Interval Reference Point Technique for Sustainable Industrial Process Selection under Uncertainties

Jingzheng Ren[a*], Sara Toniolo[b]

[a]Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China

[b]University of Padova, Department of Industrial Engineering, Via Marzolo 9, 35131, Padova, Italy

Email: jzhren@polyu.edu.hk (JZ Ren)

Abstract: Sustainability of industrial processes has been a hot spot and draws the attentions of more and more stakeholders, because determining the synthetic sustainability index can help the decision-makers/stakeholders to make informed decisions towards sustainable industrial processes, especially for selecting the most sustainable industrial process among multiple alternatives. This study aims to develop an interval reference point technique for the prioritization of industrial processes under uncertainties. The hierarchy best-worst method which can determine the weights of the criteria, the local weights of the sub-criteria and the global weights of the sub-criteria simultaneously was employed for weights determination. An interval reference point technique which allows the users to set the aspiration point (the level of acceptance) and the reservation point (the level of desirable expectation) was developed to determine the dual synthetic sustainability indexes, and both the weak synthetic sustainability index (compensatory sustainability index) and the strong synthetic sustainability index (non- compensatory sustainability index) were used to prioritize the industrial processes. A case study with five alternative energy storage technologies has been studied to illustrate the developed framework, and the results reveal that the developed framework is feasible and efficient for sustainability assessment and prioritization of industrial processes.

Keywords: Sustainability; sustainability assessment; synthetic sustainability index; reference point technique; multi-criteria decision making

1. Introduction

Industrial processes which can transform the raw materials and energy into the desirable products or services for people play a significant important role in sustaining the high quality of life, the economic growth and the modernization of the world. The decision-makers/stakeholders are aware of the importance of investigating the sustainability of industrial processes for selecting the most sustainable process with effective tools and techniques (Yi *et al*., 2014). The modern concept of sustainability usually consists of three building blocks including economy, environment and society, and the indicators/criteria in economic, environmental, and social pillars are usually used for sustainability assessment (Olinto, 2014). These are the triple bottom-line (TBL) elements. Traditionally, corporations focused on only one bottom line – the economic profits. However, corporations should emphasize on not only the economic profits, but also the influences of their actions on all the stakeholders (e.g., the customers, the investors and the suppliers) for pursuing the development in a sustainable approach, and the influences include both the social acceptability and the environmental impacts (Lombardi Netto et al., 2020). Accordingly, the indicators/criteria in economic, environmental, and social dimensions are usually used for sustainability assessment or sustainability measurement. However, the indicators/criteria in these three main dimensions are usually influenced by (or can influence) the indicators/criteria in other dimensions such as technological, political, and cultural dimensions. Therefore, sustainability assessment requires a set of criteria/indicators in multiple dimensions including not only the economic, environmental, and social dimensions, but also other dimensions (e.g., the technological, political, and cultural dimensions).

After sustainability assessment of industrial processes and obtaining the data of the industrial processes with respect to the indicators/criteria in different dimensions, a question arises: how to determine the sustainability sequence or rank these alternative industrial processes (sustainability-oriented decision-making)? Sustainability-oriented decision-making based on the considerations of multiple criteria in multiple dimensions for the prioritization of alternative industrial systems is a typical multi-criteria decision analysis problem, and the selection of the indicators for sustainability assessment and the determination of the weights of these indicators are critical for making informed decisions, and they will have significantly important effects on the final result (Bell and Morse, 2008). In addition, Munda (2005) pointed out that the improvement of the alternatives with respect to some criteria/indicators would lead to the deterioration with respect to other criteria/indicators, and it is of vital importance to answer the question – "how could the criteria/indicators for sustainability assessment be aggregated into a unique sustainability index to represent the overall sustainability?" (Munda, 2005). In this context, the synthetic sustainability index which can aggregate the criteria/indicators in multiple dimensions into an aggregated sustainability index and represent the overall sustainability of each alternative has been increasingly recognized as a useful tool for policy making and public communication by providing information on sustainability performance (Ruiz et al., 2020). With the synthetic sustainability index after aggregation, the decision-makers/stakeholders can make informed decision towards sustainability industrial process selection.

Besides the introduction section, the remaining parts of this paper has been organized as follows: the literature review was firstly carried out in section 2; subsequently, the methods including the hierarchy best-worst method and the interval reference point technique were presented in section 3; then, the case study with five energy storage technologies, the results and sensitivity analysis were presented in section 4; and finally, this study has been concluded in section 5.

2. Literature review

There are various methods for constructing the synthetic sustainability indices (also called composite or aggregated indices), and they were summarized in Table 1. Multi-criteria decision making or multi-criteria decision analysis methods (including AHP, TOPSIS, ELECTRE, DEA and SAW, etc) are the most commonly used methods. Besides these multi-criteria decision making (MCDM) methods, principal component analysis (PCA), factor analysis, benchmarking, statistical additive aggregation, goal programming, binary extended goal programming, distance-principal component and reference point technique were also employed in many studies. Normalization, weighting, and aggregation are the basic procedures for aggregating multiple indicators into an integrated index. MCDM methods are usually combined with life cycle tools for determining the aggregated/composite sustainability index (Soltani *et al*., 2016). For instance, life cycle sustainability assessment (LCSA) which usually consists of life cycle costing (LCC), life cycle assessment (LCA) and social life cycle assessment (SLCA) was usually employed to determine the data of the alternative industrial processes with respect to the criteria in economic, environmental and social dimension, respectively. After determining the decision-making matrix with the data of the alternative industrial processes with respect to the criteria in LCSA, MCDM methods can be used to determine the aggregated/composite sustainability index of each alternative.

These methods for determining the composite or synthetic sustainability indexes are usually based on the aggregation of the normalized data with respect to multiple indicators in the decision-making matrix into an aggregated index, and the aggregated sustainability index determined by most of these methods are the so-called "weak synthetic index" in which the bad environmental sustainability can be compensated by good sustainability performances in economic, social or other dimensions. However, the users of these methods can only determine the aggregated sustainability index of each alternative, but they do not know how to take appropriate actions and draft effective policies for sustainability improvement. Differently, the reference point technique employs both the weak synthetic sustainability index and the strong synthetic sustainability index to measure the sustainability of different alternatives, and the strong synthetic sustainability index is a measure of sustainability without the compensation of bad environmental performances by economic or social sustainability, and it can measure the worst status (sustainability performance) of the alternatives with respect to these indicators (Cabello *et al*., 2014). In addition, the reference point technique allows the users to set double reference points for each sustainability indicator as the aspiration level which can satisfy the desirable expectation of the decision-makers/stakeholders on each indicator and the reservation level which is the threshold value of each indicator for making the decision-makers/stakeholders feel acceptable (Ruiz *et al*., 2018). However, the applications of the reference point technique in the previous studies lack the integration of any smart weighting methods for the users to determine the weights of the indicators accurately. In addition, these methods cannot address uncertainties when the data of the alternatives with respect to some indicators are interval numbers or fuzzy numbers rather than the crisp numbers.

Table 1: The summary of the composite, synthetic and aggregated sustainability indexes

	Name of index	Method	Filed of application	References
1	Synthetic index	Analytic Hierarchy Process (AHP)	Urban ecosystem	Zhang et al., 2006
	Composite sustainability performance index		Steel industry	Singh *et al*. 2007
	Composite indicator		Waste water treatment	Molinos-Senante et al., 2014
	Synthetic index		Water supply	Molinos-Senante et al., 2019

	Synthetic competitiveness index		Geothermal resources	Kurek et al., 2020
2	Overall sustainability performance	Principal Component Analysis (PCA)	Manufacturing companies	Li et al., 2012
	Aggregated SD (sustainable development) index		EU 27	Bolcárová and Kološta, 2015
	Energy technology sustainability index		rural electrification	Mainali and Silveira, 2015
	Syntetic sustainability index		Tourism destinations	Lorenzo Linares et al., 2019
3	Composite indicator of sustainability	Data envelopment analysis (DEA) and multicriteria decision making	Farm	Rei‑Martínez *et al.*, 2011
	Synthetic index		Waste water treatment	Gomez et al., 2017
	Eco-(in) Efficiency Index		Agricultural	Vlontzos et al., 2017
4	Complex Performance Indicator	Benchmarking	Corporate	Dočekalová and Kocmanová, 2016
	Composite sustainability index		Farming systems	Stylianou et al., 2020
	Performance indicator		Corporate	Tasdemir et al., 2020
5	Synthetic sustainability index	Simple additive weighting (SAW)	Low impact infrastructure development	Maiolo *et al.*, 2017
	Composite sustainability index		Real estate projects	Dobrovolskienė *et al.*, 2019
	Sustainability index		City sustainability	Yi et al., 2018
6	Composite indicators	PCA, AHP	Agricultural systems	Gómez-Limón and Riesgo, 2009
	Composite index		Territorial growth	Lee and Chou, 2018
7	Composite indicator	PCA, DEA	Tourism destinations.	Perez *et al.*, 2013
	Composite sustainability index		Agriculture	Dong. et al., 2016
8	Composite environmental impact index	Statistical aggregation	Agriculture	Sabiha *et al.*, 2016
	Synthetic index		Corporate	Escrig-Olmedo et al., 2017
	National Corporate Social Responsibility Practices Index		Corporate	Amor-Esteban et al., 2019

9	Goal programming synthetic indicators	Goal programming technicques	Tourism	Blancas et al., 2010
	Vectorial Dynamic Composite Indicator (VDCI)		Tourism	Blancas et al., 2018
	Differential Dynamic Index		Tourism	Lozano-Oyola et al., (2019)
	Synthetic indicator	Distance-principal component and goal programming synthetic indicator	Waste water companies	Molinos-Senante et al., 2016
	Regional composite indicator	Binary extended goal programming	Agriculture	Xavier *et al.*, 2018
	Life cycle aggregated sustainability index	Interval preference relation based goal programming model and projection-based aggregated sustainability index	Electricity	Ren, 2018
10	Composite transport sustainability index	PCA, factor analysis	Transport	Reisi *et al.*, 2014
	Transportation sustainability index		Transport	Mandinia et al., 2018
11	Aggregated sustainability indices	ELECTRE I and TOPSIS	Residential modular buildings	Kamali *et al.*, 2018
	Sustainability indices		Urban areas	Zinatizadeh et al., 2017
12	Composite sustainability index	Reference point tecnique	Industry	Zhou *et al.*, 2012
	Multidimensional index		Automotive	Sikdar et al., 2012
	Synthetic sustainability indicator		Territorial unit	Ruiz *et al.*, 2011

Based on the above-mentioned literature reviews, it is apparent that there are two main research gaps:

(1) Some of the studies in literature for determining the aggregated sustainability indexes based on different MCDM methods (such as AHP, TOPSIS, ELECTRE, DEA and SAW, PCA, factor analysis, benchmarking, statistical additive aggregation, goal

programming, binary extended goal programming and distance-principal component) cannot determine the weak synthetic sustainability index and the strong synthetic sustainability index simultaneously. Although some of these methods can address uncertainties in decision-making process but they can only be used to determine the weak synthetic sustainability index in which the bad environmental performances can be compensated by other dimensions, without considering the strong synthetic sustainability index; and

(2) Some of the studies in literature can determine both the weak synthetic sustainability index and the strong synthetic sustainability index. However, these methods cannot address uncertainty in decision-making process.

In order to fill in the above-mentioned two research gaps, this study aims to develop an integrated multi-criteria decision making method for determining the weak synthetic sustainability performance and the strong synthetic sustainability performance under data uncertainties. In other words, this study aims to propose a new approach with the intent of determining both the weak synthetic sustainability index (compensatory sustainability index) and the strong synthetic sustainability index (non-compensatory sustainability index) under data uncertainties for the decision-makers to make informed decisions on the selection of alternative industrial systems. Hierarchy best-worst method which can help the users to determine the weights of the indicators accurately has been employed for weights determination, and the interval reference point technique has been developed to address data uncertainties by extending the traditional point technique to interval conditions.

3. Methods

Sustainability is a multi-dimensional concept which usually consists of multiple objectives/criteria. We can optimize the sustainability by n dimensional objectives/criteria, see the following for the multi-objective optimization problem (Ruzi *et al.*, 2011):

$$\begin{aligned} &\textit{Maximize} \quad f(x) = \left[f_1(x), f_2(x), \cdots, f_n(x) \right] \\ &\qquad\qquad\quad x \in X \end{aligned} \tag{1}$$

These n objective functions in different dimensions of sustainability usually conflict. We can assume that the stakeholders/decision-makers set a desirable value for each objective as the goal and the reference point, denoted by $g = [g_1, g_2, \cdots, g_n]$. Then, the reference point technique (Wierzbicki, 1980; Ruzi *et al.*, 2011) can transform problem (1) into:

$$\textit{Maximize} \quad s\left(f(x), \mu, g\right) = \min_{i=1,2,\cdots,n} \left\{ \mu_i \left[f_i(x) - g_i \right] \right\} + \rho \sum_{i=1}^{n} \left[f_i(x) - g_i \right] \tag{2}$$

where $s\left(f(x), \mu, g\right)$ is the achievement-scalarizing function, μ_i represents the weight (relative importance) of the *i*-th objective $f_i(x)$, and ρ represents a small positive number.

The first component of (2), namely $\min_{i=1,2,\cdots,n} \left\{ \mu_i \left[f_i(x) - g_i \right] \right\}$ represents the greatest unachievement, and the second component of (2), namely $\rho \sum_{i=1}^{n} \left[f_i(x) - g_i \right]$ is an augmentation term which can guarantee the optimum solution of (2) and is an efficient solution of Problem (1).

Inspired by the reference point technique for multi-objective optimization models, it has also been used for multi-criteria decision making problems by developing two indexes (one is the strong synthetic index, and another is the weak synthetic index) as presented in Eq. 3 and Eq.4, respectively (Ruzi *et al*., 2011)).

$$I_S = \min_{i=1,2,\cdots,n} \left\{ \mu_i s\left[f_i(x), g_i^a, g_i^r \right] \right\} \tag{3}$$

$$I_W = \rho \sum_{i=1}^{n} s\left[f_i(x), g_i^a, g_i^r \right] \tag{4}$$

where I_S represents the strong synthetic index, I_W represents the weak synthetic index, $s\left[f_i(x), g_i^a, g_i^r \right]$ represents the achievement-scalarizing function, μ_i represents the weight (relative importance) of the *i*-th objective function, g_i^a represents the aspiration point of the the *i*-th objective/criterion, and g_i^r represents the reservation point of the i-th objective/criterion.

Double reference points were used in establishing the achievement-scalarizing function for determining the achievement of each alternative with respect to each objective/criterion. The reservation point is lower limit of acceptance for each objective/criterion, and the aspiration point represents the desirable expectation of the decision-makers/stakeholders on each objective/criterion. The reference point technique has been extended to interval condition for determining the weak synthetic sustainability index and the synthetic sustainability index under uncertainties in this study. The Hierarchical Best-Worst Method was employed to determine the weight of the objective/criteria, and the interval reference point technique has been developed for the weak synthetic sustainability index and the synthetic sustainability index under uncertainties.

3.1 Hierarchical Best-Worst Method

There are various weighting methods that can be used to determine the weights of the criteria and the sub-criteria for sustainability assessment, e.g., Entropy weighting method, Analytic Hierarchy Process (AHP), fuzzy AHP, and best-worst method (BWM). Among these methods, BWM as an advanced weighting method developed by Rezaei (2015) has two significant advantages: (i) less comparisons comparing with other traditional weighting methods and (ii) easy for achieving relatively higher consistency comparing with other traditional weighting methods. Accordingly, it has been widely used in various fields recently, e.g., location selection (Kheybari *et al.*, 2019), service quality evaluation (Gupta, 2018), supplier selection (Rezaei *et al.*, 2016), technology selection (van de Kaa *et al.*, 2017), risk management (Wang *et al.*, 2019), and emergency facility planning (Nyimbili and Erden, 2020), etc.

The Hierarchical Best-Worst Method (HBWM) was developed by Tabatabaie *et al.* (2019) based on the work of Rezaei (2015), it is an efficient tool for hierarchical decision-making which enables the users to determine the weights of the criteria as well as that of the sub-criteria in each criterion in one programming model simultaneously. The most significant advantage of this method is that it requires only one running to determine the global weights of the sub-criteria with high consistency. The four steps of HBWM have been summarized as follows based on the work of Tabatabaie *et al.* (2019):

Step 1: Identifying the best criterion and the worst criterion among all the criteria, and identifying the best sub-criterion and the worst sub-criterion among all the sub-criteria in each criterion.

We can assume that there are N criteria (C_1, C_2,…, C_N) and K sub-criteria (C_{j1}, C_{j2},…, C_{iK}) in the *j*-th criterion (C_j) which denote the best (i.e. the most important and the most preferable) and the worst (i.e. the least important and the least preferable) criteria by C_B and C_W, respectively. Similarly, the best and the worst sub-criteria in each criterion can also be determined, denoted by C_{jB} and C_{jW}, respectively.

Step 2: Determining the Best-to-Others (BO) and Others-to-Worst (OW) vectors. The BO vector can be determined by comparing C_B with each of the N criteria. For instance, the preference or priority of C_B over C_j (j=1,2,…,N) can be determined by the comparison of C_B with C_j (j=1,2,…,N), denoted by a_{Bj}. In a similar way, the comparison of C_{jB} as the best criterion in the *j*-th criterion (C_j) with each of the sub-criteria in the *j*-th criterion can be used to determine the preference or priority of C_{Bj} over C_{jt} (t=1,2,…,K). Then, the BO vectors can be determined, see Eq. 5 and Eq. 6. The nine numbers (from 1 to 9) and their reciprocals can be used to determine the relative preference or priority of one criterion (sub-criterion) over another criterion (sub-criterion), as presented in Table 2.

Table 2: The pairwise comparison scales used in BWM (Saaty, 1980; 2008)

Scale	The comparison one criterion (sub-criterion) over another criterion (sub-criterion)
1	Equal importance
2	Between equal importance and moderate importance
3	Moderate importance

4	Between moderate importance and strong importance
5	Strong importance
6	Between strong importance and demonstrated importance
7	Demonstrated importance
8	Between demonstrated importance and absolute importance
9	Absolute importance
Reciprocals above	The relative preference/priority of the *i*-th activity comparing with the *j*-th activity is denoted by one of the above-mentioned non-zero numbers, then its reciprocal will be used to describe the preference/priority of the *j*-th activity comparing with the *i*-th activity

It is worth pointing out that $a_{Bj}=1$ when $j=B$. Similarly, $a_{Bt}^{j}=1$ when $t=B$.

$$BO=\begin{bmatrix} a_{B1} & a_{B2} & \cdots & a_{BN} \end{bmatrix} \quad (5)$$

where BO represents the Best-to-Others (BO) by comparing the best criterion with other criteria, and $a_{Bj}\,(j=1,2,\cdots,N)$ represents the relative importance/priority of the best criterion comparing with the *j*-th criterion.

$$BO^{j}=\begin{bmatrix} a_{B1}^{j} & a_{B2}^{j} & \cdots & a_{BK}^{j} \end{bmatrix} \quad (6)$$

where BO^{j} represents the Best-to-Others (BO) by comparing the best sub-criterion in the *j*-th criterion with other sub-criteria in the *j*-th criterion, and $a_{Bt}^{j}\,(t=1,2,\cdots,K)$ represents the relative importance/priority of the best sub-criterion in the *j*-th criterion comparing with each sub-criterion in the *j*-th criterion.

OW vector can be determined by comparing each of the N criteria with the worst criterion (C_W). For instance, the preference or priority of C_j (j=1,2,…,N) over C_W can be determined by the comparison of C_j (j=1,2,…,N) with C_W, denoted by a_{jW}. In a similar way, the comparison of each of the sub-criteria in the *j*-th criterion with the worst sub-criterion (C_{jW}) in the *j*-th criterion (C_j) can be used to determine the preference or priority of C_{jt} (t=1,2,…,K) over C_{jW}. Then, the OW vectors can be determined, see Eq. 7 and Eq.8.

$$OW = \begin{bmatrix} a_{1W} & a_{2W} & \cdots & a_{NW} \end{bmatrix} \tag{7}$$

where OW represents the Others-to-Worst (OW) by comparing each criterion with the worst criterion, and $a_{jW}\left(j=1,2,\cdots,N\right)$ represents the relative importance/priority of the *j*-th criterion comparing with the worst criterion (C_W).

$$OW^j = \begin{bmatrix} a_{1W}^j & a_{2W}^j & \cdots & a_{KW}^j \end{bmatrix} \tag{8}$$

where OW^j represents the Others-to-Worst (OW) by comparing each sub-criterion in the *j*-th criterion with the worst sub-criterion (C_{jW}) in the *j*-th criterion, and $a_{tW}^j\left(t=1,2,\cdots,K\right)$ represents the relative importance/priority of each sub-criterion in the *j*-th criterion comparing with the worst sub-criterion in the *j*-th criterion.

Step 3: Determine the weights of the criteria, the local weights of the sub-criteria in each criterion and the global weights of the sub-criteria. A programming model was established in this step, see the programming model presented in (9). The objective function of this model is to minimize the total deviations of the comparisons for each pair of criteria or each pair of sub-criteria made by the users. Many constraints have also been incorporated in this programming model. For instance, the sum of the weights of the criteria and that of the local weights of the sub-criteria in each criterion equal to one, and the weights of the criteria and the local weights of the sub-criteria in each criterion should be non-negative.

$$
\begin{aligned}
& Min \quad \xi+\xi_j \\
& \left|\omega_B - a_{Bj}\omega_j\right| \le \xi \quad (j=1,2,\cdots,N) \\
& \left|\omega_j - a_{jW}\omega_W\right| \le \xi \quad (j=1,2,\cdots,N) \\
& \left|\omega_{jB} - a^j_{Bt}\omega_{jt}\right| \le \xi_j \quad (t=1,2,\cdots,K) \\
& \left|\omega_{jt} - a^j_{tW}\omega_{jW}\right| \le \xi_j \quad (t=1,2,\cdots,K) \\
& \omega^G_{jt} = \omega_{jt}\omega_j \quad (j=1,2,\cdots,N;t=1,2,\cdots,K) \\
& \sum_{j=1}^{N}\omega_j = 1 \quad (j=1,2,\cdots,N) \\
& \omega_j > 0 \\
& \sum_{t=1}^{K}\omega_{jt} = 1 \quad (j=1,2,\cdots,N;t=1,2,\cdots,K) \\
& \omega_{jt} > 0
\end{aligned}
\tag{9}
$$

where ω_B represents the weight of the best criterion, ω_j represents the weight of the j-th criterion, ω_{jB} represents the local weight of the best criterion in the j-th criterion, ω_{jt} represents the local weight of the t-th criterion in the j-th criterion, ω_{jW} represents the local weight of the worst criterion in the j-th criterion, and ω^G_{jt} represents the global weight of the t-th criterion in the j-th criterion.

Step 4: Calculate the consistency ratio for consistency check. ξ^* and $\xi^*_j\,(j=1,2,\cdots,N)$ are the optimum values of ξ and ξ_j under the optimum conditions $\omega^*_j\,(j=1,2,\cdots,N)$, $\omega^*_{jt}\,(j=1,2,\cdots,N;t=1,2,\cdots,K)$ $\omega^{G*}_{jt}\,(j=1,2,\cdots,N;t=1,2,\cdots,K)$, and the consistency ratios for judging the consistency level of the comparisons can be determined by Eq.(10) and Eq.(11). The consistency index (CI) can be obtained according to the work of Rezaei (2015), as presented in Table 3.

$$CR = \frac{\xi^*}{CI} \tag{10}$$

$$CR_j = \frac{\xi_j^*}{CI} \tag{11}$$

where CR represents the consistency ratio for the judgments in determining the BO and OW vectors, and CR_j represents the consistency ratio for the judgments in determining the BO^j and OW^j vectors.

Table 3: Consistency index in BWM (Rezaei, 2015)

a_{BW}, a_{BW}^j	1	2	3	4	5	6	7	8	9
CI	0.00	0.44	1.00	1.63	2.30	3.00	3.73	4.47	5.23

If all the consistency ratios equal to zero, namely $CR = 0$ and $CR_j = 0(j = 1, 2, \cdots, N)$, it represents that the judgments are absolutely consistent. However, it is difficult to achieve absolute consistency. Therefore, we set 0.10 as the threshold value. In other words, the judgments can be recognized as consistent if all the consistency ratios are less than 0.10, or the users need to revise the corresponding BO, OW, $BO^j(j = 1, 2, \cdots, N)$ or $OW^j(j = 1, 2, \cdots, N)$ vectors.

The global weights of the criteria/indicators determined by the HBWM will be used in the interval reference technique presented in section 2.2.

3.2 Interval reference point technique

An interval reference point technique was developed based on the work of Ruiz *et al.* (2011) to address the decision-making matrix with both crisp numbers and the interval numbers. Comparing with the traditional reference point technique as well as other multi-criteria decision making methods, the developed interval reference point technique can address various uncertainties expressed in the format of interval numbers. The interval reference point technique developed in this study consists of five steps based on the work of Ruiz *et al.* (2011):

Step 1: Determining the decision-making matrix with both crisp numbers and interval numbers. We can assume that there are M alternatives $(A_1, A_2, \cdots, A_M)$ to be evaluated by N indicators $(C_1, C_2, \cdots, C_M)$, and the decision-making matrix consists of both crisp numbers and interval numbers which can be firstly determined. It is worth pointing out that interval numbers are used to represent the data uncertainties. Among these N indicators, the data of the alternatives with respect to the first P indicators are depicted by using the crisp numbers, and those of the alternatives with respect to the last (N-P) indicators including the *(P+1)*-th indicator, the *(P+2)*-th indicator,…, and until the N-th indicator are depicted by using the interval numbers.

$$\begin{array}{cccccccc} & C_1 & C_2 & \cdots & C_P & C_{P+1} & \vdots & C_N \\ A_1 & x_{11} & x_{12} & \cdots & x_{1P} & \left[x_{1(P+1)}^L \quad x_{1(P+1)}^U\right] & \cdots & \left[x_{1N}^L \quad x_{1N}^U\right] \\ A_2 & x_{21} & x_{22} & \cdots & x_{2P} & \left[x_{2(P+1)}^L \quad x_{2(P+1)}^U\right] & \cdots & \left[x_{2N}^L \quad x_{2N}^U\right] \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_M & x_{M1} & x_{M2} & \cdots & x_{MP} & \left[x_{M(P+1)}^L \quad x_{M(P+1)}^U\right] & \cdots & \left[x_{MN}^L \quad x_{MN}^U\right] \end{array} \tag{12}$$

where $x_{ij}\left(i=1,2,...,M;\ j=1,2,...,P\right)$ which is a crisp number represents the data of the *i*-th alternative with respect to the *j*-th indicator, and

$x_{ij}\left(i=1,2,...,M;\ j=P+1,P+2,...,N\right)$ which is an interval number represents the data of the *i*-th alternative with respect to the *j*-th indicator.

Step 2: Determining the aspiration and the reservation level with respect to each indicator. The indicators for the evaluation of the alternatives can be divided into "more is better" and "less is better" types. As for the "more is better" criteria, it means that the increase of the data with respect to these criteria, the more superior the alternative will be. On the contrary, the increase of the data with respect to the "less is better" criteria, the more inferior the alternative will be. In this step, two goals are set for each objective/criterion as the reference points. One is reservation point which represents lower limit of acceptance for each objective/criterion, and another is aspiration point which represents the desirable expectation of the decision-makers/stakeholders on each objective/criterion. The double reference points (reservation point and aspiration point) will be established for each indicator. There are three commonly used way for determining the reservation point and the aspiration point, including the *neutral* scheme, the *voting* scheme and the *statistical* scheme (Wierzbicki *et al*., 2000; Ruiz *et al.,* 2011). The *statistical* scheme was used to determine the reservation point and the aspiration point.

As for the "more is better" criteria:

The aspiration point of each "more is better" criterion was set to be greater than or equal to the average value of the alternatives with respect to each "more is better" criterion, because it represents the desirable expectation of the decision-makers/stakeholders on each objective/criterion, and it cannot satisfy the desirable expectation of the decision-makers/stakeholders if the value set for the aspiration point is too small. The reservation point of each "more is better" criterion was set to be greater than or equal to the minimum value of the alternatives with respect to each "more is better" criterion, because it represents the minimum acceptance of the decision-makers/stakeholders on each objective/criterion. Similarly, it also cannot satisfy the lower limit of acceptance of the decision-makers/stakeholders if the value set for the aspiration point is too small.

(1) If all the data of the alternatives with respect to the "more is better" criterion are all crisp numbers, then the reservation point and the aspiration point can be determined by Eq. 13 and Eq.14, respectively.

$$x_j^a = \frac{\sum_{i=1}^{M} x_{ij}}{M} + \alpha_1 \left(\max_{i=1,2,..,M} x_{ij} - \frac{\sum_{i=1}^{M} x_{ij}}{M} \right) \tag{13}$$

where x_j^a represents the aspiration point for the *j*-th indicator, and α_1 can take the value from 0 to 1.

α_1 represents the aspiration of the decision-makers, and the higher the value of α_1, the higher the expectation of the decision-makers on the aspiration point. When $\alpha_1 = 0$, $x_j^a = \frac{\sum_{i=1}^{M} x_{ij}}{M}$, when $\alpha_1 = 0.5$, $x_j^a = \frac{\sum_{i=1}^{M} x_{ij}}{2M} + \frac{\max_{i=1,2,..,M} x_{ij}}{2}$, and when $\alpha_1 = 1$, $x_j^a = \max_{i=1,2,..,M} x_{ij}$.

$$x_j^r = \frac{\sum_{i=1}^{M} x_{ij}}{M} - \beta_1 \left(\frac{\sum_{i=1}^{M} x_{ij}}{M} - \min_{i=1,2,...,M} x_{ij} \right) \tag{14}$$

where x_j^r represents the reservation point for the *j*-th indicator, and β_1 an take the value from 0 to 1.

β_1 represents the reservation of the decision-makers, and the smaller the value of β_1, the higher the expectation of the decision-makers on the reservation point.

When $\beta_1 = 0$, $x_j^r = \frac{\sum_{i=1}^{M} x_{ij}}{M}$, when $\beta_1 = 0.5$, $x_j^r = \frac{\sum_{i=1}^{M} x_{ij}}{2M} + \frac{\min_{i=1,2,..,M} x_{ij}}{2}$, and when $\beta_1 = 1$, $x_j^r = \min_{i=1,2,..,M} x_{ij}$.

(2) If the data of the alternatives with respect to the "more is better" criterion are interval numbers, then the reservation point and the aspiration point can be determined by Eq. 15 and Eq.16, respectively.

$$x_j^a = \frac{\sum_{i=1}^{M}\left(x_{ij}^L + x_{ij}^U\right)}{2M} + \alpha_1 \left(\max_{i=1,2,..,M} x_{ij}^U - \frac{\sum_{i=1}^{M}\left(x_{ij}^L + x_{ij}^U\right)}{2M} \right) \tag{15}$$

where x_j^a represents the aspiration point for the *j*-th indicator, and α can take the value from 0 to 1.

When $\alpha_1 = 0$, $x_j^a = \frac{\sum_{i=1}^{M}\left(x_{ij}^L + x_{ij}^U\right)}{2M}$, when $\alpha_1 = 0.5$, $x_j^a = \frac{\sum_{i=1}^{M}\left(x_{ij}^L + x_{ij}^U\right)}{4M} + \frac{\max_{i=1,2,..,M} x_{ij}^U}{2}$, and when $\alpha_1 = 1$, $x_j^a = \max_{i=1,2,..,M} x_{ij}^U$.

$$x_j^r = \frac{\sum_{i=1}^{M}\left(x_{ij}^L + x_{ij}^U\right)}{2M} - \beta_1 \left(\frac{\sum_{i=1}^{M}\left(x_{ij}^L + x_{ij}^U\right)}{2M} - \min_{i=1,2,..,M} x_{ij}^L \right) \tag{16}$$

where x_j^r represents the reservation point for the *j*-th indicator, and β_1 an take the value from 0 to 1.

When $\beta_1 = 0$, $x_j^r = \frac{\sum_{i=1}^{M}\left(x_{ij}^L + x_{ij}^U\right)}{2M}$, when $\beta_1 = 0.5$, $x_j^r = \frac{\sum_{i=1}^{M}\left(x_{ij}^L + x_{ij}^U\right)}{2M} + \frac{\min_{i=1,2,..,M} x_{ij}^L}{2}$, and when $\beta_1 = 1$, $x_j^r = \min_{i=1,2,..,M} x_{ij}^L$.

As for the "less is better" criteria:

The aspiration point of each “less is better” criterion was defined to be less than or equal to the average value of the alternatives with respect to each “less is better” criterion, because only small value with respect to each “less is better” criterion can satisfy the desirable expectation of the decision-makers/stakeholders on each objective/criterion, and the smaller the aspiration point, the higher the desirable expectation. The reservation point of each “less is better” criterion was defined to be between the average value and the maximum value of the alternatives with respect to each “more is better” criterion, and it cannot satisfy the lower limit of acceptance of the decision-makers/stakeholders if the value is too high.

(1) If all the data of the alternatives with respect to the “less is better” criterion are all crisp numbers, then the reservation point and the aspiration point can be determined by Eq. 17 and Eq.18, respectively.

$$x_j^a = \min_{i=1,2,..,M} x_{ij} + \alpha_2 \left(\frac{\sum_{i=1}^{M} x_{ij}}{M} - \min_{i=1,2,..,M} x_{ij} \right) \tag{17}$$

where x_j^a represents the aspiration point for the *j*-th indicator, and α can take the value from 0 to 1.

α_2 represents the aspiration of the decision-makers, and the smaller the value of α_2, the higher the expectation of the decision-makers on the aspiration point. When $\alpha_2 = 0$, $x_j^a = \min_{i=1,2,..,M} x_{ij}$, when $\alpha_2 = 0.5$, $x_j^a = \frac{\sum_{i=1}^{M} x_{ij}}{2M} + \frac{\min_{i=1,2,..,M} x_{ij}}{2}$, and when $\alpha_2 = 1$, $x_j^a = \frac{\sum_{i=1}^{M} x_{ij}}{M}$.

$$x_j^r = \frac{\sum_{i=1}^{M} x_{ij}}{M} + \beta_2 \left(\max_{i=1,2,..,M} x_{ij} - \frac{\sum_{i=1}^{M} x_{ij}}{M} \right) \tag{18}$$

where x_j^r represents the aspiration point for the j-th indicator, and β_2 can take the value from 0 to 1.

β_2 represents the reservation of the decision-makers, and the smaller the value of β_2, the higher the expectation of the decision-makers on the reservation point. When $\beta_2 = 0$, $x_j^r = \frac{\sum_{i=1}^{M} x_{ij}}{M}$, when $\beta_2 = 0.5$, $x_j^r = \frac{\sum_{i=1}^{M} x_{ij}}{2M} + \frac{\max_{i=1,2,..,M} x_{ij}}{2}$, and when $\beta_2 = 1$, $x_j^r = \max_{i=1,2,..,M} x_{ij}$.

(2) If the data of the alternatives with respect to the "less is better" criterion are interval numbers, then the reservation point and the aspiration point can be determined by Eq. 19 and Eq.20, respectively.

$$x_j^a = \min_{i=1,2,..,M} x_{ij}^L + \alpha_2 \left(\frac{\sum_{i=1}^{M} \left(x_{ij}^L + x_{ij}^U \right)}{2M} - \min_{i=1,2,..,M} x_{ij}^L \right) \tag{19}$$

where x_j^a represents the aspiration point for the j-th indicator, and α_2 can take the value from 0 to 1.

When $\alpha_2 = 0$, $x_j^a = \min_{i=1,2,..,M} x_{ij}^L$, when $\alpha_2 = 0.5$, $x_j^a = \frac{\sum_{i=1}^{M} \left(x_{ij}^L + x_{ij}^U \right)}{4M} + \frac{\min_{i=1,2,..,M} x_{ij}^L}{2}$, and when $\alpha_2 = 1$, $x_j^a = \frac{\sum_{i=1}^{M} \left(x_{ij}^L + x_{ij}^U \right)}{2M}$.

$$x_j^r = \frac{\sum_{i=1}^{M}\left(x_{ij}^L + x_{ij}^U\right)}{2M} + \beta_2 \left(\max_{i=1,2,..,M} x_{ij}^U - \frac{\sum_{i=1}^{M}\left(x_{ij}^L + x_{ij}^U\right)}{2M} \right) \tag{20}$$

where x_j^r represents the aspiration point for the j-th indicator, and β_2 can take the value from 0 to 1.

When $\beta_2 = 0$, $x_j^r = \frac{\sum_{i=1}^{M}\left(x_{ij}^L + x_{ij}^U\right)}{2M}$, when $\beta_2 = 0.5$, $x_j^r = \frac{\sum_{i=1}^{M}\left(x_{ij}^L + x_{ij}^U\right)}{4M} + \frac{\max_{i=1,2,..,M} x_{ij}^U}{2}$, and when $\beta_2 = 1$, $x_j^r = \max_{i=1,2,..,M} x_{ij}^U$.

Step 3: Establishing the achievement scalarizing function for determining the achievement of each alternative with respect to the criteria. The achievement scalarizing functions under different conditions were established for measuring the positions of the alternatives with respect to the references points in terms of each indicator (Ruiz *et al.*, 2011; Cabello *et al.*, 2019).

As for the "more is better" criteria:

(1) If all the data of the alternatives with respect to the "more is better" criterion are all crisp numbers, the achievement of the i-th alternative with respect to the j-th criterion can be determined by the achievement scalarizing function ,as Eq.21 (Ruiz *et al.*, 2011).

$$s_{ij} = \begin{cases} \delta \dfrac{x_{ij} - x_j^r}{x_j^r - \min_{i=1,2,..,M} x_{ij}} & , \min_{i=1,2,..,M} x_{ij} \le x_{ij} \le x_j^r \\ \dfrac{x_{ij} - x_j^r}{x_j^a - x_j^r} & , x_j^r \le x_{ij} \le x_j^a \\ 1 + \gamma \dfrac{x_{ij} - x_j^a}{\max_{i=1,2,..,M} x_{ij} - x_j^a} & , x_j^a \le x_{ij} \le \max_{i=1,2,..,M} x_{ij} \end{cases} \tag{21}$$

where s_{ij} which is a crisp number represents the achievement of the i-th alternative with respect to the j-th indicator, $\delta>0$ which is a parameter represents the penalizing factor for the data under the reservation point, and $\gamma>0$ representing the rewarding factor for the data greater than the aspiration point.

The crisp number s_{ij} can be rewritten in the format of interval number $s_{ij} = \left[s_{ij} \quad s_{ij} \right]$.

When $\delta=1$ and $\gamma=1$, s_{ij} takes the value within the interval [-1 0] when the data of the alternatives with respect to the criteria are below the reservation point, s_{ij} takes the value within the interval [0 1] when the data of the alternatives with respect to the criteria are between the reservation point and the aspiration point, and s_{ij} takes the value within the interval [1 2] when the data of the alternatives with respect to the criteria are greater than the aspiration point.

(2) If the data of the alternatives with respect to the "more is better" criterion are interval numbers, the achievement of the i-th alternative with respect to the j-th criterion can be determined by Eq.22.

$$s_{ij}^{\pm} = \left[s_{ij}^{L} \quad s_{ij}^{U} \right] = \begin{cases} \left[\delta \dfrac{x_{ij}^{L} - x_{j}^{r}}{x_{j}^{r} - \min\limits_{i=1,2,..,M} x_{ij}} \quad \delta \dfrac{x_{ij}^{U} - x_{j}^{r}}{x_{j}^{r} - \min\limits_{i=1,2,..,M} x_{ij}} \right], & \min\limits_{i=1,2,..,M} x_{ij} \le x_{ij}^{L} \le x_{ij}^{U} \le x_{j}^{r} \\ \left[\delta \dfrac{x_{ij}^{L} - x_{j}^{r}}{x_{j}^{r} - \min\limits_{i=1,2,..,M} x_{ij}} \quad \dfrac{x_{ij}^{U} - x_{j}^{r}}{x_{j}^{a} - x_{j}^{r}} \right], & \min\limits_{i=1,2,..,M} x_{ij} \le x_{ij}^{L} \le x_{j}^{r} \text{ and } x_{j}^{r} \le x_{ij}^{U} \le x_{j}^{a} \\ \left[\dfrac{x_{ij}^{L} - x_{j}^{r}}{x_{j}^{a} - x_{j}^{r}} \quad \dfrac{x_{ij}^{U} - x_{j}^{r}}{x_{j}^{a} - x_{j}^{r}} \right], & x_{j}^{r} \le x_{ij}^{L} \le x_{ij}^{U} \le x_{j}^{a} \\ \left[\dfrac{x_{ij}^{L} - x_{j}^{r}}{x_{j}^{a} - x_{j}^{r}} \quad 1 + \gamma \dfrac{x_{ij}^{U} - x_{j}^{a}}{\max\limits_{i=1,2,..,M} x_{ij} - x_{j}^{a}} \right], & x_{j}^{r} \le x_{ij}^{L} \le x_{j}^{a} \text{ and } x_{j}^{a} \le x_{ij}^{U} \le \max\limits_{i=1,2,..,M} x_{ij} \\ \left[1 + \gamma \dfrac{x_{ij}^{L} - x_{j}^{a}}{\max\limits_{i=1,2,..,M} x_{ij} - x_{j}^{a}} \quad 1 + \gamma \dfrac{x_{ij}^{U} - x_{j}^{a}}{\max\limits_{i=1,2,..,M} x_{ij} - x_{j}^{a}} \right], & x_{j}^{a} \le x_{ij}^{L} \le x_{ij}^{U} \le \max\limits_{i=1,2,..,M} x_{ij} \end{cases}$$

(22)

where $s_{ij}^{\pm} = \left[s_{ij}^{L} \quad s_{ij}^{U} \right]$ which is an interval number represents the achievement of the *i*-th alternative with respect to the *j*-th indicator, $\delta > 0$ which is a parameter represents the penalizing factor for the data under the reservation point, and $\gamma > 0$ representing the rewarding factor for the data greater than the aspiration point.

As for the "less is better" criteria:

(1) If all the data of the alternatives with respect to the "less is better" criterion are all crisp numbers, the achievement of the *i*-th alternative with respect to the *j*-th criterion can be determined by Eq.23.

$$s_{ij} = \begin{cases} 1+\gamma \dfrac{x_j^a - x_{ij}}{x_j^a - \min\limits_{i=1,2,..,M} x_{ij}} & , \min\limits_{i=1,2,..,M} x_{ij} \le x_{ij} \le x_j^a \\ \dfrac{x_j^r - x_{ij}}{x_j^r - x_j^a} & , x_j^a \le x_{ij} \le x_j^r \\ \delta \dfrac{x_j^r - x_{ij}}{\max\limits_{i=1,2,..,M} x_{ij} - x_j^r} & , x_j^r \le x_{ij} \le \max\limits_{i=1,2,..,M} x_{ij} \end{cases} \tag{23}$$

where s_{ij} which is a crisp number represents the achievement of the *i*-th alternative with respect to the *j*-th indicator, $\delta > 0$ which is a parameter represents the penalizing factor for the data above the reservation point, and $\gamma > 0$ representing the rewarding factor for the data less than the aspiration point.

When $\delta = 1$ and $\gamma = 1$, s_{ij} takes the value within the interval [1 2] when the data of the alternatives with respect to the criteria are below the aspiration point, s_{ij} takes the value within the interval [0 1] when the data of the alternatives with respect to the criteria are between the aspiration point and the reservation point, and s_{ij} takes the value within the interval [-1 0] when the data of the alternatives with respect to the criteria are above the reservation point.

(2) If the data of the alternatives with respect to the "less is better" criterion are interval numbers, the achievement of the i-th alternative with respect to the j-th criterion can be determined by Eq.24.

$$s_{ij}^{\pm}=\left[s_{ij}^{L}\quad s_{ij}^{U}\right]=\begin{cases}\left[1+\gamma\dfrac{x_j^a-x_{ij}^U}{x_j^a-\min\limits_{i=1,2,..,M}x_{ij}}\quad 1+\gamma\dfrac{x_j^a-x_{ij}^L}{x_j^a-\min\limits_{i=1,2,..,M}x_{ij}}\right], & \min\limits_{i=1,2,..,M}x_{ij}\le x_{ij}^L\le x_{ij}^U\le x_j^a\\ \left[\dfrac{x_j^r-x_{ij}^U}{x_j^r-x_j^a}\quad 1+\gamma\dfrac{x_j^a-x_{ij}^L}{x_j^a-\min\limits_{i=1,2,..,M}x_{ij}}\right], & \min\limits_{i=1,2,..,M}x_{ij}\le x_{ij}^L\le x_j^a \text{ and } x_j^a\le x_{ij}^U\le x_j^r\\ \left[\dfrac{x_j^r-x_{ij}^U}{x_j^r-x_j^a}\quad \dfrac{x_j^r-x_{ij}^L}{x_j^r-x_j^a}\right], & x_j^a\le x_{ij}^L\le x_{ij}^U\le x_j^r\\ \left[\delta\dfrac{x_j^r-x_{ij}^U}{\max\limits_{i=1,2,..,M}x_{ij}-x_j^r}\quad \dfrac{x_j^r-x_{ij}^L}{x_j^r-x_j^a}\right], & x_j^a\le x_{ij}^L\le x_j^r \text{and } x_j^r\le x_{ij}^U\le\max\limits_{i=1,2,..,M}x_{ij}\\ \left[\delta\dfrac{x_j^r-x_{ij}^U}{\max\limits_{i=1,2,..,M}x_{ij}-x_j^r}\quad \delta\dfrac{x_j^r-x_{ij}^L}{\max\limits_{i=1,2,..,M}x_{ij}-x_j^r}\right], & x_j^r\le x_{ij}^L\le x_{ij}^U\le\max\limits_{i=1,2,..,M}x_{ij}\end{cases}$$

(24)

where $s_{ij}^{\pm}=\left[s_{ij}^{L}\quad s_{ij}^{U}\right]$ which is a crisp number represents the achievement of the i-th alternative with respect to the j-th indicator, $\delta>0$ which is a parameter represents the penalizing factor for the data above the reservation point, and $\gamma>0$ representing the rewarding factor for the data less than the aspiration point.

Step 4: Determining the weak synthetic sustainability index and the strong synthetic sustainability index. The weak synthetic sustainability index and the strong synthetic index can be determined by Eq. 25 and Eq.26, respectively.

$$I_{i,SS}^{\pm}=\left[I_{i,SS}^{L}\quad I_{i,SS}^{U}\right]=\min_{j=1,2,\cdots,N}\omega_j s_{ij}^{\pm}=\min_{j=1,2,\cdots,N}\left[\omega_j s_{ij}^{L}\quad \omega_j s_{ij}^{U}\right] \tag{25}$$

$$I_{i,WS}^{\pm}=\left[I_{i,WS}^{L}\quad I_{i,WS}^{U}\right]=\sum_{j=1}^{N}\omega_j s_{ij}^{\pm}=\sum_{j=1}^{N}\left[\omega_j s_{ij}^{L}\quad \omega_j s_{ij}^{U}\right] \tag{26}$$

where $I_{i,SS}^{\pm}=\left[I_{i,SS}^{L} \quad I_{i,SS}^{U}\right]$ which is an interval number represents the strong synthetic sustainability index of the *i*-th alternative, and $I_{i,WS}^{\pm}=\left[I_{i,WS}^{L} \quad I_{i,WS}^{U}\right]$ which is an interval number represents the weak synthetic sustainability index of the *i*-th alternative.

The weak synthetic sustainability index means that the good performances in economic aspects or other aspects can compensate the bad performances in environmental aspect, it is an overall measure of the integrated sustainability, but the strong synthetic sustainability index means that the environmental performances cannot be compensated by economic or other dimensional indictors, and it can provide the decision-makers the alarm signs in some sustainability indicators (Cabello *et al.*, 2019). The higher the weak synthetic sustainability index or the strong synthetic sustainability index, the better the alternative will be. The negative values existing in the strong synthetic sustainability index, namely the interval $I_{i,SS}^{\pm}=\left[I_{i,SS}^{L} \quad I_{i,SS}^{U}\right]$, represents that the alternative performs under the reservation point for at least one indicator, and if $I_{i,SS}^{\pm}=\left[I_{i,SS}^{L} \quad I_{i,SS}^{U}\right]$ is greater than 1, it means that the performances of this alternative on all the indicators are greater than the aspiration points. The weak synthetic sustainability index represents the overall sustainability performance of a alternative by incorporating the performances with respect to all the indicators into one index.

Step 5: Determining the overall picture of sustainability in a graphical way.

Both the weak synthetic sustainability indicator and the strong synthetic sustainability indicator are interval numbers, and it is usually difficult to rank the interval numbers. Therefore, the interval number ranking method developed by Xu and

Da (2003) was employed in this study. According to Step 4, the strong synthetic sustainability index $I_{i,SS}^{\pm} = \left[I_{i,SS}^{L} \quad I_{i,SS}^{U} \right]$ for all the M alternatives can be determined, and the probability of the strong synthetic sustainability index of the *i*-th alternative be greater than that of the *k*-th alternative can be determined by Eq. 23 based on the work of Xu and Da (2003). This method for the ranking of interval numbers has been widely used in many studies (Ren and Toniolo, 2018).

$$\begin{aligned} p_{ik}^{SS} &= p\left(I_{i,SS}^{\pm} \geq I_{k,SS}^{\pm}\right) = p\left(\left[I_{i,SS}^{L} \quad I_{i,SS}^{U}\right] \geq \left[I_{k,SS}^{L} \quad I_{k,SS}^{U}\right]\right) \\ &= \max\left\{1 - \max\left[\frac{I_{k,SS}^{U} - I_{i,SS}^{L}}{I_{i,SS}^{U} - I_{i,SS}^{L} + I_{k,SS}^{U} - I_{k,SS}^{L}}, 0\right], 0\right\} \end{aligned} \tag{27}$$

where p_{ik}^{SS} represents the probability of the strong synthetic sustainability index of the *i*-th alternative be greater than that of the *k*-th alternative

It is apparent that when $I_{i,SS}^{\pm} = \left[I_{i,SS}^{L} \quad I_{i,SS}^{U} \right] = I_{k,SS}^{\pm} = \left[I_{k,SS}^{L} \quad I_{k,SS}^{U} \right]$, $p_{ik}^{SS} = 0.50$. It is worth pointing out that Eq.23 cannot be used for if $I_{i,SS}^{\pm}$ and $I_{k,SS}^{\pm}$ are both crisp numbers, because when $I_{i,SS}^{U} = I_{i,SS}^{L}$ and $I_{k,SS}^{U} = I_{k,SS}^{L}$, the denominator is zero. As for the comparison of two crisp numbers, the probability of one crisp number being greater than another has been defined based on the following two rules:

(1) If $I_{i,SS}^{U} = I_{i,SS}^{L} > I_{k,SS}^{U} = I_{k,SS}^{L}$, then $p(\left[I_{i,SS}^{L} \quad I_{i,SS}^{U} \right] \geq \left[I_{k,SS}^{L} \quad I_{k,SS}^{U} \right]) = 1$ and $p(\left[I_{k,SS}^{L} \quad I_{k,SS}^{U} \right] \geq \left[I_{i,SS}^{L} \quad I_{i,SS}^{U} \right]) = 0$;

(2) $p(\left[I_{i,SS}^{L} \quad I_{i,SS}^{U} \right] \geq \left[I_{i,SS}^{L} \quad I_{i,SS}^{U} \right]) = 0.50$.

In a similar way, the probability matrix for the strong synthetic sustainability can be determined by comparing the strong synthetic sustainability indexes with respect to each pair of alternatives, as presented in Eq.28 (Ren and Toniolo, 2018).

$$P^{SS}=\begin{array}{c|cccc} & I_{1,SS}^{\pm} & I_{2,SS}^{\pm} & \cdots & I_{M,SS}^{\pm} \\ \hline I_{1,SS}^{\pm} & 0.50 & p_{12}^{SS} & \cdots & p_{1M}^{SS} \\ I_{2,SS}^{\pm} & p_{21}^{SS} & p_{22}^{SS} & \cdots & p_{2M}^{SS} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I_{M,SS}^{\pm} & p_{M1}^{SS} & p_{M2}^{SS} & \cdots & 0.50 \end{array} \tag{28}$$

where P^{SS} represents the probability matrix for the strong synthetic sustainability, and p_{ik}^{SS} represents the probability of the strong synthetic sustainability index of the i-th alternative be greater than that of the k-th alternative.

Then, the integrated strong synthetic sustainability (SSS) performance of the i-th alternative can be determined according to the work of Xu and Da (2003)

$$SSS_i=\frac{\sum_{j=1}^{M} p_{ij}^{SS}+M/2-1}{M(M-1)} \tag{29}$$

where SSS_i which is a crisp number represents the integrated strong synthetic sustainability performance of the i-th alternative.

In the similar way, the probability matrix for the weak synthetic sustainability can also be determined by comparing the weak synthetic sustainability indexes with respect to each pair of alternatives, as presented in Eq.30.

$$P^{WS}=\begin{array}{c|cccc} & I_{1,WS}^{\pm} & I_{2,WS}^{\pm} & \cdots & I_{M,WS}^{\pm} \\ \hline I_{1,WS}^{\pm} & 0.50 & p_{12}^{WS} & \cdots & p_{1M}^{WS} \\ I_{2,WS}^{\pm} & p_{21}^{WS} & p_{22}^{WS} & \cdots & p_{2M}^{WS} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I_{M,WS}^{\pm} & p_{M1}^{WS} & p_{M2}^{WS} & \cdots & 0.50 \end{array} \tag{30}$$

where P^{WS} represents the probability matrix for the weak synthetic sustainability, and p_{ik}^{WS} represents the probability of the weak synthetic sustainability index of the i-th alternative be greater than that of the k-th alternative.

Then, the integrated weak synthetic sustainability (WSS) performance of the i-th alternative can also be determined, as presented in Eq.31.

$$WSS_i = \frac{\sum_{j=1}^{M} p_{ij}^{WS} + M/2 - 1}{M(M-1)} \tag{31}$$

where WSS_i which is a crisp number represents the integrated weak synthetic sustainability performance of the *i*-th alternative.

The overall picture of sustainability of all the alternatives can be represented in a graphical way, as illustrated in Figure 1, and it is a two-dimensional representation, and the overall sustainability of each alternative can be represented by a point in which its value in the vertical axis represents its integrated weak synthetic sustainability performance and its value in the horizontal axis represents its integrated strong synthetic sustainability performance. Accordingly, the most sustainable alternatives should be in the top right corner of this two-dimensional representation.

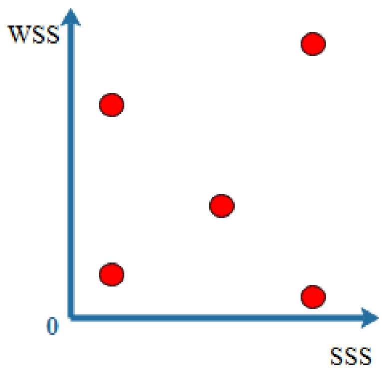

Figure 1: The overall picture of sustainability

4. Results and discussion

Energy storage is the critical technology for addressing the problem of intermittency in the use of renewable energy, and there are various technologies with different sustainability performances. In order to illustrate the developed synthetic sustainability index under uncertainties, five energy storage technologies including pumped hydro (PH), compress air (CA), Lead-Acid (LA), Lithium-ion (Li), and flywheel (F) derived from Ren and Ren (2018) were studied in this study. There are two main reasons of choosing energy storage technologies to illustrate the developed interval reference point technique developed in this study: (i) the selection of energy storage technologies is a typical problem in which we need to consider both the weak synthetic sustainability index (compensatory sustainability index) and the strong synthetic sustainability index (non-compensatory sustainability index) , because the aggregated/composite sustainability index of an energy storage technology determined by other non-reference-point-techniques are usually determined with the compensation of the bad environmental performances by the economic or social sustainability performances. Thus, the alternative with a high aggregated/composite sustainability index may perform very bad in environmental sustainability because of the good performances in other sustainability dimensions; and (ii) there are data uncertainties in the selection of the most sustainable energy storage technology-the data of some energy storage technologies with respect to some criteria for sustainability assessment are usually interval numbers rather than crisp numbers. The study on the typical problem (the sustainability-oriented selection of energy storage technologies) has a good demonstration and dissemination effect on the interval reference point technique developed in this study.

There are usually three dimensions in sustainability, and they are economy, environment, and society (the three pillars of sustainability). Just like what we have mentioned in the Introduction section, the criteria in these three dimensions are not enough to measure sustainability because the criteria in other dimensions (e.g., technological, political, cultural and psychological dimensions) usually have significant influences on the criteria in the three pillars of sustainability. For instance, subsidization scheme belonging to the policy dimension has significant influence on the net present value (NPV) which is an important criterion to measure economic performance, and technology maturity belonging to the technological dimension influences energy consumption and energy consumption dominates greenhouse gases (GHG) emissions (GHG potential is a critical criterion in environmental dimension). Accordingly, the criteria for sustainability assessment should be determined from a multi-dimensional perspective. Moreover, different stakeholders have different preferences/concerns when selecting the most sustainable option along multiple alternatives, and the selection of the criteria for sustainability assessment should also incorporate the preferences of the stakeholders. Based on the literature reviews of the current works on energy storage technologies, it is apparent that almost all the stakeholders (e.g., scientists, researchers, engineers, investors and governments) have more concerns on the economic (e.g., capital cost and operating cost), performance (e.g., materials intensity and energy intensity), technological (e.g., maturity) and environmental (CO_2 intensity and environmental impact) aspects. Therefore, four dimensional criteria including economic, performance, technological and environmental criteria were employed to determine the synthetic sustainability index of each energy storage technology. There are three sub-criteria in economic criterion, and they are capital cost, life and operating cost. There are four sub-criteria in performance criterion, and they are energy efficiency, materials intensity, energy intensity and energy density. Maturity is the only sub-criterion in technological criterion. And CO_2 and environmental impact are the two sub-criteria used to measure the environmental criterion. The data of the five energy storage technologies with respect to each sub-criterion were presented in Table 4.

Table 4: The data of the five energy storage technologies with respect to each sub-criterion

Criteria	Sub-criteria	Unit	Pumped hydro	Compressed Air	Lead-Acid	Lithium-ion	Flywheel	
	Capital cost (EC_1)	€/kWh	22.5-45	6.5-37.5	148.33-185	900-1300	400-800	Díaz-González, 2012
Economic (EC)	Life (EC_2)	years	40-50	35-35	5-15	14-16	20-20	Díaz-González, 2012
	Operating cost (EC_3)	\$/MJ	0.0006-0.0014	0.0001-0.0019	0.0008-0.0028	0.0019-0.0047	0.0008-0.0017	Ashby and Polyblank, 2012
	Energy efficiency (P_1)	%	69-74	38-39.25	72.5-80	83-83	85-85	Díaz-González, 2012
Performance (P)	Materials intensity (P_2)	kg/MJ	60-120	2-12	4.5-12	1.5-2.7	17-500	Ashby and Polyblank, 2012
	Energy intensity (P_3)	$MJ_{embodied}$/MJ	100-200	74-74	110-980	330-580	750-760	Ashby and Polyblank, 2012
	Energy	Wh/kg	0.5-1.5	30-60	30-50	75-200	10-30	Chen *et*

	density (P_4)							*al.*, 2009
Technological (T)	Maturity (T_1)	/	0.5008	0.1971	0.0775	0.1123	0.1123	Calculated by Ren and Ren (2018) based on the work of Beaudin *et al.*, 2010
Environmental (EN)	CO_2 intensity (EN_1)	kg/MJ	8-16	5.3-5.3	5-130	19-50	90-100	Ashby and Polyblank, 2012
	Environmental impact (EN_2)	/	0.0941	0.1525	0.0583	0.2660	0.4291	Calculated by Ren and Ren (2018) based on the work of Evans *et al.*, 2012

Sources: adapted from Ren and Ren (2018).

The HBWM was firstly employed to determine the weights of the four criteria, the local weights of the sub-criteria in each criterion and the global weights of the sub-criteria, and the procedures were specified as follows:

Step 1: As for the four criteria, economic (EC) criterion and environmental (EN) criterion have been recognized as the most important criterion and the least important criterion, respectively. Among the three sub-criteria in the economic criterion, capital cost (EC_1) has been resigned as the most important and operating cost (EC_3) has been recognized as the least important. Among these four sub-criteria in performance criterion, energy efficiency (P_1) and materials intensity (P_2) have been recognized as the most important and the least important, respectively. As for the two sub-criteria in environmental criterion, environmental impact (EN_1) has been recognized as more important than CO_2 intensity (EN_2).

Step 2: The BO and OW vectors can be determined by the comparison of each pair of criteria and of each pair of sub-criteria in each criterion, as presented in Table 5.

Table 5: The BO and OW vectors

Goal	The most important: EC		The lease important: EN	
	EC	P	T	EN
BO	1	2	3	6
OW	6	3	2	1
Economic	The most important: EC_1		The lease important: EC_3	
	EC_1	EC_2	EC_3	
BO	1	3	5	
OW	5	2	1	
Performance	The most important: P_1		The most important: P_2	
	P_1	P_2	P_3	P_4
BO	1	7	3	4
OW	7	1	3	2
Environmental	The most important: EN_1		The most important: EN_2	
	EN_1	EN_2		
BO	4	1		

OW	1	4

Step 3: After determining the BO and OW vectors, the following programming model has been established:

$$
\begin{aligned}
&Min \quad \xi + \xi_{EC} + \xi_P + \xi_T + \xi_{EN} \\
&|\omega_{EC} - 2\omega_P| \le \xi \\
&|\omega_{EC} - 3\omega_T| \le \xi \\
&|\omega_{EC} - 6\omega_{EN}| \le \xi \\
&|\omega_P - 3\omega_{EN}| \le \xi \\
&|\omega_T - 2\omega_{EN}| \le \xi \\
&|\omega_{EC_1} - 3\omega_{EC_2}| \le \xi_{EC} \\
&|\omega_{EC_1} - 5\omega_{EC_3}| \le \xi_{EC} \\
&|\omega_{EC_2} - 2\omega_{EC_3}| \le \xi_{EC} \\
&|\omega_{P_1} - 7\omega_{P_2}| \le \xi_P \\
&|\omega_{P_1} - 3\omega_{P_3}| \le \xi_P \\
&|\omega_{P_1} - 4\omega_{P_4}| \le \xi_P \\
&|\omega_{P_3} - 3\omega_{P_2}| \le \xi_P \\
&|\omega_{P_4} - 2\omega_{P_2}| \le \xi_P \\
&|\omega_{EN_2} - 4\omega_{EN_1}| \le \xi_{EN} \\
&\omega_{EC_1}^G = \omega_{EC_1}\omega_{EC} \\
&\omega_{EC_2}^G = \omega_{EC_2}\omega_{EC} \\
&\omega_{EC_3}^G = \omega_{EC_3}\omega_{EC} \\
&\omega_{P_1}^G = \omega_{P_1}\omega_P \\
&\omega_{P_2}^G = \omega_{P_2}\omega_P \\
&\omega_{P_3}^G = \omega_{P_3}\omega_P \\
&\omega_{P_4}^G = \omega_{P_4}\omega_P \\
&\omega_{T_1}^G = \omega_{T_1}\omega_T \\
&\omega_{EN_1}^G = \omega_{EN_1}\omega_{EN} \\
&\omega_{EN_2}^G = \omega_{EN_2}\omega_{EN} \\
&\omega_{EC} + \omega_P + \omega_T + \omega_{EN} = 1 \\
&\omega_{EC}, \omega_P, \omega_T, \omega_{EN} > 0 \\
&\omega_{EC_1} + \omega_{EC_2} + \omega_{EC_3} = 1 \\
&\omega_{P_1} + \omega_{P_2} + \omega_{P_3} + \omega_{P4} = 1 \\
&\omega_{T_1} = 1 \\
&\omega_{EN_1} + \omega_{EN_2} = 1 \\
&\omega_{EC_1}, \omega_{EC_2}, \omega_{EC_3}, \omega_{P_1}, \omega_{P_2}, \omega_{P_3}, \omega_{P4}, \omega_{T_1}, \omega_{EN_1}, \omega_{EN_2} > 0
\end{aligned}
\tag{32}
$$

Table 6: The weights of the four criteria, the local weights of the sub-criteria in each criterion and the global weights of the ten sub-criteria determined by HBWM

Criteria	Weights by HBWM	Weights by NLFP (Ren and Ren, 2018)	Sub-criteria	Local weights by HBWM	Local weights by NLFP (Ren and Ren, 2018	Global weights by HBWM	Global weights by NLFP (Ren and Ren, 2018
Economic	**0.5000**	0.4592	Capital cost	**0.6500**	0.5000	**0.3250**	0.2296
			Life	**0.2250**	0.3333	**0.1125**	0.1531
			Operating cost	**0.1250**	0.1667	**0.0625**	0.0765
Performance	**0.2500**	0.2656	Energy efficiency	**0.5714**	0.5595	**0.1429**	0.1486
			Materials intensity	**0.0772**	0.0872	**0.0193**	0.0232
			Energy intensity	**0.2008**	0.2165	**0.0502**	0.0575
			Energy density	**0.1506**	0.1367	**0.0376**	0.0363
Technological	**0.1667**	0.1936	Maturity	**1**	1	**0.1667**	0.1936
Environmental	**0.0833**	0.0816	CO_2 intensity	**0.2000**	0.2000	**0.01667**	0.0163
			Environmental impact	**0.8000**	0.8000	**0.0667**	0.0653

After solving programming (32), the weights of the four criteria, the local weights of the sub-criteria in each criterion and the global weights of the ten sub-criteria can be determined, and the results were presented in Table 6. The consistency check results reveal that all the judgments for determining the BO and the OW vectors can be recognized as consistent, and the consistency check results were presented in Table 7. Therefore, the global weights of the ten sub-criteria determined by HBWM can be used for determining the synthetic sustainability index of each alternative energy storage technology.

Table 7: Consistency Check results

Goal	The most important: EC		The lease important: EN	
	EC	P	T	EN
BO	1	2	3	6
OW	6	3	2	1
ξ^*=0， $CI=1.63, CR=\frac{\xi^*}{CI}=0<0.10$				
Economic	The most important: EC_1		The lease important: EC_3	
	EC_1	EC_2	EC_3	
BO	1	3	5	
OW	5	2	1	
ξ_{EC}^*=0.0250， $CI=1.00, CR=\frac{\xi^*}{CI}=0.0250<0.10$				
Performance	The most important: P_1		The most important: P_2	
	P_1	P_2	P_3	P_4

BO	1	7	3	4
OW	7	1	3	2
ξ_{EC}^{*}=0.0309, $CI=1.63, CR=\frac{\xi^{*}}{CI}=0.0190<0.10$				
Environmental	The most important: EN_1		The most important: EN_2	
	EN_1	EN_2		
BO	4	1		
OW	1	4		
ξ_{EC}^{*}=0, $CI=0.44, CR=\frac{\xi^{*}}{CI}=0<0.10$				

The results determined by HBWM are similar and consistent to that determined by Ren and Ren (2018), and they employed the Non-Linear Fuzzy Prioritization (NLFP) method developed by Mikhailov (2003) to determine the global weights of the sub-criteria.

Among these ten sub-criteria for determining the synthetic sustainability index, five sub-criteria are the “more is better” criteria including life, energy efficiency, energy density, maturity and environmental impact (the data with respect to this criterion represents the relative performances on environmental impact). The data of these five energy storage technologies with respect to maturity and environmental impact are crisp numbers, but the data with respect to other three sub-criteria (life, energy efficiency and energy density) are interval numbers. The other five are “less are better” sub-criteria, including capital cost, operating cost, materials intensity, energy intensity and CO_2

intensity. The data of these five energy storage technologies with respect to these five "less are better" sub-criteria are all interval numbers. Let's set $\alpha = 0.50$ and $\beta = 0.50$ for determining the aspiration point and the reservation point, and the aspiration point and the reservation point with respect to each sub-criterion have been presented in Table 8.

Table 8: The aspiration point and the reservation point with respect to each sub-criterion

	T_1	EN_2	EC_2	P_1	P_4	EC_1	EC_3	P_2	P_3	EN_1
Aspiration point	0.3504	0.3145	37.5000	77.9375	124.3500	195.4915	0.0009	37.3350	234.9000	23.9300
Reservation point	0.1388	0.1291	15.0000	54.4375	24.6000	842.2415	0.0032	286.5850	687.9000	86.4300

After determining the aspiration point and the reservation point with respect to each sub-criterion, we set δ=1 and γ=1 in the achievement scalarizing function for determining the achievement of each energy storage technology with respect to each sub-criterion. Taking the data of flywheel (F) with respect to capital cost ([400 800] €/kWh) as an example, capital cost (EC_1) is an "less is better" criterion and its data is an interval number, therefore, Eq.24 can be used to determine the achievement of CA with respect to EC_1. The aspiration point and the reservation point with respect to EC_1 are 195.4915 and 842.2415 €/kWh, respectively. According to Eq.24, the achievement of flywheel (F) with respect to capital cost can be determined:

$$\begin{aligned} s_{ij}^{\pm} = \left[s_{ij}^{L} \quad s_{ij}^{U} \right] &= \left[\frac{x_j^r - x_{ij}^U}{x_j^r - x_j^a} \quad \frac{x_j^r - x_{ij}^L}{x_j^r - x_j^a} \right] \quad , x_j^a \le x_{ij}^L \le x_{ij}^U \le x_j^r \\ &= \left[\frac{842.2415 - 800}{842.2415 - 195.4915} \quad \frac{842.2415 - 400}{842.2415 - 195.4915} \right] \\ &= \left[0.0653 \quad 0.6838 \right] \end{aligned} \tag{33}$$

In a similar way, the achievements of these five energy storage technologies with respect to each sub-criterion can be determined, as presented in Table 9.

Table 9: The achievements of these five energy storage technologies with respect to each sub-criterion

	Pumped hydro	Compressed Air	Lead-Acid	Lithium-ion	Flywheel
T_1	2.0000	0.2757	-1.0000	-0.4318	-0.4318
EN_2	-0.4947	0.1259	-1.0000	0.7381	2.0000
EC_2	[1.2000 2.0000]	0.889	[-1.0000 0]	[-0.1000 0.0444]	0.2222
P_1	[0.6197 0.8324]	[-1.0000 - 0.9240]	[0.7686 1.2920]	1.7168	2.0000
P_4	[-1.0000 - 0.9585]	[0.0541 0.3549]	[0.0541 0.2546]	[0.5053 2.0000]	[-0.6058 0.0541]
EC_1	[1.7963 1.9153]	[1.8360 2.0000]	[1.0555 1.2495]	[-1.0000 - 0.1262]	[0.0653 0.6838]
EC_3	[0.7761	[0.5587	[0.1674	[-1.0000	[0.6457

	1.3631]	2.0000]	1.1083]	0.5587]	1.1083]	
P2	[0.7806	[1.7070	[1.7070	1.9665	[-1.0000	
	1.0617]	1.9860]	1.9163]	2.0000[]	1.2632]	
P3	[1.2169	2.0000	[-1.0000	[0.2382	[-0.2468	-
	1.8384]		1.9784]	0.7901]	0.2126]	
EN1	[1.4189	1.9842	[-1.0000	[0.5829	[-0.3115	-
	1.8415]		1.8689]	1.2604]	0.0819]	

According to the global weights of the sub-criteria in Table 5 and the achievements of the energy storage technologies with respect to the sub-criteria in Table 8, both the weak synthetic sustainability index and the strong synthetic sustainability index can be determined, and the results were presented in Figure 2 and Figure 3, respectively.

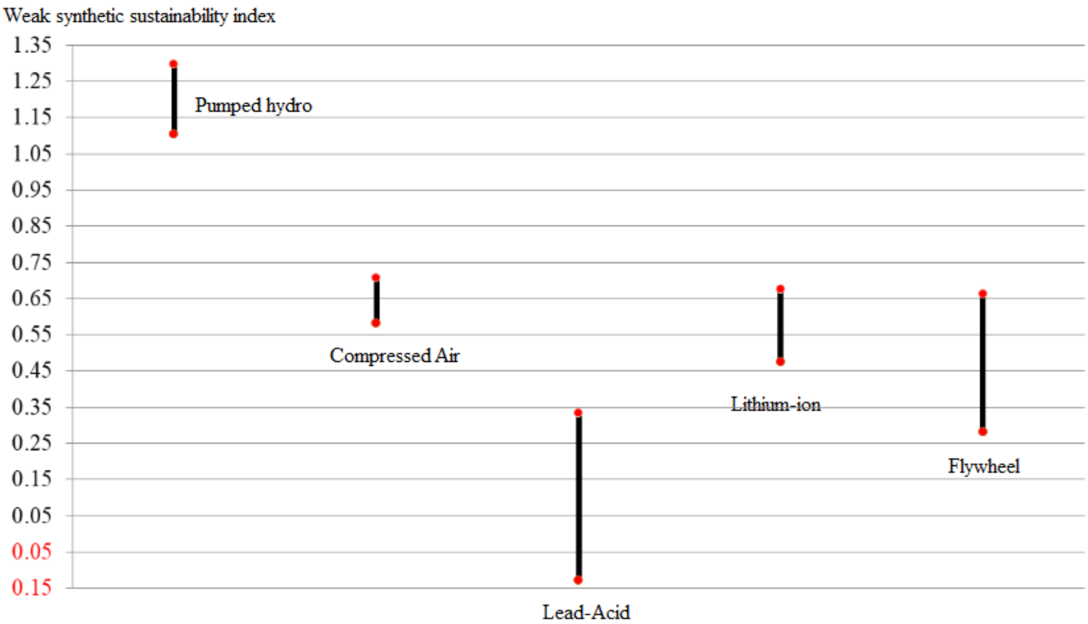

Figure 2: The weak sustainability index of each energy storage technology

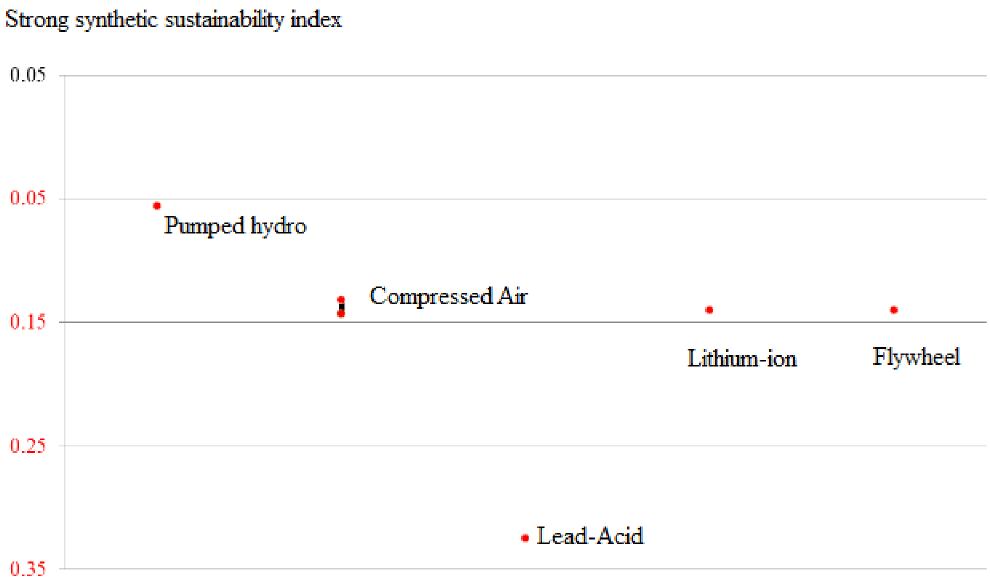

Figure 3: The strong sustainability index of each energy storage technology

The weak sustainability index of each energy storage technology is an interval number and thus Eq.27 can be used to determine the probability degree of the weak sustainability index of one energy storage technology be greater than that of another. Taking the probability degree of the weak sustainability index of Lithium-ion and that of Flywheel as an example,

$$
\begin{aligned}
p_{Li,F}^{SS} &= p\left(I_{LAiSS}^{\pm} \geq I_{F,SS}^{\pm}\right) = p\left(\left[I_{Li,SS}^{L} \quad I_{Li,SS}^{U}\right] \geq \left[I_{F,SS}^{L} \quad I_{F,SS}^{U}\right]\right) \\
&= \max\left\{1 - \max\left[\frac{I_{F,SS}^{U} - I_{Li,SS}^{L}}{I_{Li,SS}^{U} - I_{Li,SS}^{L} + I_{F,SS}^{U} - I_{F,SS}^{L}}, 0\right], 0\right\} \\
&= \max\left\{1 - \max\left[\frac{0.6629 - 0.4744}{0.6747 - 0.4744 + 0.6629 - 0.2086}, 0\right], 0\right\} \\
&= 0.2880
\end{aligned}
\tag{34}
$$

In a similar way, other elements in the probability matrix for the weak synthetic sustainability can also be determined, and the results were presented in Eq.35.

$$\begin{array}{cccccc} & PH & CA & LA & Li & F \\ PH & 0.5000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ CA & 0 & 0.5000 & 1.0000 & 0.7122 & 0.8586 \\ LA & 0 & 0 & 0.5000 & 0 & 0.1363 \\ Li & 0 & 0.2878 & 1.0000 & 0.5000 & 0.7120 \\ F & 0 & 0.1414 & 0.8637 & 0.2880 & 0.5000 \end{array} \tag{35}$$

Then, the integrated weak synthetic sustainability performance of each energy storage technology can be determined according to Eq.31. Taking the integrated weak synthetic sustainability performance of CA as an example:

$$WSS_{CA} = \frac{(0+0.5000+1.0000+0.7122+0.8586) + 5/2 - 1}{5(5-1)} = 0.2285 \tag{36}$$

where WSS_{CA} represents the integrated weak synthetic sustainability performance of CA.

The integrated weak synthetic sustainability performance of other energy storages can also be determined, as presented in Table 10.

Table 10: The integrated weak and strong synthetic sustainability performances of these five energy storage technologies

	Pumped hydro	Compressed Air	Lead-Acid	Lithium-ion	Flywheel
Integrated weak synthetic sustainability performance	0.3000	0.2285	0.1068	0.2000	0.1647
Ranking	1	2	5	3	4

Integrated strong synthetic sustainability performance	0.3000	0.2261	0.1000	0.1869	0.1869
Ranking	1	2	4	3	3

According to Figure 2, the strong sustainability index of each energy storage technology is less than zero, and it means that each energy storage technology performs worse than the reservation point for at least one sub-criterion. The data representing the weak sustainability index of PH, LA, Li and F are crisp numbers, but the data representing the weak sustainability index of CA is an interval number. The elements in the probability matrix for the weak synthetic sustainability can also be determined, as presented in Eq.37.

$$\begin{array}{cccccc} & PH & CA & LA & Li & F \\ PH & 0.5000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ CA & 0 & 0.5000 & 1.0000 & 0.7615 & 0.7615 \\ LA & 0 & 0 & 0.5000 & 0 & 0 \\ Li & 0 & 0.2385 & 1.0000 & 0.5000 & 0.5000 \\ F & 0 & 0.2385 & 1.0000 & 0.5000 & 0.5000 \end{array} \tag{37}$$

The integrated strong synthetic sustainability performance of each energy storage technology can then be determined, and the results were also presented in Table 10. Pumped hydro and compressed air have been ranked as the most sustainable two energy storage technologies. Pumped hydro has been recognized as the most sustainable followed by the compressed air by both the integrated weak synthetic sustainability performance and the integrated strong synthetic sustainability performance. However, the sustainability rankings of Lead-Acid, Lithium-ion and Flywheel determined by these two methods are different (see Table 10). The overall picture of sustainability of these five energy storage technologies was presented in Figure 4. The alternatives located in the top right corner are more sustainable than that located in the bottom left corner. Therefore, the priority sequence of these five energy storage technologies from the most sustainable to the least is pumped hydro, compressed air, Lithium-ion, flywheel, and Lead-Acid.

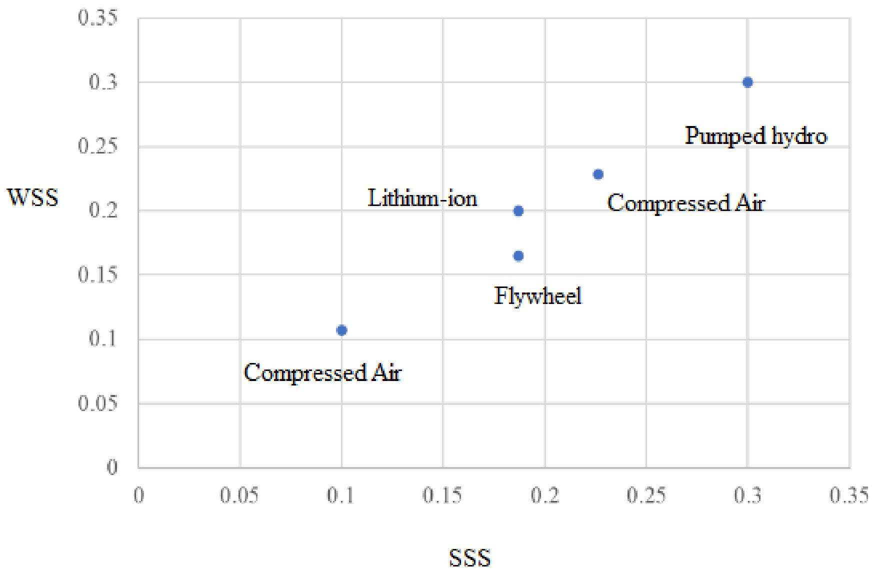

Figure 4: The overall picture of sustainability

In order to validate the effectiveness of the developed method in this study, the results were compared with those determined by the interval multi-attribute decision analysis method（Ren & Ren, 2018. The results especially the rankings of pumped hydro and compressed air (as presented in Table 11) are comparable to those determined by Ren and Ren (2018). To some extent, it reveals that the developed synthetic sustainability index is a valid method for ranking the alternatives according to their sustainability performances.

Table 11: Sustainability rankings

	Pumped hydro	Compressed Air	Lead-Acid	Lithium-ion	Flywheel
Ranking by both the WSS and SSS	1	2	5	3	4
Ranking by Ren and Ren (2018)	1	2	4	5	3

In order to investigate the effects of the weights of the sub-criteria on the rankings of these five energy storage technologies, the following scenarios have been studied by assigning different weights for the criteria or sub-criteria in different scenarios:

(1) Equal weights to the sub-criteria (Scenario 1): an equal weight was assigned to all these ten sub-criteria;

(2) Equal weight to the four criteria (Scenario 2): the global weights of the ten sub-criteria were determined according to the ratios of their local weights; and

(3) A dominant weight was assigned to one criterion and an equal weight was assigned

to other three criteria (Scenarios 3-6): a dominant weight (0.40) was assigned to each of the four criteria (economic, performance, technological and environmental criteria) and an equal weight (0.20) was assigned to all other criteria. The global weights of the ten sub-criteria were determined according to the ratios of their local weights.

The results of sensitivity analysis were presented in Figure 5, and it is apparent that the sustainability sequence of these energy storage technologies highly depends on the weights of the criteria/sub-criteria. In Scenario 1, it is impossible to determine the most sustainable energy storage technology, PH and CA have been ranked as the most sustainable alternatives according to their integrated weak synthetic sustainability performances, but they both perform worse than Li and F according to their integrated strong synthetic sustainability performances. In Scenario 2, PH has been recognized as the most sustainable, followed by F, and CA and Li that are also more sustainable than LA. In Scenario 3, PH has been ranked as the most sustainable, followed by CA and F, Li has been ranked in the fourth position, and LA has been ranked in the last position. In Scenario 4, the sustainability sequence from the most sustainable to the least is PA, F, Li, LA, and CA, respectively. In Scenario 5, the sustainability sequence from the most sustainable to the least is PH, CA, F, Li, and LA, respectively. In Scenario 6, it is impossible to determine the sustainability sequence of these energy storage technologies, PH and F have been ranked in the first and the second position according to their integrated weak synthetic sustainability performances, and F and Li have been ranked in the first and the second position according to their integrated strong synthetic sustainability performance, but it is no doubt that LA has been recognized as the least sustainable comparing with other four energy storage technologies.

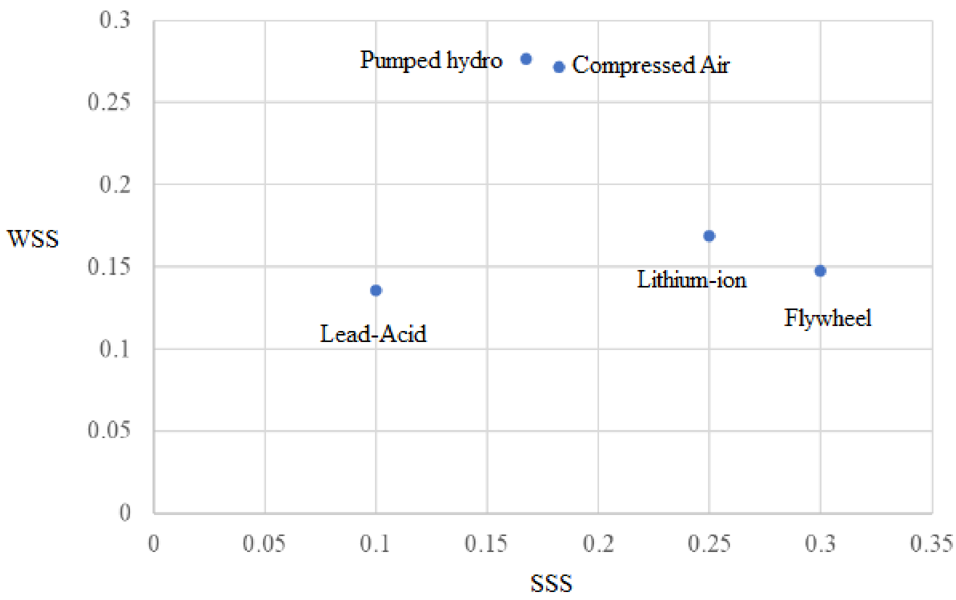

Figure 5 (a): The results of Scenario 1 in sensitivity analysis

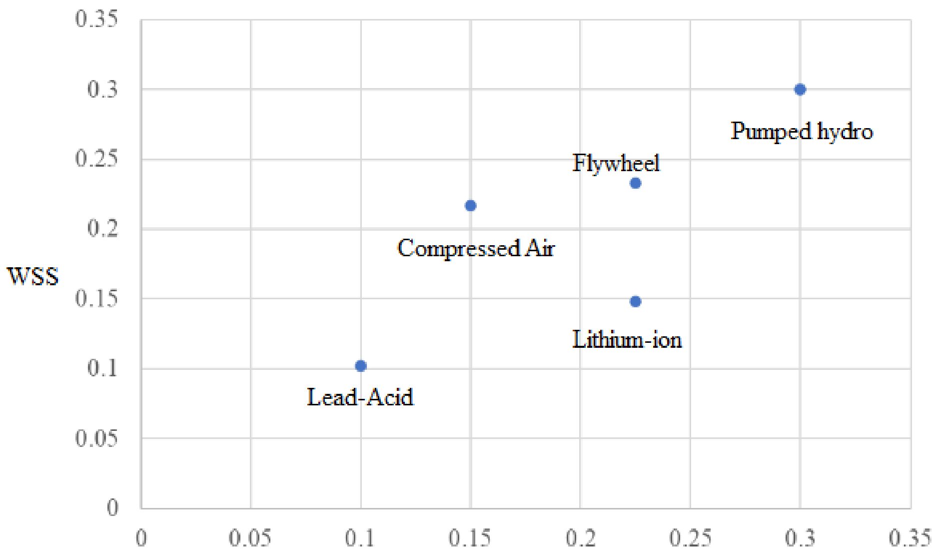

Figure 5 (b): The results of Scenario 2 in sensitivity analysis

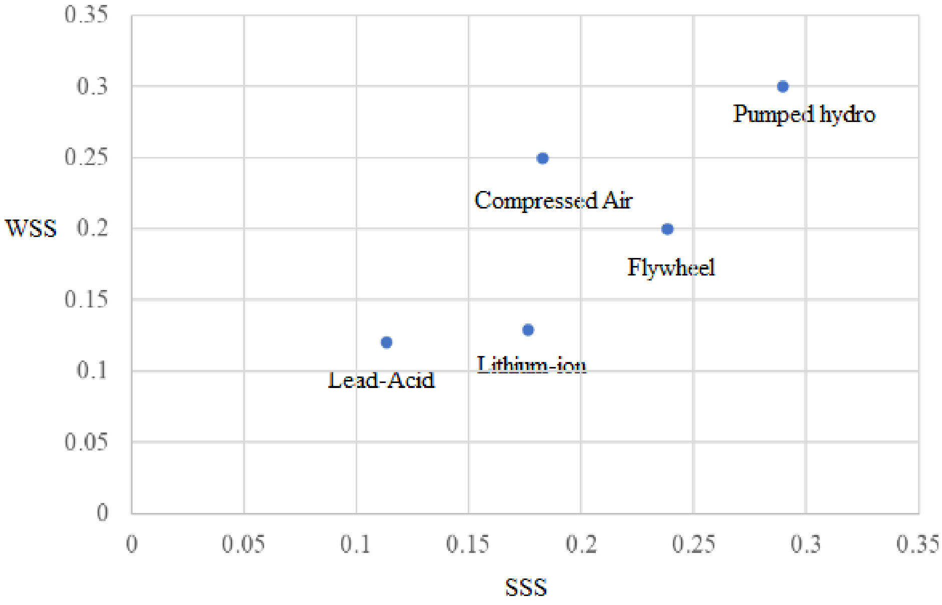

Figure 5 (c): The results of Scenario 3 in sensitivity analysis

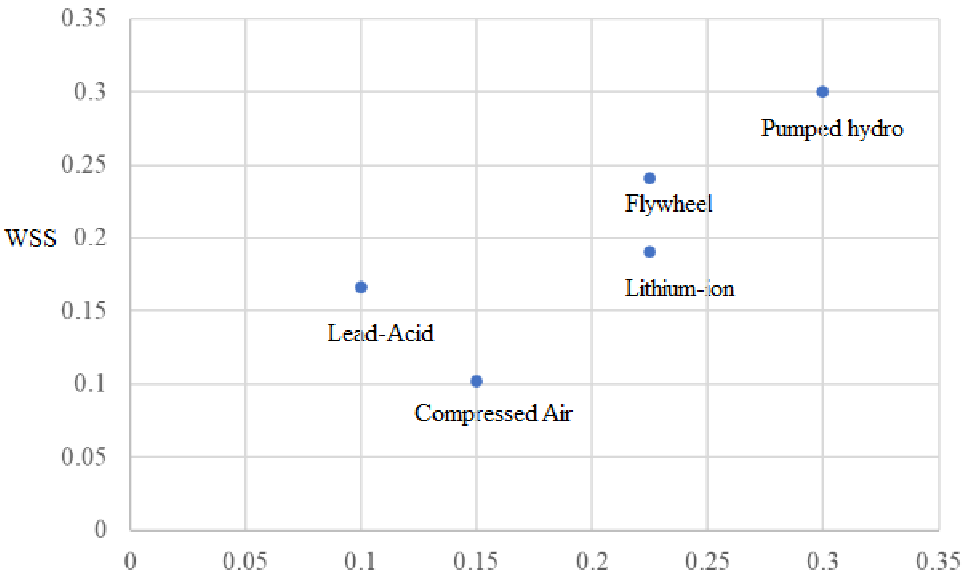

Figure 5 (d): The results of Scenario 4 in sensitivity analysis

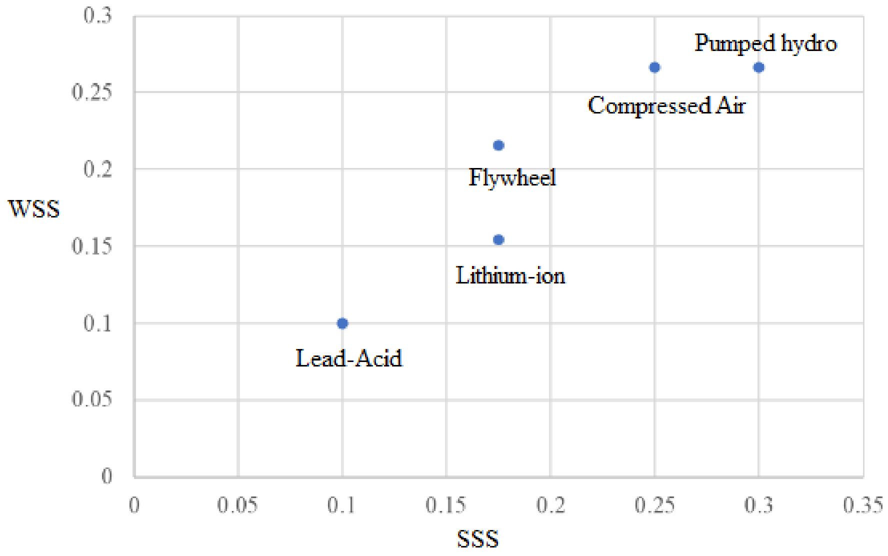

Figure 5 (e): The results of Scenario 5 in sensitivity analysis

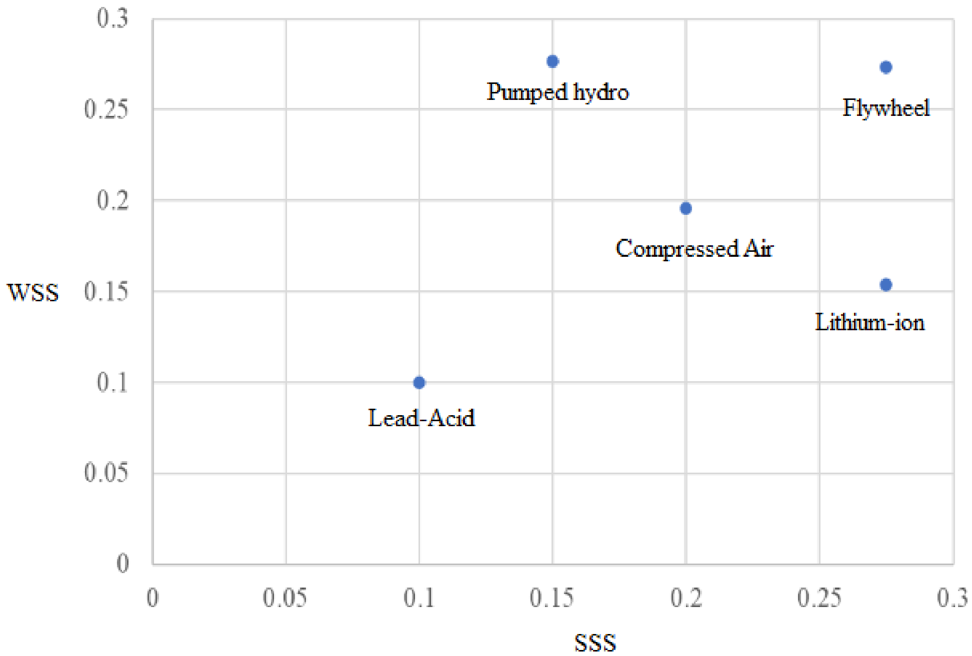

Figure 5 (f): The results of Scenario 6 in sensitivity analysis

Figure 5: The results of Sensitivity analysis

5. Conclusion

This study developed a dual synthetic sustainability index under uncertainties by using the reservation-aspiration reference point scheme, and the users can determine the aspiration point and the reservation point of each criterion/sub-criterion for sustainability assessment. The hierarchy best-worst method was employed to determine the weights of the criteria/sub-criteria for sustainability assessment. Comparing with the NLFP method, the HBMW method employed in this study has the following advantages:

(1) It requires less comparisons comparing with AHP or the methods belonging to some other methods belonging to AHP family;

(2) The weights of the criteria, the local weights of the sub-criteria and the global weights of the sub-criteria can be determined in a programming model simultaneously; and

(3) It is easy for the users to ensure that the judgments for determining the BO and the OW vectors are consistent.

The interval reference point technique which can address the decision-making matrix with both crisp numbers and interval numbers was developed for determining the integrated weak sustainability performance and the integrated strong sustainability performance of each alternative. The users of the interval reference point technique are allowed to set the aspiration point and the reservation point for each sub-criterion for sustainability assessment, and both the strong synthetic sustainability index and the weak synthetic sustainability index can be used to measure the sustainability of each alternative. Comparing with other synthetic or composite sustainability indexes, the developed interval reference point technique has the following advantages:

(1) It can address data uncertainties by using interval numbers. In other words, the data

of the alternatives with respect to the criteria for sustainability assessment can be not only crisp numbers, but also interval numbers.

(2) Both the strong synthetic sustainability performance and the weak synthetic sustainability performance are used for comparing the sustainability of different alternatives, and the synthetic or composite sustainability indexes determined by most of the methods in the previously published works are the weak synthetic sustainability indexes, and they lack the consideration of the strong synthetic sustainability performance.

Besides these advantages, there is also a weak point - the lack of considering the conflict preferences of different stakeholders. The selection of the most sustainable industrial process usually involves different stakeholders (e.g., investors, engineers, governments, and non-governmental organizations), and different stakeholders usually have different preferences, and their preferences and concerns are usually conflict rather than consistent. However, the method developed in this study for determining the dual synthetic sustainability indexes under uncertainties still cannot incorporate the conflict preferences of different stakeholders. In over to overcome this, a game theoretic interval reference point technique which can consider the conflict preferences of different stakeholders will be developed in future by combining the developed interval reference point technique and game theory.

Acknowledgements

The work described in this paper was supported by the Hong Kong Research Grants Council for Early Career Scheme (Grand No. 25208118) and Departmental

General Research Funds (UAFT) of Department of Industrial and Systems Engineers, The Hong Kong Polytechnic University (G-UAFT).

References:

Ashby, M. F., & Polyblank, J. (2012). Materials for Energy Storage Systems—A White Paper.

Amor-Esteban, V., Galindo-Villardon. M.P., Garcia-Sanchez, I.M. (2019). A Multivariate Proposal for a National Corporate Social Responsibility Practices Index (NCSRPI) for International Settings. Social Indicators Research, 143(2), 525-560.

Beaudin, M., Zareipour, H., Schellenberglabe, A., & Rosehart, W. (2010). Energy storage for mitigating the variability of renewable electricity sources: An updated review. Energy for Sustainable Development, 14(4), 302-314.

Bell, S., Morse, S. (2008) Sustainability Indicators Measuring the Immeasurable? Second Edition. ISBN-13: 978-1-84407-299-6 Earthscan, London.

Blancas, F. J., Lozano-Oyola, M., González, M., Caballero, R. (2016). Sustainable tourism composite indicators: A dynamic evaluation to manage changes in sustainability. Journal of Sustainable Tourism, 24(10), 1403-1424.

Blancas, F. J., Caballero, R., González, M., Lozano-Oyola, M., Pérez, F. (2010). Goal programming synthetic indicators: An application for sustainable tourism in Andalusian coastal counties. Ecological Economics, 69(11), 2158-2172.

Blancas, F. J., Lozano-Oyola, M., Gonzales, M., Caballero, R. (2018). A dynamic sustainable tourism evaluation using multiple benchmarks. Journal of Cleaner Production, 174, 1190-1203.

Bolcárová, P., Kološta, S. (2015). Assessment of sustainable development in the EU 27 using aggregated SD index. Ecological Indicators, 48, 699-705.

Cabello, J. M., Navarro-Jurado, E., Rodríguez, B., Thiel-Ellul, D., & Ruiz, F. (2019). Dual weak–strong sustainability synthetic indicators using a double reference point scheme: the case of Andalucía, Spain. Operational Research, 19(3), 757-782.

Cabello, J. M., Navarro, E., Prieto, F., Rodríguez, B., Ruiz, F. (2014). Multicriteria development of synthetic indicators of the environmental profile of the Spanish regions. Ecological indicators, 39, 10-23.

Chen, H., Cong, T. N., Yang, W., Tan, C., Li, Y., & Ding, Y. (2009). Progress in electrical energy storage system: A critical review. Progress in Natural Science,19(3), 291-312.

Díaz-González, F., Sumper, A., Gomis-Bellmunt, O., & Villafáfila-Robles, R. (2012). A review of energy storage technologies for wind power applications.Renewable and Sustainable Energy Reviews, 16(4), 2154-2171.

Dočekalová, M. P., Kocmanová, A. (2016). Composite indicator for measuring corporate sustainability. Ecological Indicators, 61, 612-623.

Dobrovolskienė, N., Tamošiūnienė, R., Banaitis, A., Ferreira, F. A., Banaitienė, N., Taujanskaitė, K., Meidutė-Kavaliauskienė, I. (2019). Developing a composite sustainability index for real estate projects using multiple criteria decision making. Operational Research, 19(3), 617-635.

Dong, F., Mitchell, P.D., Knuteson, D., Wyman, J., Bussan, A.J., Conley, S. (2016). Assessing sustainability and improvements in US Midwestern soybean production systems using a PCA-DEA approach. Renewable Agriculture and Food Systems, 31(6), 524-539.

Evans, A., Strezov, V., & Evans, T. J. (2012). Assessment of utility energy storage options for increased renewable energy penetration. Renewable and Sustainable Energy Reviews, 16(6), 4141-4147.

Escrig-Olmedo, E., Munoz-Torres, M.J., Fernandez-Izquierdo, M.A., Rivera-Lirio, J.M. (2017). Measuring Corporate Environmental Performance: A Methodology for Sustainable Development. Business Strategy and The Environment, 26(2), 142-162.

Gómez-Limón, J. A., Riesgo, L. (2009). Alternative approaches to the construction of a composite indicator of agricultural sustainability: An application to irrigated agriculture in the Duero basin in Spain. Journal of environmental management, 90(11), 3345-3362.

Gomez, T., Gemar, G., Molinos-Senante, M., Sala-Garrido, R., Caballero R. (2017). Assessing the efficiency of wastewater treatment plants: A double bootstrap approach. Journal of Cleaner Production, 164, 315-324.

Gupta, H. (2018). Evaluating service quality of airline industry using hybrid best worst method and VIKOR. Journal of Air Transport Management, 68, 35-47.

Kamali, M., Hewage, K., Milani, A. S. (2018). Life cycle sustainability performance assessment framework for residential modular buildings: Aggregated sustainability indices. Building and Environment, 138, 21-41.

Kheybari, S., Kazemi, M., Rezaei, J. (2019). Bioethanol facility location selection using best-worst method. Applied energy, 242, 612-623.

Kurek, K.A., Heijman, W., van Ophem, J., Gedek, S., Strojny, J. (2020). Dataset for the model of a municipality competitiveness in relation to the geothermal resources exploitation in Poland. Data in brief, 31, 105687.

Lee, C-M, Chou, H-H (2018) Green growth in Taiwan - an application of the OECD green growth monitoring indicators. Singapore Economic Review, 63(2), 249-274.

Li, T., Zhang, H., Yuan, C., Liu, Z., Fan, C. (2012). A PCA-based method for construction of composite sustainability indicators. The International Journal of Life Cycle Assessment, 17(5), 593-603.

Lombardi Netto, A., Salomon, V.A.P., Ortiz-Barrios, M.A., Florek-Paszkowska, A. K., Petrillo, A., De Oliveira, O.J. (2020) Multiple criteria assessment of sustainability programs in the textile industry. International Transactions in Operational Research, DOI: 10.1111/itor.12871.

Lorenzo Linares, H., Betancourt Garcia, M.E., Falcon Rodriguez, M.C. (2019). Sustainability and Local Development: Procedure for Integrated Assessment of Tourist Destinations. Rosa dos Ventos-Turismo e Hospitalidade, 11(1), 84-102.

Lozano-Oyola, M., Blancas, F. J., Gonzales, M., Caballero, R. (2019). Sustainable tourism tags to reward destination management. Journal of Environmental Management, 250, 109458.

Mainali, B., Silveira, S. (2015). Using a sustainability index to assess energy technologies for rural electrification. Renewable and Sustainable Energy Reviews, 41, 1351-1365.

Maiolo, M., Carini, M., Capano, G., Piro, P. (2017). Synthetic sustainability index (SSI) based on life cycle assessment approach of low impact development in the

Mediterranean area. Cogent Engineering, 4(1), 1410272.

Mandinia, I., Habibian, M., Hatamzadeh, Y., Gudmundsson, H. (2018). An indicator-based algorithm to measure transportation sustainability: A case study of the US states. Ecological Indicators, 89, 738-754.

Mikhailov, L. (2003). Deriving priorities from fuzzy pairwise comparison judgements. Fuzzy sets and systems, 134(3), 365-385.

Molinos-Senante, M., Gómez, T., Garrido-Baserba, M., Caballero, R., Sala-Garrido, R. (2014). Assessing the sustainability of small wastewater treatment systems: a composite indicator approach. Science of the total environment, 497, 607-617.

Molinos-Senante, M., Marques, R. C., Pérez, F., Gómez, T., Sala-Garrido, R., Caballero, R. (2016). Assessing the sustainability of water companies: A synthetic indicator approach. Ecological indicators, 61, 577-587.

Molinos-Senante, M., Munoz S., Chamorro, A. (2019) Assessing the quality of service for drinking water supplies in rural settings: A synthetic index approach. Journal of Environmental Management, 247, 613-623.

Munda, G. (2005). "Measuring sustainability": a multi-criterion framework. Environment, Development and Sustainability, 7(1), 117-134.

Nyimbili, P. H., Erden, T. (2020). Comparative evaluation of GIS-based best–worst method (BWM) for emergency facility planning: perspectives from two decision-maker groups. Natural Hazards, 1-37.

Olinto, A. C. (2014). Vector space theory of sustainability assessment of industrial processes. Clean Technologies and Environmental Policy, 16(8), 1815-1820.

Perez, V., Guerrero, F., Gonzalez, M., Perez, F., Caballero, R. (2013). Composite indicator for the assessment of sustainability: The case of Cuban nature-based tourism destinations. Ecological Indicators, 29, 316-324.

Reig‐Martínez, E., Gómez‐Limón, J. A., Picazo‐Tadeo, A. J. (2011). Ranking farms with a composite indicator of sustainability. Agricultural economics, 42(5), 561-575.

Reisi, M., Aye, L., Rajabifard, A., Ngo, T. (2014). Transport sustainability index: Melbourne case study. Ecological Indicators, 43, 288-296.

Ren, J., Toniolo, S. (2018). Life cycle sustainability decision-support framework for ranking of hydrogen production pathways under uncertainties: An interval multi-criteria decision making approach. Journal of Cleaner Production, 175, 222-236.

Ren, J. (2018). Life cycle aggregated sustainability index for the prioritization of industrial systems under data uncertainties. Computers & Chemical Engineering, 113, 253-263.

Ren, J., Ren, X. (2018). Sustainability ranking of energy storage technologies under uncertainties. Journal of cleaner production, 170, 1387-1398.

Rezaei, J. (2015). Best-worst multi-criteria decision-making method. Omega, 53, 49-57.

Rezaei, J., Nispeling, T., Sarkis, J., Tavasszy, L. (2016). A supplier selection life cycle approach integrating traditional and environmental criteria using the best worst method. Journal of Cleaner Production, 135, 577-588.

Ruiz, F., Cabello, J. M., Luque, M. (2011). An application of reference point techniques to the calculation of synthetic sustainability indicators. Journal of the Operational

Research Society, 62(1), 189-197.

Ruiz, F., Cabello, J. M., Pérez-Gladish, B. (2018). Building Ease-of-Doing-Business synthetic indicators using a double reference point approach. Technological Forecasting and Social Change, 131, 130-140.

Ruiz, F., El Gibari, S., Cabello, J.M., Gómez, T. (2020). MRP-WSCI: Multiple reference point based weak and strong composite indicators. Omega, 95, 102060.

Saaty, T.L. The Analytic Hierarchy Process, 3rd ed.; McGraw-Hill: New York, NY, USA, 1980

Saaty, T. L. (2008). Decision making with the analytic hierarchy process. International journal of services sciences, 1(1), 83-98.

Sabiha, N. E., Salim, R., Rahman, S., Rola-Rubzen, M. F. (2016). Measuring environmental sustainability in agriculture: A composite environmental impact index approach. Journal of environmental management, 166, 84-93.

Sikdar, S.K., Sengupta, D., Harten, P. (2012). More on aggregating multiple indicators into a single index for sustainability analyses. Clean Technologies and Environmental Policy, 14(5), 765-773.

Singh, R. K., Murty, H. R., Gupta, S. K., Dikshit, A. K. (2007). Development of composite sustainability performance index for steel industry. Ecological Indicators, 7(3), 565-588.

Soltani, A., Sadiq, R., Hewage, K. (2016). Selecting sustainable waste-to-energy technologies for municipal solid waste treatment: a game theory approach for group decision-making. Journal of Cleaner Production, 113, 388-399.

Stylianou, A., Sdrali D., Apostolopoulos, C.D. (2020). Integrated Sustainability Assessment of Divergent Mediterranean Farming Systems: Cyprus as a Case Study. Sustainability, 12(15), 6105.

Tabatabaei, M. H., Amiri, M., Ghahremanloo, M., Keshavarz-Ghorabaee, M., Zavadskas, E. K., Antucheviciene, J. (2019). Hierarchical Decision-making using a New Mathematical Model based on the Best-worst Method. International Journal of Computers Communications & Control, 14(6), 710-725.

Tasdemir, C., Gazo, R., Quesada, H.J. (2020). Sustainability benchmarking tool (SBT): theoretical and conceptual model proposition of a composite framework. Environment Development and Sustainability, 22(7), 6755-6797.

van de Kaa, G., Kamp, L., Rezaei, J. (2017). Selection of biomass thermochemical conversion technology in the Netherlands: A best worst method approach. Journal of Cleaner Production, 166, 32-39.

Vlontzos, G., Niavis, S., Pardalos, P. (2017) Testing for Environmental Kuznets Curve in the EU Agricultural Sector through an Eco-(in) Efficiency Index. Energies, 10(12), 1992.

Wang, Z., Xu, G., Lin, R., Wang, H., Ren, J. (2019). Energy performance contracting, risk factors, and policy implications: Identification and analysis of risks based on the best-worst network method. Energy, 170, 1-13.

Wierzbicki, A.P. (1980). The use of reference objectives in multiobjective optimization. In: Fandel, G., Gal, T. (eds). Lecture notes in economics and mathematical sysems, vol 177. Springer, Berlin, pp. 468-486.

Wierzbicki, A.P., Makowski, M., Wessels, J. (Eds.) (2000). Model-based decision support methodology with environmental applications. Kluwer Academic Publisher: Dordrecht.

Xavier, A., Freitas, M. D. B. C., Fragoso, R., & do Socorro Rosário, M. (2018). A regional composite indicator for analysing agricultural sustainability in Portugal: A goal programming approach. Ecological Indicators, 89, 84-100.

Yi, H. S., Hau, J. L., Ukidwe, N. U., Bakshi, B. R. (2004). Hierarchical thermodynamic metrics for evaluating the environmental sustainability of industrial processes. Environmental Progress, 23(4), 302-314.

Yi, P., Li, W., Li, L. (2018) Evaluation and Prediction of City Sustainability Using MCDM and Stochastic Simulation Methods. Sustainability, 10(10), 3771.

Xu, Z., Da, Q., 2003. A possibility-based method for priorities of interval judgments matrices. Chin. J. Manag. Sci. 11, 63-65.

Zhang, Y., Yang, Z., Yu, X. (2006) Measurement and evaluation of interactions in complex urban ecosystem. Ecological modelling, 196 (1-2), 77-89.

Zhou, L., Tokos, H., Krajnc, D., Yang, Y. (2012). Sustainability performance evaluation in industry by composite sustainability index. Clean Technologies and Environmental Policy, 14(5), 789-803.

Zinatizadeh, S., Azmi, A., Monavari, S.M., Sobhanardakani, S. (2017), Evaluation and prediction of sustainability of urban areas: A case study for Kermanshah city, Iran. Cities, 66, 1-9.