 1, 1)   |
| Poor (P)            | (0.7, 0.9, 1) |
| Medium poor (MP)    | (0.5, 0.7, 0.9) |
| Medium (M)          | (0.3, 0.5, 0.7) |
| Medium good (MG)    | (0.1, 0.3, 0.5) |
| Good (G)            | (0.01, 0.1, 0.3) |
| Very good (VG)      | (0.01, 0.01, 0.1) |

5.1 Comparative analysis

In addition, in the same example, we make a comparison of our method with the existing fuzzy ordinal priority approach (OPA-F) (Mahmoudi et al. 2021).

After converting cost criterion C₁, C₂, C₃ into the benefit criterion, the linguistic representation of Bagg, using the linguistic semantics in LS, is as follows.

\[
\begin{align*}
\text{MG} & \quad \text{MG} & \quad \text{VG} & \quad \text{MP} & \quad \text{MG} & \quad \text{M} & \quad \text{G} \\
\text{MG} & \quad \text{MG} & \quad \text{VG} & \quad \text{MP} & \quad \text{M} & \quad \text{VP} \\
\text{MG} & \quad \text{MP} & \quad \text{MG} & \quad \text{VG} & \quad \text{G} & \quad \text{VG} & \quad \text{G} \\
\text{M} & \quad \text{M} & \quad \text{VG} & \quad \text{VP} & \quad \text{G} & \quad \text{G} & \quad \text{M} \\
\text{MG} & \quad \text{MG} & \quad \text{VG} & \quad \text{M} & \quad \text{G} & \quad \text{G} \\
\text{M} & \quad \text{M} & \quad \text{MP} & \quad \text{G} & \quad \text{VG} & \quad \text{VG} & \quad \text{G} \\
\text{MP} & \quad \text{MG} & \quad \text{VG} & \quad \text{G} & \quad \text{MG} & \quad \text{G} \\
\text{M} & \quad \text{G} & \quad \text{G} & \quad \text{M} & \quad \text{VG} & \quad \text{M} & \quad \text{M}
\end{align*}
\]

Mahmoudi et al. (2021) described the representation of each linguistic term into a triangular fuzzy number (TFN) using the fuzzy scale presented in Table 8. Each of the entries of Bagg are converted into TFN applying this scale.

Following the steps of the OPA-F algorithm by Mahmoudi et al. (2021) and explained toward end of Sect. 3, we formulated the linear programming attributes model of Mahmoudi et al. (2021) to obtain the criteria weights. This model involves 38 constraints and 45 variables and the obtained triangular fuzzy weights of criteria are exhibited in Table 9.

We then solve the alternatives model of the OPA-F algorithm and compute the scores of the alternatives. The problem model, when written for the data in the Example 1, involves 300 constraints and 339 variables. The triangular fuzzy weights of the alternatives and the ranking function of their total fuzzy scores are depicted in Table 10, where the total fuzzy score of i-th alternative is \( TS_i = (L, M, U) = \sum_{j=1}^{3} (w_i^L, w_i^M, w_i^U), (\tilde{u}_i^L, \tilde{u}_i^M, \tilde{u}_i^U) \), with \( w_i^L \) and \( w_i^U \) representing j-th criterion and i-th alternative triangular weights, respectively, and \( R(TS) = \frac{L+M+U}{6} \).

The values \( R(TS) \) yield the ranking of alternatives as follows.

\[ A_3 > A_1 > A_7 > A_6 > A_5 > A_8 > A_2 > A_4. \]

We observe that the top two ranks and the bottom two ranks of alternatives by our proposed approach with \( \eta = 0.5 \) and the OPA-F method are the same. However, it is the ranks in the middle where one notice a difference. While we obtain A₈ to be mainly at fourth rank, the OPA-F ranking placed it at the sixth spot; A₇ got a lower rank of fifth or sixth place in our method, but it got the third rank by the OPA-F method.

We tried to see if the two rankings statistically differ. We compared the rankings of our method when \( \eta = 0.5 \) with that of the OPA-F method. Set the null hypothesis, \( H^0 \) : there is no concordance in rankings vs the alternate hypothesis and the alternative hypothesis \( H^1 \) : there is a concordance in
Table 12 Fuzzy criteria weights in F-TOPSIS

| Criteria | Triangular weight |
|----------|-------------------|
| C₁       | (0.1, 0.575, 1)   |
| C₂       | (0.01, 0.275, 0.7) |
| C₃       | (0.01, 0.339, 1)  |
| C₄       | (0.01, 0.451, 1)  |
| C₅       | (0.01, 0.178, 0.7) |
| C₆       | (0.01, 0.228, 0.7) |
| C₇       | (0.01, 0.313, 1)  |

rankings. Here, \( \tau = 2, m = 8 \), and the test statistics \( W = 0.6486, \chi^{2} = 9 \), and the \( p \) value equals 0.2527, indicating we fail to reject \( H^0 \) at almost 25.27% of significance level. In other words, the statistical test agrees with our noting of different ranking of alternatives by our proposed and OPA-F methods. The \( W \) value close to 0.65 also indicates that the two rankings are partly in unanimity but this unanimity is not very strong.

Though it is hard to explain one reason for the ranking difference, to point out the major ones, unlike the OPA-F algorithm where the usual linear program computes the criteria weights, our method uses prospect theory-based value function to capture the perspective difference on the cost and benefit criteria. We also used the cross-efficiency scores that consider the other alternatives’ efficiency while calculating the self-efficiency of an alternative. Another limitation is that the OPA-F algorithm uses fuzzy equivalence scales to capture the semantics of the linguistic term. This scale is subjective and cumbersome to use when the cardinality of the linguistic term set increases, and its granularity becomes finer with complex interpretation.

In order to expand the scope, we include the fuzzy TOPSIS (we shall be calling it by F-TOPSIS) in our comparative analysis. On applying the linguistic semantics in \( LS \), Table 12 depicts the triangular fuzzy weights of criteria.

Following the procedure for F-TOPSIS of Nădăban et al. (2016) on the aggregated matrix \( B_{agg} \), we obtain the alternatives rank \( A_2 > A_1 > A_4 > A_8 > A_5 > A_7 > A_6 > A_3 \).

Comparing the F-TOPSIS ranking with that of ours for \( \eta = 0.5 \), the two rankings appeared to be quite different. For instance, \( A_2 \) is ranked seventh by our proposed approach while F-TOPSIS ranks it at the top. Also, \( A_3 \) gets the top rank by our approach but ranked the lowest by F-TOPSIS. The difference in ranking of \( A_4 \) and \( A_6 \) is also quite evident. We test our observation statistically by setting the null hypothesis: \( H^0 \) : there is no concordance in rankings of alternatives by two approaches (ours one with \( \eta = 0.5 \) and F-TOPSIS); against the alternative hypothesis: \( H^1 \) : there is a concordance in rankings of alternatives by the two approaches.

Here, \( \tau = 2, m = 8 \), and the test statistics \( W = 0.238 \), and \( \chi^{2} = 3.33 \) with very high \( p \) value equals 0.853. We fail to reject the null hypothesis at a very high significance level of almost 85.3%. It is in line with what we have observed. The \( W \)-value is far away from one and more close to zero, clearly signaling that the two rankings differ more or less, and any agreement between them is fairly weak.

The rankings by F-TOPSIS are sensitive to the choice of the pre-defined criteria weights. The same is summarized in Table 15. We have picked four weight vectors to solve the F-TOPSIS method, depicted in Table 14 and obtained the rankings of alternatives. The first weight vector \( W^1 \) is derived using column-wise minimum of the first component, average of the middle component and maximum of the third component of the TFN from the equivalent representation of matrix \( B_{agg} \) with triangular fuzzy entries applying the scale given in Table 12. The weight vector \( W^2 \) is the one obtained from OPA-F linear programming problem, while the other two weight vectors \( W^3 \) and \( W^4 \) are randomly chosen.

From Table 15, we observe that altering the weight vector causes change in rankings of alternatives by F-TOPSIS method.

In Table 16, we report the correlation of rankings between our method and F-TOPSIS or that between OPA-F and F-TOPSIS, for different choices of the pre-defined weight vectors on seven criteria in the F-TOPSIS. We observed that result in both the positive and negative correlation values.

The summary of our observations on comparative analysis for Example 1 is as follows:

(i) One can obtain different rankings of alternatives by different MCGDM methodologies.
(ii) Ranking of the alternatives by F-TOPSIS is sensitive to the choice of the fuzzy weight vectors of the criteria, which are assumed to be pre-specified.
(iii) The correlation coefficient in different rankings between our method, OPA-F, and F-TOPSIS, could take both negative and positive values.
(iv) Our method involves the four parameters \( \alpha, \phi, \theta \) in the value function of CPT and the fourth parameter

Table 13 Rank of alternatives by our method with \( \eta = 0.5 \), OPA-F method (Mahmoudi et al. 2021) and F-TOPSIS (Nădăban et al. 2016) with criteria weights in Table 12

| Alternatives | Our method | OPA-F | F-TOPSIS |
|--------------|------------|-------|----------|
| A₁           | 2          | 2     | 2        |
| A₂           | 7          | 7     | 1        |
| A₃           | 1          | 1     | 8        |
| A₄           | 8          | 8     | 3        |
| A₅           | 6          | 5     | 5        |
| A₆           | 3          | 4     | 7        |
| A₇           | 5          | 3     | 6        |
| A₈           | 4          | 6     | 4        |
Table 14 Four weights vectors selected to conduct the F-TOPSIS analysis on the aggregated the linguistic representation of $B_{agg}$ matrix

| Criteria | Fuzzy triangular weights $W^1$ | Fuzzy triangular weights $W^2$ |
|----------|-------------------------------|-------------------------------|
| C1       | (0.1, 0.575, 1)               | (0.345, 0.716, 0.774)         |
| C2       | (0.01, 0.275, 0.7)            | (0.446, 0.585, 0.616)         |
| C3       | (0.01, 0.339, 1)              | (0.258, 0.296, 0.381)         |
| C4       | (0.01, 0.451, 1)              | (0.288, 0.328, 0.344)         |
| C5       | (0.01, 0.178, 0.7)            | (0.188, 0.241, 0.241)         |
| C6       | (0.01, 0.228, 0.7)            | (0.115, 0.256, 0.682)         |
| C7       | (0.01, 0.313, 1)              | (0.127, 0.277, 0.414)         |

Table 15 Rank of alternatives with our proposed method with $\eta = 0.5$, OPA-F, and F-TOPSIS with four different criteria weight vectors $W^a$, $a = 1, \ldots, 4$

| Alternatives | Our method | OPA-F | F-TOPSIS $W^1$ | F-TOPSIS $W^2$ | F-TOPSIS $W^3$ | F-TOPSIS $W^4$ |
|--------------|------------|-------|---------------|---------------|---------------|---------------|
| A1           | 5          | 2     | 2             | 4             | 1             |
| A2           | 7          | 1     | 1             | 1             | 4             |
| A3           | 1          | 6     | 7             | 8             | 5             |
| A4           | 4          | 8     | 4             | 5             | 3             |
| A5           | 2          | 5     | 8             | 7             | 8             |
| A6           | 3          | 4     | 3             | 4             | 6             |
| A7           | 8          | 3     | 7             | 6             | 2             |
| A8           | 6          | 6     | 5             | 3             | 5             |

Table 16 The correlation between different rankings of the alternatives

| Approach       | F-TOPSIS $W^1$ | F-TOPSIS $W^2$ | F-TOPSIS $W^3$ | F-TOPSIS $W^4$ |
|----------------|---------------|---------------|---------------|---------------|
| OPA-F          | 0.2143        | 0.9284        | -0.1191       | -0.1429       |
| Our method $\eta = 0.5$ | -0.4286 | 0.523 | -0.1429 | 0.3095 |
| Our method $\eta = 0.5$, OPA-F | 0.4524 |

$\eta$ in the $(CE)$ model, Sect. 3. The values of the first three parameters are set following several research articles; we analyze the effect of change in $\eta$ values on the ranking of alternatives. The OPA-F and F-TOPSIS are free from such subjectivity of setting parameters. However, these methods have their share of concerns on the reverse conversion scale of fuzzy numbers to linguistic variables and sensitivity in ranking due to pre-selected weights vector.

6 Case study

In this section, we present a decision-making problem requiring to evaluate the overall teaching performance of faculties in the respective courses they teach in a semester based on the feedback received from the students on the end-semester questionnaire for the course.

A class of students is asked to fill the end-semester questionnaire and provide feedback on five core courses taught to them by five different faculties in the semester on the following nine criterion. The first two criteria are related to examination and evaluation policy in the course while the remaining seven criteria pertain to the course organization, delivery, and learning in the course.

C1 The exams are lengthy and involve concepts not covered in the lectures
C2 The evaluation is not fair and not effective to understand mistakes and improve on basics
C3 Course plan and lectures met the course objectives
C4 Effective in response and addressing queries
C5 Addresses and answers to queries
C6 Effective and clear communication and concepts explanation
C7 Inspire to pursue the subject area further
C8 Encourage independent and logical thinking
C9 Punctuality in holding lectures and exams

Among these, $C_1$ and $C_2$ are cost criteria thus considered as inputs while $C_3$ to $C_9$ are beneficial criteria taken as outputs in DEA model $(M1)$. The assessment of five alternatives (faculty) on nine criteria is carried out by decision-makers (students) from the linguistic term sets $L_{S1}$ (for criteria $C_1$ and $C_2$) and $L_{S2}$ (for criteria $C_3$ to $C_9$) having semantic representations as follows:

- $L_{S1} = \{t_{2(-2)} : \text{Strongly disagree (SD)}, t_{2(-1)} : \text{Disagree (D)}, t_{2(0)} : \text{Undecided (U)}, t_{2(1)} : \text{Agree (A)}, t_{2(2)} : \text{Strongly agree (SA)}\}$
- $L_{S2} = \{t_{2(-2)} : \text{Highly Ineffective (HI)}, t_{2(-1)} : \text{Average (A)}, t_{2(0)} : \text{Good (G)}, t_{2(1)} : \text{Highly effective (HE)}, t_{2(2)} : \text{Excellent (E)}\}$

The class size participated in providing feedback by filling questionnaire was more than 30 but after scrutinizing the responses, the incomplete forms where the responding students left over the questions unanswered, are discarded. We are then left with 30 acceptable responses which we call as
30 decision makers \(D_k, k = 1, \ldots, 30\). For the sake of brevity, we present only one assessment of \(D_1\) in Table 17.

The optimal weights from model (M1) are listed in Table 18.

Applying Step 3 of Algorithm 2, the aggregated matrix \(B_{agg}\) obtained is as follows:

\[
\begin{pmatrix}
0.571 & 0.598 & 1.404 & 1.99 & 1.23 & 1.24 & 1.223 & 1.22 & 1.594 \\
0.6709 & 0.682 & 1.454 & 1.322 & 1.819 & 0.873 & 1.186 & 1.44 & 1.217 \\
0.684 & 0.718 & 1.339 & 0.646 & 1.381 & 2.168 & 1.226 & 0.652 & 1.841 \\
0.512 & 0.6 & 1.00 & 2.24 & 1.99 & 1.195 & 1.4304 & 0.692 & 1.72 \\
0.596 & 0.413 & 0.684 & 2.199 & 0.9675 & 2.058 & 1.049 & 1.861 & 1.798
\end{pmatrix}
\]

We set \(\alpha = 0.89\), \(\phi = 0.92\) and \(\theta = 2.25\), in the prospect value function and apply Algorithm 3, by varying the trade-off parameters \(\eta \in (0, 1)\), to see the change it brought in the ranks of the alternatives. The results are summarized in Table 19.

We consider the pessimistic case, \(\eta = 0.1, 0.2, 0.3, 0.4, 0.5\), so, \(\tau = 5\), and \(m = 5\) alternatives \((A_1, \ldots, A_5)\), to test the null hypothesis against the alternative hypothesis: \(H^0\): there is no concordance in rankings of alternatives in pessimistic case; \(H^1\): there is a concordance in rankings of alternatives in pessimistic case.

The test statistics is \(W = 0.896\), and \(\chi^2 = 17.92\) with \(p\) value equals 0.000195. The null hypothesis gets rejected. Analogously, in the optimistic case when \(\eta = 0.5, 0.6, 0.7, 0.8, 0.9\), the test statistics turns out to be \(W = 1.957\), and \(\chi^2 = 68.5\) with the \(p\) value equals 1.2\(^{-8}\). The null hypothesis gets rejected.

In conclusion, one can say that, in general, faculty \(A_5\) is top ranked, while faculties \(A_2\) and \(A_3\) get the lower rank among the five faculties.

### 6.1 Comparative analysis

After converting the first two cost criteria into the benefit ones, we applied OPA-F procedure on the \(3 \times 9\) aggregated matrix. Table 20 provides the fuzzy scores, and subsequently the alternatives are ranked as

\(A_5 > A_2 > A_1 > A_3 > A_4\).

We tested the hypothesis on concordance between rankings by OPA-F and ours approach with \(\eta = 0.5\). We obtain \(W = 0.65\), and \(\chi^2 = 5.2\) with \(p\) value equals 0.2674. We fail to reject the null hypothesis of ‘no concordance’ on ranking by the two approaches at almost 26.74% level of significance.

We next examine the ranking with F-TOPSIS on \(B_{agg}\). Table 21 lists the criteria weights.
The ranking is $A_4 > A_5 > A_1 > A_3 > A_2$. The test statistics is $W = 0.9$, with $\chi^2 = 7.2$ and the $p$ value 0.1257. We fail to reject the null hypothesis on ‘no concordance’ between rankings by the F-TOPSIS and our approach with $\eta = 0.5$.

The example and the case study indicate that the ranking of alternatives by our approach and OPA-F and F-TOPSIS is not a concordance. More experiments are needed to conduct a thorough comparative analysis of decision-making problems involving many decision makers, evaluating a large number of alternatives on various numbers of cost–benefit criteria. All these parameters can be varied, and we can simulate assessment matrices for the decision makers from the linguistic term set. A large-scale experimental investigation can reveal the benefits and drawbacks of our suggested technique compared to OPA-F and F-TOPSIS.

7 Conclusions

The highlighting contributions of our paper are four-folds. Firstly, we have used the 2-tuple unbalanced linguistic term set that provides a broader and more realistic domain set for DMs in supplying the alternative–criteria judgment matrices. Unlike several other studies on MCGDM problems demanding precise numerical quantification, the linguistic terms embrace decision-makers’ multi-granularity and cognitive process in comprehending the available alternatives on different criteria.

Secondly, we present the DEA optimization model to compute the weights of each DM on each alternative–criteria pair in the group decision-making. The DEA model incorporates complete information from the input criteria (or cost criteria) and output criteria (or beneficial criteria) from the decision matrices to designate weights that enable the computation of one aggregated decision matrix.

Thirdly, we determine the positive ideal (the best) and negative ideal (the worst) alternatives and apply the prospect theory to determine the alternatives’ gain and loss values concerning the two ideal options. The value function allows for segregating the gain and loss domains based on the risk perspective, where guaranteed small gains are preferred over uncertain large gains and large uncertain losses get preferred over certain small losses.

Finally, we put forward a cross-efficiency DEA optimization model and aggregation of cross-efficiency scores to rank the alternatives.

Thus, allowing unbalanced 2-tuple linguistic term set in decision-making, usage of DEA in computing weights of DMs on each alternative–criteria and the prospect value theory function for determining criteria weights in different loss–gain domains to compute cross-efficiency of alternatives offers a comprehensive package to rank the alternatives.

We present a numerical example to illustrate the proposed methodology. The rankings obtained on varying the parameter in the cross-efficiency model get statistically tested for concordance or agreement. The null hypothesis of no concordance is rejected at a very high confidence level, thereby showing the efficacy and robustness of the proposed methodology in ranking the alternatives both in the pessimistic and optimistic situations.

For future works, a promising research avenue could be to use personalized individual semantics to represent the distinct understanding of each linguistic term by individual decision-maker. We may explore extending the proposed method to deal with more complex linguistic expressions in decision-making applying multi-granular hesitant fuzzy linguistic term sets and incomplete criteria weight information. Some work in this direction is already reported (see, Ren et al. 2018; Zhang et al. 2021a) yet more needs to be done specifically to handle large scale MCGDM problems that often arise in data science and explainable artificial intelligence. We can also delve into applications of other utility functions of cumulative prospect theory in MCGDM. In addition, we can employ different cross-efficiency DEA models, including satisfaction and consensus information for ranking alternatives.

Acknowledgements The authors express their gratitude to the editor and the anonymous reviewers for their critical observations and valuable suggestions. The authors are grateful to the editor-in-chief for providing us opportunity to improve upon our manuscript.

Author Contributions IK performed writing original draft, developing methodology, carrying out formal analysis, and empirical evaluations. AG contributed to research and methodology ideas, and review of mathematical and analysis parts and editing. AM performed supervision, writing, empirical evaluations and analysis, and review and editing.

Funding The authors did not receive any funds, grants, or other support.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Ethical approval The research does not contain any studies performed with human participants or animals by any of the authors.

Informed consent The research does not involve any human participants and animals.

References

Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2(6):429–444
Dong Q, Cooper O (2016) A peer-to-peer dynamic adaptive consensus reaching model for the group AHP decision making. Eur J Oper Res 250(2):521–530
large-scale group decision making: A linguistic distribution-based approach. IEEE Trans Fuzzy Syst 28:2875–2889

Zhang Z, Gao J, Gao Y, Yu W (2021) Two-sided matching decision making with multi-granular hesitant fuzzy linguistic term sets and incomplete criteria weight information. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2020.114311

Zhang Z, Li Z, Gao Y (2021b) Consensus reaching for group decision making with multi-granular unbalanced linguistic information: a bounded confidence and minimum adjustment-based approach. Inf Fusion 74:96–110

Zhao M, Wei G, Wei C, Wu J (2021) Improved TODIM method for intuitionistic fuzzy MAGDM based on cumulative prospect theory and its application on stock investment selection. Int J Mach Learn Cybern 12(3):891–901

Zou L, Pei Z, Karimi HR, Shi P (2012) The unbalanced linguistic aggregation operator in group decision making. Proc IEEE Int Sympos Intell Control 250(2):703–707

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.