



# Consistent Construction of Evaluation Threshold Values and Rules for Heterogeneous Linguistic Input Information

Gang Qian<sup>1</sup> · ErZi Zhang<sup>2</sup> · ZhiSong Chen<sup>3</sup> · Radko Mesiar<sup>4</sup> · Ronald R. Yager<sup>5</sup> · LeSheng Jin<sup>3</sup> 

Received: 17 May 2021 / Accepted: 22 July 2021  
© The Author(s) 2021

## Abstract

This study firstly proposes a simpler method for evaluating one certain object's quality with multiple criteria according to some preset evaluation threshold values that are real numbers. In real life, numerous individual valuations are provided with distributional linguistic input information and with multiple criteria, and thus they can become heterogeneous. Against this background, by using OWA weight functions we propose an extended setting and some methods to generate distributional evaluation threshold values which are suitable for the corresponding thresholds-based evaluation method. Some special definitions and formulations are also well provided with necessary analyses and comments. A numerical example of reservoir evaluation and effect are also illustrated.

**Keywords** Decision making · Evaluation · Linguistic information · Preference information · OWA weight function

## 1 Introduction

Numerous evaluation problems are based on quantitative analysis and its related decision taking [1–4]. In most of those problems, to judge the evaluation objects involves

several evaluation criteria rather than a single criterion. Therefore, merging those collected individual information from different criteria into a single one can significantly facilitate the corresponding overall evaluation and decision making.

As commonly known, multi-criteria decision making (MCDM) [5–9] is a widely used evaluation and decision making tool which considers numerical inputs (generally real numbers) and applies an elicited weight vector to perform weighted average (WA) over the inputs, merging them and then yielding a single evaluation value which is still a real number. If there are several evaluation objects (belonging to a same kind) under evaluation and comparison, then by the returned outcomes using MCDM or other evaluation methods, we can select one or some few objects as the desired ones having better properties and outcomes. If there is only one evaluation object under consideration, then in general decision makers may need to judge whether the object is qualified by comparing its evaluation outcome with a predetermined threshold value, which is actually still a decision making problem.

Recall that aggregation operators [10, 11] have been widely applied in evaluation and information fusion, and the related theories have been fast developed during the last few decades [11–15]. A type of well-known aggregation operators used in MCDM include weighted average, geometrical weighted average, and weighted harmonic mean [10], etc.

---

✉ LeSheng Jin  
jls1980@163.com

Gang Qian  
qgmail@vip.sina.com

ErZi Zhang  
601631862@qq.com

ZhiSong Chen  
zhisongchen@nynu.edu.cn

Radko Mesiar  
radko.mesiar@stuba.sk

Ronald R. Yager  
yager@panix.com

<sup>1</sup> Nanjing Audit University, Nanjing, China

<sup>2</sup> Jiangning District Water Affairs Bureau of Nanjing City, Nanjing, China

<sup>3</sup> Business School, Nanjing Normal University, Nanjing, China

<sup>4</sup> Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, 810 05 Bratislava, Slovakia

<sup>5</sup> Machine Intelligence Institute, Iona College, New Rochelle, NY 10801, USA

A common feature in those operators is that they all apply a well determined normalized weight vector (or weight function). Note that with such a weight function but performed in a very different order of inputs positions, Yager proposed another aggregation scheme called ordered weighted averaging (OWA) operators [16], which has been widely applied in considerably many areas [17–19]. It is also noteworthy that the OWA operators can be regarded as a generalization of order statistic (OS) [10] and can also be understood as a special case of the well-known Choquet Integrals with symmetrical capacities (note also that WA operators can be seen as the Choquet Integrals with additive capacities) [10].

All of the above-discussed operators take real numbers as inputs and yield real number output for further decision taking and evaluation. However, in a myriad of evaluation practices, the quantities provided and collected by experts and respondents, and the overall evaluation results for decision makers to take further judgment, are expressed by linguistic information. Researchers studied and proposed some linguistic information and aggregation models [20–22]. The linguistic information considered in this study is a very common one which is much easier to be obtained and collected via different approaches such as by direct inquiring with experts or customers, internet questionnaires, and group meeting and voting.

Put simply, in real evaluation and decision making problems, sometimes the collected input data are not always real values which can be easily transformed or standardized into unit interval  $[0, 1]$ . For example, some individual evaluations are usually obtained by a familiar linguistic evaluation based on a linearly ordered set  $H^{(r)} = (\{1, \dots, r\}, \leq)$  (later sometimes we may only consider  $\{1, \dots, r\}$  and neglect its associated order relation  $\leq$  which will not make any confusion arise). Such linearly ordered set can be embodied or realized by some *linguistic term set* such as  $(\{1 \text{ “excellent”, } 2 \text{ “good”, } 3 \text{ “average”, } 4 \text{ “substandard”}\}, \leq)$  in which  $\leq$  indicates a preference relation to show that linguistic evaluation term  $i$  is “better” than term  $j$  whenever  $i < j$ . Another instance of this type of evaluation information is  $(\{1 \text{ “recommended”, } 2 \text{ “satisfied”, } 3 \text{ “unqualified”}\}, \leq)$ .

With one such linguistic term set, an expert, stakeholder or customer of a certain evaluation problem can provide his/her own judgment over some evaluation object (like a product or a type of service). For example, a customer can be inquired about the quality of a product and offer only one linguistic term as his/her evaluation, say, “good” or “average.” Note that such survey can be carried out by involving multiple persons rather than only one single evaluation subject like that customer. Therefore, with simple statistics, a normalized distribution can be naturally obtained over the linguistic term set  $H^{(r)}$ , which, in this study using some conventional way, is expressed by a nonnegative function  $p : H^{(r)} \rightarrow [0, 1]$  satisfying  $\sum_{i=1}^r p(i) = 1$ .

Since the forgoing mentioned linguistic information is more complex in structure than the simple real number, and for different involved criteria there need differently designed linguistic term sets for making surveys, then the handling, judgment and possible merging of them need special techniques. The pervasiveness of such type of linguistic information in real decision making problems and the feasibility for collecting and dealing with such linguistic information, make the corresponding studies important and meaningful. This study will discuss some special and relevant automatic judgment and evaluation techniques which are mainly based on some well-designed partitioning and thresholds determination rules. One of the advantages of this study lies in that it can provide some relatively objective evaluation scheme within the complex and subjectivity permeated comprehensive decisional scenarios. The theoretical value of the study can be also found in the area of computational intelligence and aggregation theories.

The study will also revolve around some concepts and methods of OWA weight functions’ defining and determination. In addition, we need to make a special note that the OWA weight functions are normally closely associated with OWA operators, but in this study they will be independently used in different decision scenarios and no longer be linked with OWA operators for taking corresponding aggregations.

The remainder of this study is organized as follows. Section 2 firstly formulates a simpler partitioning method for real valued inputs and then presents necessary preparations for later discussions. Section 3 elaborates a comprehensive evaluation method using linguistic evaluation thresholds for heterogeneous inputs. Section 4 provides an application in the evaluation of reservoir operation quality and effect. Section 5 concludes and remarks this study.

## 2 Some Preparation for the Evaluation Method with Heterogeneous Evaluation Information

In this section, we firstly formulate a relatively simpler evaluation method using partitioning method with real valued inputs, and then some necessary review, definitions and comments are prepared for the later discussed method for heterogeneous evaluation information.

### 2.1 The Formulation of the Evaluation Method Under Real Valued Inputs

Some terminologies and expressions are fixed in what follows. The normal numerical input information (for evaluation) is defined as a nonnegative bounded real function  $x : \{1, \dots, n\} \rightarrow [0, 1]$  ( $n \geq 2$ ). The collection of all such

nonnegative bounded real functions defined on  $\{1, \dots, n\}$  is conventionally denoted by  $[0, 1]^n$ .

With the above information, we next design an evaluation method based on given threshold values and summarize it into the following procedures. One may observe that only few human interventions are involved, which can provide more objectivity and efficiency in some real decision making and evaluation problems. The method is suitable for the situation where we only need to decide a linguistic evaluation value for one certain object under

evaluation. Nevertheless, it is noteworthy that the following model might become unsuitable for the decision situation where it is needed to compare several alternative evaluation objects and select only one optimal object. This is because the method is mainly based on qualitative evaluation, and thus often two or more objects may have a same evaluation (such as qualified or unqualified). Besides, quantitative evaluations usually are sensitive, and they may not be very suitable for comparing individual evaluation values which cannot be commensurable.

#### *A basic evaluation method using threshold values*

Step 1: For an evaluation object  $A$ , determine all of the  $n$  relevant criteria  $\{C_i\}_{i=1}^n$  over which a comprehensive evaluation will be performed.

Step 2: For each of the  $n$  criteria, collect individual quantitative evaluation information  $\tilde{x}_i$ , corresponding to  $C_i$ .

Step 3: Those  $n$  pieces of collected information  $(\tilde{x}_i)_{i=1}^n$  are required to be commensurable or can be standardized. Without loss of generality, after standardization a nonnegative bounded real function  $x: N^{(n)} \rightarrow [0, 1]$  will be obtained where  $x(i)$  becomes the new individual evaluation information (instead of  $\tilde{x}_i$ ) with respect to  $C_i$ .

Step 4: Assign a normalized weight function  $u: \{1, \dots, n\} \rightarrow [0, 1]$  to  $\{C_i\}_{i=1}^n$ , which can be obtained by different weights determination methods.

Step 5: Using above determined weight function to perform weighted average and obtain a comprehensive final numerical evaluation result  $E(A) = \sum_{i=1}^n u(i)x(i)$ .

Step 6: Select a linguistic term set  $H^{(3)} = (\{1 \text{ "optimal", } 2 \text{ "average", } 3 \text{ "substandard"}\}, \_)$  as the evaluation scale. (In some different decision environments, one may freely and easily extend it to a linguistic term set with higher dimension.)

Step 7: Determine a pair  $(a, b) \in [0, 1]^2$  (with  $a \leq b$ ) as the normal real valued evaluation thresholds. (This pair can be also easily extended to a tuple with multiple elements according to detailed decision requirement. The pair is suggested to be determined by some experienced experts. In general, some larger parameters (such as  $a, b$ ) embody some stricter decision preferences than those smaller parameters can embody.)

Step 8: With the determined evaluation thresholds  $(a, b)$ , some final and overall linguistic judgment can be made to the evaluation object by the following rules:

- if  $E(A) \geq b$ , then the final linguistic judgment is 1 "optimal";
- if  $E(A) \leq a$ , then the final linguistic judgment is 3 "substandard";
- else, the final linguistic judgment is 2 "average".

**Remark** The choices of evaluation thresholds  $(a, b)$  may influence the final decision making results. In practices, several different experts can be invited to determine their individual suggested thresholds  $(a_i, b_i)$  and then apply an average of those thresholds.

### 2.2 Some Definitions for Dealing with Heterogeneous Evaluation Information

As mentioned in Introduction, the individual evaluation information collected for some certain criteria can be with the form of a nonnegative function  $p : H^{(r)} \rightarrow [0, 1]$  satisfying  $\sum_{i=1}^r p(i) = 1$ . For different criteria  $\{C_i\}_{i=1}^n$ , it may apply some different linguistic term sets  $H^{(r)} = (\{1, \dots, r\}, \leq)$  with dimension  $r$  varying in  $\{2, 3, \dots\}$ . Since those different linguistic term sets are heterogeneous and non-commensurable, then for better formulation and convenient analysis, we should design a set of strict definitions and concepts for further formulating purpose.

We firstly review, rephrase or redefine some basic concepts relating to OWA weight vectors which will serve as the main ingredients in the discussed methods in this study.

**Definition 2.1** [16] In this study, a weight function defined on the linearly ordered set  $H^{(r)} = (\{1, \dots, r\}, \leq)$ ,  $w^{(r)} : \{1, \dots, r\} \rightarrow [0, 1]$ , is called an OWA weight function with dimension  $r$ . The set of all  $r$ -dimensional OWA weight functions is denoted by  $\mathcal{W}^{(r)}$ .

In this paper, when discussing the domain or range of a function, we do not distinguish a linearly ordered set, say,  $H^{(r)} = (\{1, \dots, r\}, \leq)$  from its underlying set  $\{1, \dots, r\}$ . Due to the linearity structure, Yager’s orness definition is very natural and acceptable to measure the extent of bipolar preference within OWA weight vector in numerous applications.

**Definition 2.2** [16] The orness of any OWA weight function  $w^{(r)}$  is defined as a function  $orness : \mathcal{W}^{(r)} \rightarrow [0, 1]$  such that

$$orness(w^{(r)}) = \sum_{i=1}^r w^{(r)}(i) \cdot \frac{r-i}{r-1} \tag{2.1}$$

Dually, the andness of any OWA weight function  $w^{(r)}$  is defined as a function  $andness : \mathcal{W}^{(r)} \rightarrow [0, 1]$  by

$$andness(w^{(r)}) = \sum_{i=1}^r w^{(r)}(i) \cdot \frac{i-1}{r-1} = 1 - orness(w^{(r)}) \tag{2.2}$$

In many applications, the orness/andness can conveniently embody some bipolar decision preference such as optimism/pessimism. In this study, however, we consider

the bipolar strictness-indifference decision preference which can also be suitably and well embodied by orness/andness. For example, a 3-dimensional OWA weight vector  $w^{(3)} = (0.6, 0.3, 0.1)$  is with  $orness(w^{(3)}) = 0.75$ , representing some strict decision and judgment attitude; a 5-dimensional OWA weight vector  $w^{(5)} = (0.2, 0.2, 0.2, 0.2, 0.2)$  has  $orness(w^{(5)}) = 0.5$ , embodying a neutral attitude to the involved evaluation threshold; and a 4-dimensional OWA weight vector  $w^{(4)} = (0.1, 0.2, 0.3, 0.4)$  is with  $orness(w^{(4)}) = 1/3$ , indicating a slightly indifferent attitude about the related evaluation standard.

Next, we review the standard definitions related to the space of parameterized OWA weight vectors (of dimension  $r$ ) and a partial ordering relation on it.

**Definition 2.3** [13] For any OWA weight function  $w^{(r)} \in \mathcal{W}^{(r)}$ , an associated function  $\vec{w}^{(r)} : \{1, \dots, r\} \rightarrow [0, 1]$  such that  $\vec{w}^{(r)}(i) = \sum_{j=1}^i w^{(r)}(j)$  is called the *accumulation function* of  $w^{(r)}$ .

**Definition 2.4** [13] Given  $w_1^{(r)}, w_2^{(r)} \in \mathcal{W}^{(r)}$  and  $\vec{w}_1^{(r)}, \vec{w}_2^{(r)}$  being their accumulation functions, respectively, if  $\vec{w}_2^{(r)} \leq \vec{w}_1^{(r)}$  (i.e.,  $\vec{w}_2^{(r)}(i) \leq \vec{w}_1^{(r)}(i)$  for all  $i \in \{1, \dots, r\}$ ), then we define  $w_2^{(r)} \leq w_1^{(r)}$ . When  $\vec{w}_2^{(r)} \leq \vec{w}_1^{(r)}$ , and there is at least one  $k \in \{1, \dots, r-1\}$  such that  $\vec{w}_2^{(r)}(k) < \vec{w}_1^{(r)}(k)$ , we write  $w_2^{(r)} < w_1^{(r)}$ . With partial ordering  $\leq$ ,  $(\mathcal{W}^{(r)}, \leq)$  is a complete lattice with top element  $w_{\max}^{(r)}$  (i.e.,  $w_{\max}^{(r)}(1) = 1$ ) and bottom element  $w_{\min}^{(r)}$  (i.e.,  $w_{\min}^{(r)}(r) = 1$ ).

It is not difficult to find that we have an equivalent expression for the orness of OWA weight function  $w^{(r)}$ ; that is,

$$orness(w^{(r)}) = \sum_{i=1}^r w^{(r)}(i) \cdot \frac{r-i}{r-1} = \frac{(\sum_{i=1}^r \vec{w}^{(r)}(i)) - 1}{r-1} \tag{2.3}$$

The above formula directly shows that for any two OWA weight functions  $w_1^{(r)}, w_2^{(r)} \in \mathcal{W}^{(r)}$ , if  $w_2^{(r)} \leq w_1^{(r)}$ , then  $orness(w_2^{(r)}) \leq orness(w_1^{(r)})$ , indicating that the strictness expressed by  $w_2^{(r)}$  is lower than the one expressed by  $w_1^{(r)}$ .

Next, we review the concept of parameterized family of OWA weight functions (of dimension  $r$ ).

**Definition 2.5** [13] Given  $r \in \{2, 3, \dots\}$ , a function cluster  $\mathcal{W}^{(r)} = \{w^{(r;\alpha)}\}_{\alpha \in [0,1]} \subseteq \mathcal{W}^{(r)}$  is called a *parameterized family* of OWA weight functions (of dimension  $r$ ) if the following two conditions hold:

1. for any  $\alpha \in [0, 1]$ ,  $w^{(r;\alpha)} \in \mathcal{W}^{(r)}$  and  $orness(w^{(r;\alpha)}) = \alpha$ ;
2. for any  $0 \leq \alpha_1 < \alpha_2 \leq 1$ ,  $w^{(r;\alpha_1)} < w^{(r;\alpha_2)}$ .

With the accumulation functions  $\vec{w}^{(r;\alpha)}$ , we observe that  $\vec{w}^{(r;\alpha)}(r) = 1$  for any  $\alpha \in [0, 1]$ ; and for any  $i \in \{1, \dots, r - 1\}$ ,  $\vec{w}^{(r;0)}(i) = 0$  and  $\vec{w}^{(r;1)}(i) = 1$ .

We recall two-parameterized family of OWA weight functions. The binomial family  $\mathcal{W}^{(r;B)} = \{w^{(r;\alpha;B)}\}_{\alpha \in [0,1]}$  [17] satisfies

$$w^{(r;\alpha;B)}(i) = \binom{r-1}{i-1} (1-\alpha)^{i-1} \alpha^{r-i} \tag{2.4}$$

The Hurwicz family,  $\mathcal{W}^{(r;H)} = \{w^{(r;\alpha;H)}\}_{\alpha \in [0,1]}$  such that  $w^{(r;\alpha;H)}(1) = \alpha$  and  $w^{(r;\alpha;H)}(r) = 1 - \alpha$ . Note that when  $r = 2$ , there only exists the sole family  $\mathcal{W}^{(2)} = \{w^{(2;\alpha)}\}_{\alpha \in [0,1]} = \{(\alpha, 1 - \alpha)\}_{\alpha \in [0,1]} = \mathcal{W}^{(2)}$ . In the rest of this paper, we will only adopt the Binomial family for use.

Since in the evaluation environment of this study it is needed to handle heterogeneous linguistic input information which will be expressed as several different OWA weight functions with varying dimensions, then we next extend the concept of set of OWA weight functions  $\mathcal{W}^{(r)}$  in Definition 2.1 and propose the extended set of OWA weight functions.

**Definition 2.6** An extended space of OWA weight functions  $\mathcal{W}^{(r)}$  with varying dimensions  $r \in \{2, 3, \dots\}$  is denoted by  $\mathbb{W} = \bigcup_{r=2}^{\infty} \mathcal{W}^{(r)}$ .

With this definition, we will have the following extended inputs information which accommodates heterogeneous linguistic input information and can be also defined by a function.

**Definition 2.7** The extended inputs information (for evaluation in this study) is defined as a function  $x : \{1, \dots, n\} \rightarrow \mathbb{W}$ .

For example, an extended inputs information could be a function  $x : N^{(3)} \rightarrow \mathbb{W}$  such that  $x(1) = (0.6, 0.4)$ ,  $x(2) = (0.4, 0.2, 0.3, 0.1)$  and  $x(3) = (0.3, 0.5, 0.2)$ .

For making the discussion and formulation better and clear, we distinguish and strictly present the following two definitions about linguistic evaluation and distributional linguistic evaluation.

**Definition 2.8**

1. When referring to a sole value on a linguistic term set  $H^{(r)}$ , a 0–1 valued function  $G : H^{(r)} \rightarrow \{0, 1\}$ , satisfying  $\sum_{i=1}^r G(i) = 1$ , is called a linguistic evaluation.

2. When a normalized distribution is obtained on a linguistic term set  $H^{(r)}$ , a normalized distribution  $p : H^{(r)} \rightarrow [0, 1]$  (satisfying  $\sum_{i=1}^r p(i) = 1$ ) is called a distributional linguistic evaluation.

**Remark** In some different decisional scenarios, with the same function  $p : H^{(r)} \rightarrow [0, 1]$  we also call it an OWA weight function without any confusion. Besides,  $p : H^{(2)} \rightarrow [0, 1]$  can be equivalently regarded as a real value  $a \in [0, 1]$  if  $a = p(1)$ .

### 3 Comprehensive Evaluation Using Linguistic Evaluation Thresholds for Heterogeneous Inputs

As we have discussed in the preceding section, a 3-scale linguistic term set  $H^{(3)} = (\{1 \text{ “optimal”, } 2 \text{ “average”, } 3 \text{ “sub-standard”}\}, \leq)$  is taken as the evaluation scale for judging the real evaluation result  $E(A) \in [0, 1]$  of the evaluation object  $A$ . In there an evaluation thresholds  $(a, b) \in [0, 1]^2$  is applied to partition  $[0,1]$  and to decide the linguistic evaluation of  $A$  according to which area  $E(A)$  falls into.

When the single evaluation result or individual evaluation input is a function value of the extended inputs information  $x : N^{(n)} \rightarrow \mathbb{W}$  (i.e.,  $x(i) = (v_1, \dots, v_{r_i})$  is the individual distributional linguistic evaluation for  $A$  with respect to criterion  $C_i$ ), then we need to partition  $\mathcal{W}^{(r_i)}$  into three pairwise disjoint subsets and observe which subset  $x(i)$  falls into. To fulfill this aim, we firstly still need a real-valued evaluation thresholds  $(a, b) \in [0, 1]^2$  (with  $a \leq b$ ) to derive an OWA weight-valued evaluation thresholds  $(w^{(r_i;a;B)}, w^{(r_i;b;B)}) \in (\mathcal{W}^{(r_i;B)})^2$  with  $w^{(r_i;a;B)} \leq w^{(r_i;b;B)}$ . Then, we can categorize  $x(i)$  into the three partitioned subsets by the following rules, similar to what have been put in Step 8 of the basic evaluation method using threshold values near the beginning of Sect. 2:

- If  $w^{(r_i;b;B)} \leq x(i)$ , then the linguistic judgment of  $A$  with respect to  $C_i$  is 1 “optimal”;
- if  $x(i) \leq w^{(r_i;a;B)}$ , then the linguistic judgment is 3 “sub-standard”;
- else, the linguistic judgment is 2 “average”.

A complete set of procedures are organized and proposed in what follows.

An evaluation method using linguistic evaluation thresholds for heterogeneous inputs with three stages

**Stage I** Evaluation object and criteria determination with Data preparing

*Step 1:* Select an evaluation object  $A$ , determine all of the  $n$  relevant criteria  $\{C_i\}_{i=1}^n$  over which a comprehensive evaluation will be carried out.

*Step 2:* Assign a normalized weight function  $u: \{1, \dots, n\} \rightarrow [0, 1]$  to  $\{C_i\}_{i=1}^n$ , which can be obtained by different weights determination methods or offered by experienced experts.

*Step 3:* Collect the heterogeneous linguistic input information  $x: \{1, \dots, n\} \rightarrow \mathbb{W}$ . Each individual input  $x(i): (\{1, \dots, r_i\}, \preceq) \rightarrow [0, 1]$  corresponds to a criterion  $C_i$  and is an OWA weight function having the dimension which can be different from the other individual inputs and corresponds to different linguistic term sets  $H_i^{(r_i)}$ . By simple statistics,  $x(i) = (p_1, \dots, p_{r_i})$  can be obtained by inviting a group of experts or customers to evaluate which linguistic evaluation scale in  $H_i^{(r_i)}$  is suitable for  $A$  with respect to  $C_i$ , and  $p_t$  ( $t \in \{1, \dots, r_i\}$ ) represents 100  $p_t$  % of the responses suggest the linguistic evaluation  $t \in \{1, \dots, r_i\}$ .

**Stage II** Linguistic evaluation thresholds defining and generating

*Step 4:* Select a 3-scale linguistic term  $H^{(3)} = (\{1 \text{ "optimal", } 2 \text{ "average", } 3 \text{ "substandard"}\}, \_)$ , as the evaluation scale for the decision maker.

*Step 5:* Determine a pair  $(a, b) \in [0, 1]^2$  (with  $a \leq b$ ) as the real valued evaluation thresholds.

The pair is suggested to be determined by some experienced experts. The parameter  $a$  represents the real valued threshold of dropping to the linguistic evaluation 3 "substandard", and parameter  $b$  represents the real valued threshold of attaining 1 "optimal". In general, some larger  $a, b$  embody some stricter decision preferences than those smaller parameters can embody.

*Step 6:* Apply the real pair  $(a, b)$  as a parameter and use Binomial OWA family to derive an OWA weight valued evaluation thresholds  $(w^{(3;a;B)}, w^{(3;b;B)}) \in \binom{(3;B)}{2}$  with  $w^{(3;a;B)} \_ w^{(3;b;B)}$  for the outer layer evaluation.

*Step 7:* Apply the same real pair  $(a, b)$  as the parameter and still use Binomial OWA family to derive a group of  $n$  OWA weight valued evaluation thresholds  $\{(w_i^{(r_i;a;B)}, w_i^{(r_i;b;B)})\}_{i=1}^n$  with each  $(w_i^{(r_i;a;B)}, w_i^{(r_i;b;B)}) \in \binom{(r_i;B)}{2}$  and  $w_i^{(r_i;a;B)} \_ w_i^{(r_i;b;B)}$  for the inner layers evaluation.

**Stage III** A two-layer linguistic evaluating process

*Step 8:* Perform the inner layers single criterion linguistic judgment. For each  $i \in \{1, \dots, n\}$  and criterion  $C_i$ , judge which linguistic evaluation in  $H^{(3)}$  is suitable for evaluating  $A$  using a function

$G: \{1, \dots, n\} \rightarrow H^{(3)}$  such that:

if  $w_i^{(r_i; b; B)} \preceq x(r_i)$ , then  $G(i) = 1$  (i.e., the linguistic judgment of  $A$  with respect to  $C_i$  is 1 “optimal”);

if  $x(r_i) \preceq w_i^{(r_i; a; B)}$ , then  $G(i) = 3$  (i.e., the linguistic judgment is 3 “substandard”);

else,  $G(i) = 2$  (i.e., the linguistic judgment is 2 “average”).

Step 9: Perform the outer layer linguistic judgment as the final comprehensive evaluation result. With the obtained function  $G$  and the normalized weight function  $u$  assigned to  $\{C_i\}_{i=1}^n$ , we generate an OWA weight function  $v: H^{(3)} \rightarrow [0, 1]$  as a comprehensive distributional linguistic evaluation such that

$$v(j) = \sum_{s \in \{i: G(i)=j\}} u(s) \tag{3.1}$$

Step 10: With the obtained distributional linguistic evaluation  $v$  in Step 9 and the OWA weight valued evaluation thresholds  $(w^{(3; a; B)}, w^{(3; b; B)})$  in Step 6, obtain a final comprehensive linguistic evaluation for  $A$ ,  $E(A) \in H^{(3)}$ , by the following rules:

If  $w^{(3; b; B)} \preceq v$ , then  $E(A) = 1$  (i.e., the final comprehensive linguistic evaluation of  $A$  is 1 “optimal”);

if  $v \preceq w^{(3; a; B)}$ , then  $E(A) = 3$  (i.e., the comprehensive linguistic judgment is 3 “substandard”);

else,  $E(A) = 2$  (i.e., the comprehensive linguistic judgment is 2 “average”).

**Remark 3.1** The OWA weight-valued evaluation thresholds used in the above procedures all correspond to a 3-scale linguistic term  $H^{(3)} = (\{1 \text{ “substandard”, } 2 \text{ “average”, } 3 \text{ “optimal”}\}, \preceq)$ , which can be extended to some terms with more scales if wanted. The 3-scale linguistic term set we used is practical and workable in application because it has a good affinity and close relation to intuitionistic fuzzy sets [23] which is commonly applied in numerous applications [24]. In addition, to adopt the linguistic term set with relatively lower dimension could present a clearer illustration for practitioners to understand the proposed evaluation problem.

### 4 An application in Reservoir Operation Quality and Effect Evaluation

This section provides a numerical case of the proposed evaluation model in the evaluation of reservoir operation and effect. Reservoir operation is related to the water

resource management and evaluation [25] and is important in many aspects of social development and environment conservation.

We will evaluate a certain reservoir  $A$  in the southern area of Nanjing about its operation quality and effect and will adopt a set of four criteria after consulting with some experts working in that reservoir. The evaluation result is useful for further planning and possible improvement or adjustment of that reservoir.

We next elaborate the evaluation procedures proposed in the preceding section. All the initial data and linguistic information are provided by some experts working in that reservoir or studying in water recourse management and planning.

Reservoir operation quality and effect evaluation

**Stage I** Evaluation object and criteria determination with Data preparing

*Step 1:* Denote  $A$  the reservoir under evaluation. The four criteria  $\{C_i\}_{i=1}^4$  are provided by experts and engineers (including two of the coauthors of this study) and listed below.

$C_1$ : Management, construction and hardware

$C_2$ : Flood preventing and drought resisting ability

$C_3$ : Environment production and water quality

$C_4$ : Economic effect

*Step 2:* A normalized weight function  $u: \{1, 2, 3, 4\} \rightarrow [0, 1]$  is provided by some experts and allocated to  $\{C_i\}_{i=1}^4$  with  $u = (0.2, 0.4, 0.3, 0.1)$ .

*Step 3:* The four pieces of collected heterogeneous linguistic input information are evaluated by experts and are defined on the following four different linguistic term sets respectively:

$$H_1^{(2)} = (\{1 \text{ "Qualified"}, 2 \text{ "Not qualified"}\}, \preceq);$$

$$H_2^{(3)} = (\{1 \text{ "Strong"}, 2 \text{ "Average"}, 3 \text{ "Weak"}\}, \preceq);$$

$$H_3^{(4)} = (\{1 \text{ "excellent"}, 2 \text{ "good"}, 3 \text{ "average"}, 4 \text{ "substandard"}\}, \preceq);$$

$$H_4^{(3)} = (\{1 \text{ "Profitable"}, 2 \text{ "break even"}, 3 \text{ "At a loss"}\}, \preceq).$$

And the corresponding four pieces of heterogeneous linguistic input information  $x: \{1, 2, 3, 4\} \rightarrow \mathbb{W}$  are listed in the following:

$$x(1) = (0.8, 0.2); \quad x(2) = (0.9, 0.1, 0); \quad x(3) = (0.7, 0.2, 0.1, 0); \quad x(4) = (0, 1, 0).$$

**Stage II** Linguistic evaluation thresholds defining and generating

*Step 4:* Select a 3-scale linguistic term  $H^{(3)} = (\{1 \text{ "optimal"}, 2 \text{ "average"}, 3 \text{ "substandard"}\}, \preceq)$ , as the evaluation scale.

*Step 5:* Determine a pair  $(0.3, 0.7) \in [0, 1]^2$  as the real valued evaluation thresholds. The pair is suggested by some experienced engineers and experts.

*Step 6:* Apply the real pair  $(0.3, 0.7)$  as a parameter and use Binomial OWA family to derive an OWA weight valued evaluation thresholds  $(w^{(3;0.3;B)}, w^{(3;0.7;B)}) \in (\mathcal{W}^{(3;B)})^2$  with

$$w^{(3;0.3;B)} \preceq w^{(3;0.7;B)} \text{ for the outer layer evaluation. In detail, we obtain } w^{(3;0.3;B)} = (0.09, 0.42, 0.49)$$

$$\text{and } w^{(3;0.7;B)} = (0.49, 0.42, 0.09).$$

*Step 7:* Apply the same real pair (0.3, 0.7) as the parameter and still use Binomial OWA family to derive a group of 4 OWA weight valued evaluation thresholds  $\{(w_i^{(r_i;a;B)}, w_i^{(r_i;b;B)})\}_{i=1}^4$  for the inner layers evaluation. In detail, we obtain

$$w_1^{(2;0.3;B)} = (0.3, 0.7), \quad w_1^{(2;0.7;B)} = (0.7, 0.3);$$

$$w_2^{(3;0.3;B)} = (0.09, 0.42, 0.49), \quad w_2^{(3;0.7;B)} = (0.49, 0.42, 0.09);$$

$$w_3^{(4;0.3;B)} = (0.027, 0.189, 0.441, 0.343), \quad w_3^{(4;0.7;B)} = (0.343, 0.441, 0.189, 0.027);$$

$$w_4^{(3;0.3;B)} = (0.09, 0.42, 0.49), \quad w_4^{(3;0.7;B)} = (0.49, 0.42, 0.09).$$

**Stage III** A two-layer linguistic evaluating process

*Step 8:* Perform the inner layers single criterion linguistic judgment. For each  $i \in \{1, 2, 3, 4\}$  and criterion  $C_i$ , judge which linguistic evaluation in  $H^{(3)}$  is suitable for evaluating  $A$ . In detail, by comparing we have  $G(1) = 1, G(2) = 1, G(3) = 1, G(4) = 2$ .

*Step 9:* Perform the outer layer linguistic judgment as the final comprehensive evaluation result. With the obtained function  $G$  and the normalized weight function  $u = (0.2, 0.4, 0.3, 0.1)$  assigned to  $\{C_i\}_{i=1}^4$ , an OWA weight function  $v: H^{(3)} \rightarrow [0, 1]$  as a distributional linguistic evaluation is generated with  $v = (0.9, 0.1, 0)$ .

*Step 10:* With the obtained distributional linguistic evaluation  $v$  and the OWA weight valued evaluation thresholds  $(w^{(3;0.3;B)}, w^{(3;0.7;B)})$ , obtain a final comprehensive evaluation for  $A$ :  $E(A) = 1$ , indicating the final linguistic evaluation of  $A$  (the reservoir) is “optimal”.

We make some final comments and discussions. With the heterogeneous linguistic input information  $x: \{1, \dots, n\} \rightarrow \mathbb{W}$  and a set of linguistic term set  $\{H_i^{(r_i)}\}_{i=1}^n$  with different numbers of scale, we can indeed devise a set of valuating rules to directly transform  $x$  into a real function just like a piece of real valued input information. For example,  $x(3) = (0.7, 0.2, 0.1, 0)$  can be transformed into a real value by performing a weighted average with a reasonably designed score vector, say,  $q = (1, 0.7, 0.3, 0)$ , and obtain a real value  $y(3) = (0.7)(1) + (0.2)(0.7) + (0.1)(0.3) + (0)(0) = 0.87$ . Then, we can perform some commonly known evaluation methods using the obtained real function  $y$ . However, the score vectors for transforming are not always easy to design. As for the evaluation method using linguistic evaluation thresholds, apart from its clear reasonability and feasibility, another advantage of it lies in that one can easily and flexibly add some more restrictions to the rules deciding the linguistic evaluation results. For example, at some certain decision situations, one may add some more conditions for

achieving a 1 “optimal” evaluation; that is, for example, we can add one more condition “ $(x(r_i))(1) \geq 0.8$ ” to the original condition “ $w_i^{(r_i;b;B)} \leq x(r_i)$ ”. Clearly, the proposed model is more adaptive and flexible than the commonly used evaluation methods based on merging real values and then judging.

## 5 Conclusions

Thresholds-based evaluation is commonly seen, workable and effective in many practical evaluation problems including the one involving in multiple criteria. When the individual inputs corresponding to each evaluation criteria are real numbered, we can be relatively easy to devise some reasonable evaluation schemes and procedures to perform the desired evaluation.

When the individual inputs are provided by heterogeneous linguistic information with distributional forms, the extended space of OWA weight functions was defined to

help strictly devise some partitioning methods for those distributional inputs. The proposed extended space has good adaptivity since it can accommodate different dimensions. In addition, we defined a linguistic evaluation to be a function on linearly ordered linguistic term set, while we also defined a distributional linguistic evaluation to be an OWA weight function. Besides, the distributional information derived can be obtained from statistics and thus may have more objectivity than other linguistic decision making methods.

We adopted a 3-scale linguistic term set mainly for practical purpose and illustrative convenience. In actual, the linguistic term set can be extended into higher dimension according to real needs. By a two-layer evaluation procedures, a linguistic comprehensive evaluation is well built and applied to reservoir evaluation.

**Acknowledgements** We give thanks to reviewers and area editor for their constructive comments which helped us significantly revise our work.

## Declarations

**Conflict of interest** There is no conflict of interest of any form and no animal test.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Jin, L., Mesiar, R., Stupnanova, A., Yager, R.R.: Some generalized integrals applied in scientometrics and related evaluation. *IEEE Trans. Emerg. Top. Comput. Intell.* (2020). <https://doi.org/10.1109/TETCI.2020.3005736>
- Jin, L., Mesiar, R., Yager, R.R.: The properties of crescent preference vectors and its usage in decision making with risk and preference. *Fuzzy Sets Syst.* (2020). <https://doi.org/10.1016/j.fss.2020.06.008>
- Jin, L., Mesiar, R., Yager, R.R.: Parameterized preference aggregation operators with improved adjustability. *Int. J. Gener. Syst.* (2020). <https://doi.org/10.1080/03081079.2020.1786822>
- Kumar, P.S.: Developing a New Approach to Solve Solid Assignment Problems Under Intuitionistic Fuzzy Environment. *International Journal of Fuzzy System Applications* **9**(1), 1–34 (2020)
- Chen, Z., Yu, C., Chin, K., Martínez, L.: An enhanced ordered weighted averaging operators generation algorithm with applications for multicriteria decision making. *Appl. Math. Model.* **71**, 467–490 (2019)
- Hwang, C.L., Yoon, K.: *Multiple Attribute Decision Making*. Springer, Berlin (1981)
- Keeney, R.L., Raiffa, H.: *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Wiley, New York (1976)
- Pedrycz, W., Chen, S.M.: *Granular Computing and Decision-Making: Interactive and Iterative Approaches*. Springer, Heidelberg (2015)
- Saaty, T.L.: Axiomatic foundation of the analytic hierarchy process. *Manag. Sci.* **32**(7), 841–855 (1986)
- Grabisch, M., Marichal, J.L., Mesiar, R., Pap, E.: *Aggregation Functions*. Cambridge University Press, Cambridge (2009). (ISBN:1107013429)
- Klement, E.P., Mesiar, R., Pap, E.: *Triangular Norms*. Springer, Dordrecht (2000)
- Almahasneh, R., Tüü-Szabó, B., Kóczy, L.T., Földesi, P.: Optimization of the time-dependent traveling salesman problem using interval-valued intuitionistic fuzzy sets. *Axioms* **9**(2), 53 (2020). <https://doi.org/10.3390/axioms9020053>
- Pu, X., Jin, L., Mesiar, R., Yager, R.R.: Continuous Parameterized families of RIM quantifiers and quasi-preference with some properties. *Inf. Sci.* **481**, 24–32 (2019)
- Boczek, M., Hovana, A., Hutník, O., Kaluszka, M.: Hölder-Minkowski type inequality for generalized Sugeno integral. *Fuzzy Sets Syst.* **396**, 51–71 (2020)
- Boczek, M., Hovana, A., Hutník, O., Kaluszka, M.: New monotone measure-based integrals inspired by scientific impact problem. *Eur. J. Oper. Res.* **290**, 346–357 (2021)
- Yager, R.R.: On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Trans. Syst. Man Cybern.* **18**(1), 183–190 (1988)
- Leon, T., Zuccarello, P., Ayala, G., de Ves, E., Domingo, J.: Applying logistic regression to relevance feedback in image retrieval systems. *Pattern Recognit.* **40**, 2621–2632 (2007)
- Llamazares, B.: Choosing OWA operator weights in the field of Social Choice. *Inf. Sci.* **177**, 4745–4756 (2007)
- Ouyang, Y.: Improved minimax disparity model for obtaining OWA operator weights: Issue of multiple solutions. *Inf. Sci.* **320**(1), 101–106 (2015)
- Chen, Z., Yang, Y., Wang, X., Chin, K., Tsui, K.: Fostering linguistic decision-making under uncertainty: a proportional interval type-2 hesitant fuzzy TOPSIS approach based on hamacher aggregation operators and andness optimization models. *Inf. Sci.* **500**, 229–258 (2019)
- Labella, Á., et al.: Computing with comparative linguistic expressions and symbolic translation for decision making: ELICIT information. *IEEE Trans. Fuzzy Syst.* **28**(10), 2510–2522 (2020)
- Dutta, B., et al.: Aggregating interrelated attributes in multi-attribute decision-making with ELICIT information based on Bonferroni mean and its variants. *Int. J. Comput. Intell. Syst.* **12**(2), 1179–1196 (2019)
- Atanassov, K.T.: Intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **20**(1), 87–96 (1986)
- Muhiuddin, G., Al-Kadi, D., Balamurugan, M.: Anti-intuitionistic fuzzy soft a-ideals applied to BCI-algebras. *Axioms* **9**(3), 79 (2020). <https://doi.org/10.3390/axioms9030079>
- Huang, Q., Xu, J.: Rethinking environmental bureaucracies in river chiefs system (RCS) in China: a critical literature study. *Sustainability* **11**(6), 1608 (2019). <https://doi.org/10.3390/su11061608>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.