Knowledge-Based Systems 56 (2014) 253–263

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

# A fuzzy multi-criteria group decision making method for individual research output evaluation with maximum consensus

Zongmin Li [a,b], Merrill Liechty [b], Jiuping Xu [a,*], Benjamin Lev [b]

[a] Uncertainty Decision-Making Laboratory, Sichuan University, Chengdu 610064, PR China
[b] Decision Sciences Department, LeBow College of Business, Drexel University, Philadelphia, PA 19104, USA

ARTICLE INFO

*Article history:*
Received 24 June 2013
Received in revised form 19 November 2013
Accepted 26 November 2013
Available online 4 December 2013

*Keywords:*
Multi-criteria group decision making
Individual research output
Bibliometric measures
Peer review
Group consensus

ABSTRACT

Individual research output (IRO) evaluation is both practically and theoretically important. Current research tends to only consider either bibliometric measures or peer review in IRO evaluation. This paper argues that bibliometric measures and peer review should be applied simultaneously to evaluate IRO. Moreover, in real life situations IRO evaluations are often made by groups and inevitably contain evaluators' subjective judgments. Accordingly, this paper develops a fuzzy multi-criteria group evaluation method which considers objective and subjective evaluations, i.e., bibliometric measures and peer review opinions simultaneously. The goals here are to conquer weighting difficulty and achieve maximum group consensus. This requires determining criteria weights, which we do with an intuitionistic fuzzy weighted averaging operator and then determining evaluator weights, which we do with a fuzzy distance-based method. Thereafter, we use a revised TOPSIS method to aggregate the objective and subjective ratings. A practical case study is used to test the feasibility of the methodology. Finally, we discuss the effectiveness of the proposed method.



## 1. Introduction

As pointed out by Van Raan [28], the fundamental purpose of evaluation is to promote quality, therefore evaluation is without any doubt a necessity. Nowadays individual research output (IRO) evaluation inevitably takes place every time a new professor is appointed or promoted, or a learned society or government body allocates a grant.

The main methods of IRO evaluation can be classified as bibliometric measures (objective) and peer review (subjective). Because bibliometric analyses cannot be usefully applied across the board to all departments in large number of universities [19], peer review has become the principal method of university assessment [20]. Although the objective evaluation approach represented by citation-based models and bibliometric indicators cannot replace the subjective evaluation based on an in-depth peer-review analysis of scientific products, it is helpful to aggregate large quantities of data when peer reviewing becomes difficult to implement. This is further illustrated by the fact that IRO evaluation is a complex multifaceted endeavor [12,32]. Some aspects of IRO, such as research quantity and impact can be measured accurately and easily by bibliometric indicators, but for some other aspects, such as research utility and viability, we must draw support from peer opinions. Therefore, objective and subjective evaluation, i.e., bibliometric measures and peer review should be applied simultaneously to evaluate IRO.

Current research tends to only consider either bibliometric measures or peer review in IRO evaluation. In the bibliometric measures field, many single indicators have been developed, such as the total number of papers published, total number of citations garnered, and the mean number of citations per paper [16]. In particular, the proposal of the *h* index [12] has taken the world of research assessment by storm. On the basis of the *h* index, scientists have proposed several '*h*-type' indicators with the intention of either replacing or complementing the original index. Examples include the *g* index [6], the *AR* index [15], the *R* index [15], and the *h* index weighted by citation impact [7]. New bibliometric indicators are continuously coming forth [3,8]. As for peer review, it is like democracy, it may not be the perfect system, but it is the best we have for the evaluation of quality [28]. Nederhof and Van Raan used practical examples to illustrate how peer review and bibliometric measures are complementary and mutually supportive [21]. Van Raan gave an overview of the potentials and limitations of bibliometric methods and presented practical examples of research performance indicators as well as peer review [28]. There are several studies that have compared bibliometric scores with available peer review judgements [1,9,24].

Practically, IRO evaluations are often made by groups (scientific community or evaluating agency), not only because of the problem complexity but also because of wider implications of the decision in terms of responsibility [36]. As a result, IRO evaluation is in

* Corresponding author. Tel.: +86 28 85418191; fax: +86 28 85415143.
*E-mail address:* xujiuping@scu.edu.cn (J. Xu).

0950-7051/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.knosys.2013.11.018

nature a multi-criteria group decision making (MCGDM) problem. Employing MCGDM methods has been proven to be a very effective technique to increase the level of overall satisfaction for the final decision across the group and particularly in evaluation decision-making [18]. Moreover, IRO evaluation, especially the subjective evaluation part, inevitably contains the evaluators' subjective judgments and preferences. This situation has further amplified the uncertainty of assessments inherent in the IRO evaluation process. Evaluators cannot apply precise numbers to describe their assessments, however, they can utilize linguistic variables according to their professional knowledge and experience. Hence, the concept of fuzzy numbers can be integrated into the multi-criteria group evaluation of IRO. In the recent decade, many attempts have been made to propose compromise solution methods for the MCGDM problems under a fuzzy environment [5,31,35,38]. However, there are still two long standing key issues that are apparently not solved well in MCGDM. First, determining a set of suitable weights for multiple evaluation criteria as well as multiple evaluators is often considered to be a very difficult task. Setting arbitrary weights for each criterion in terms of subjective judgments of decision-makers will add to the subjectivity and reduce the decision accuracy. Second, evolving an effective group consensus out of different judgments from different evaluators is still an unsolved issue in the previous studies [10,36].

In this paper, to overcome the weighting difficulty, a fuzzy distance-based method is developed for determining evaluators' weights. We minimize the sum of Euclidean distances between all pairs of ratings and thus make them achieve maximum consensus. In addition, an intuitionistic fuzzy weighted averaging operator is applied to determine criteria weights, so that the importance of criteria can be determined by experts' intuitionistic description. Considering that the ratings contain both crisp and fuzzy numbers, this paper uses a revised TOPSIS method to aggregate all evaluation results.

MCGDM methods for IRO evaluation under fuzzy environments is still largely unexplored. This paper considers for the first time objective, subjective and fuzzy evaluation in IRO evaluation. The remaining of this paper is structured as follows: in Section 2, a set of objective and subjective criteria are identified. Section 3 formulates a fuzzy multi-criteria group evaluation process for IRO evaluation which includes the general framework, the determination of weights and the method of aggregation. A case study along with a discussion is presented in Section 4. Section 5 concludes with a summary.

## 2. Identification of criteria

First of all, a set of objective and subjective evaluation criteria have to be identified for IRO evaluation. For different evaluation objectives or different disciplines and scientific communities, different criteria may be employed. However, generally speaking, they all tend to fall into four research output measures categories: *volume, impact, quality, and utility* [11]. There exists well-developed bibliometric indicators for measuring volume and impact, therefore, we choose objective evaluation criteria for volume and impact. For quality and utility, we have to draw support from peer review opinions as there are not any effective bibliometric indicators to measure them, so subjective evaluation criteria need to be used.

### 2.1. Objective evaluation criteria

Volume can be measured by scientific productivity. A natural candidate to measure scientific productivity of a scientist is the number of papers published ($N_p$) [4]. However, when there are multiple authors, $N_p$ gives every author full credit and thus each author's contribution is indiscriminated. This may not be a fair measurement because authors' contributions are often not treated as equal unless there is a statement: "All authors contributed equally to all aspects of this work". Considering that the average number of authors per paper is keeping growing, Xu et al. [32] proposes an improved $N_p$, i.e. $N'_p$ indicator which considers author rank in measuring scientific productivity to address this multiple authorship problem.

According to Xu et al. [32], suppose that an author has $N_p$ papers published. There are $n_a$ authors in the $a$th paper, $1 \leqslant a \leqslant N_p$ and the author is ranked $k_a$. The author takes $N_a$ of all the credits for the $a$th paper:

$$N_a = 2\left(\frac{n_a - k_a + 1}{n_a^2 + n_a}\right). \tag{1}$$

Summing up the proportion of credit of each of the $N_p$ papers, the author's $N'_p$ is

$$N'_p = \sum_{a=1}^{N_p} N_a. \tag{2}$$

We use $N'_p$, i.e., the adjusted number of papers published (denoted as "$C_1$") to measure volume.

Citation analysis has been increasingly used to judge and quantify the importance of scientists and scientific research. Among various citation analysis based indicators, the $h$ index [12] is usually considered as one of the most important indicators of the scientific impact of a scientist in his/her particular field. The $h$ index is now a well-established standard tool for the evaluation of the scientific impact. It is so widely accepted as an bibliometric impact indicator, that it is even used to measure the impact of academic journals [13]. Therefore, we use the $h$ index (denoted as "$C_2$") to measure impact.

### 2.2. Subjective evaluation criteria

Quality and utility are general in concept. A lot of efforts have been made to select appropriate criteria to evaluate research quality and utility. The United Kingdom has developed one of the most advanced research evaluation systems in Europe. The Research Assessment Exercise in the UK includes the assessment of "relevance to the needs of commerce and industry, as well as to the public and voluntary sectors", "the invention and generation of ideas, images, performances", "the use of existing knowledge in experimental development", "substantially improved materials, devices, products and processes", etc. [11]. In the 1990s, the universities in the Netherlands jointly established a new assessment system for research quality. For each discipline one peer review committee of 5–7 members is set up. Four aspects of research quality are considered [28]: (1) Scientific quality in general; (2) Scientific productivity; (3) Scientific, and where appropriate, societal and/or technological relevance; and (4) Long-term viability. In the research evaluation of six economic research groups, Nederhof and Van Raan [21] put forward that peer judgement focuses on specific and mainly cognitive aspects of the research work: next to a "general" impression, scores are given for "progress toward objectives", "quality of analysis", "contribution to methodology", "contribution to the theory", "dissemination", and "value of money". Among these criteria, "scientific productivity" and "dissemination" can be seen as volume and impact respectively, so we have eliminated these two criteria as members of our subjective evaluation criteria since we have classified them as objective evaluation criteria.

Therefore, based on all these practices, we use the following subjective evaluation criteria for quality and utility evaluation:

(1) General impression (denoted as "$C_3$").
(2) Practical and technological relevance (denoted as "$C_4$").
(3) Improvement to materials/technology/processes (denoted as "$C_5$").
(4) Long-term viability (denoted as "$C_6$").
(5) Innovative invention and generation of new methodology and theory (denoted as "$C_7$").
(6) Expand and apply existing knowledge to contribute existing methodology and theory (denoted as "$C_8$").

$C_4$ and $C_5$ are related to utility, $C_7$ and $C_8$ are related to quality, while criteria $C_3$ and $C_6$ are related to both.

## 3. Formulating a fuzzy MCGDM method for IRO evaluation

This section introduces and formulates a fuzzy multi-criteria group evaluation process for IRO evaluation, including the general framework, determination of weights and the method of aggregation.

### 3.1. General framework

The general framework for fuzzy multi-criteria group evaluation of IRO methodology is shown in Fig. 1.

As shown in Fig. 1, there are $p$ evaluators, i.e., $E_i(i = 1, 2, \ldots, p), m$ criteria, i.e., $C_j(j = 1, 2, \ldots, m)$, and $n$ scientists to be evaluated, i.e., $A_k(k = 1, 2, \ldots, n)$. This paper considers two hierarchies, i.e., the hierarchy of evaluators (peers) and the hierarchy of criteria which are necessary for MCGDM as pointed out by Ma et al. [18]. A fuzzy distance-based method and an intuitionistic fuzzy weighted averaging operator are applied respectively for determination of evaluator and criteria weights. After each evaluator gives the ratings of each scientist's IRO for each criteria, a revised TOPSIS aggregation method is used to get the final evaluation result.

### 3.2. A fuzzy distance-based method for evaluator weight determination

In the process of IRO group evaluation, a group of evaluators (i.e., members of a scientific community or evaluation agency) are involved in a complex decision process and each of them plays a different role. Usually, the weights of evaluators differ because of their position, prestige, experience and scientific insight, etc. Previous studies tend to give arbitrary weights or linguistic variables to generate fuzzy weights of evaluators [38]. However, sometimes it is difficult to judge evaluators' importance in IRO evaluations. A practical situation that occurs frequently for example, is when one evaluator is high in position but is relatively low in experience, and another evaluator is high in experience and prestige, but low in position. Moreover, if all evaluators are almost equal in professional experience and scientific profile, giving each evaluator an equal weight is not necessarily the best solution. As discussed above, group consensus is an important indication of group agreement or reliability. Lack of satisfactory consensus during IRO evaluation will directly lead to disagreement towards decisions regarding personnel selection, promotion and awarding of grants, etc. In order to fully reflect the real behavior of IRO group evaluation, a final decision should be made at a significant level of consensus [5]. Therefore, this paper develops a fuzzy distance-based method for determining evaluators' weights ($W^E = \{w_i^E, i = 1, \ldots, p\}$) to achieve maximum consensus between all evaluators. Our general idea is to minimize the sum of the Euclidean distance from one evaluator's result to another.

Evaluators first make their own judgments of a scientists' IRO based on subjective evaluation criteria $C_3$–$C_8$. Ratings under subjective evaluation criteria are considered as linguistic variables. A linguistic variable is a variable whose value is a natural language phrase. It is very useful in dealing with situations which are ill-defined to be described properly in conventional quantitative expressions [38]. Scientists' research output performance under each subjective evaluation criteria can be expressed on a 7-point rating scale: "very good", "good", "medium good", "fair", "medium poor", "poor", and "very poor". Such linguistic variables are converted into triangular intuitionistic fuzzy numbers (IFNs) [2,38] as shown in Table 1. IFNs are commonly used for solving decision-making problems, where the available information is imprecise. There are different shapes or forms of IFNs, among those, trapezoidal IFNs and triangular IFNs are the most commonly used. For example, Shaw and Roy [25] used trapezoidal IFNs for analysing fuzzy system reliability, while Vahdani et al. [29] applied triangular IFNs to fuzzy group decision-making problems with an application to the contractor selection. This paper chooses to use the triangular IFNs because of their conceptual and computational simplicities.

**Table 1**
Linguistic variables and triangular IFNs for ratings under the subjective evaluation criteria.

| Linguistic variables | Triangular IFNs |
|---|---|
| Very good (VG) | (0.9,1.0,1.0) |
| Good (G) | (0.7,0.9,1.0) |
| Medium good (MG) | (0.5,0.7,0.9) |
| Fair (F) | (0.3,0.5,0.7) |
| Medium poor (MP) | (0.1,0.3,0.5) |
| Poor (P) | (0.0,0.1,0.3) |
| Very poor (VP) | (0.0,0.0,0.1) |

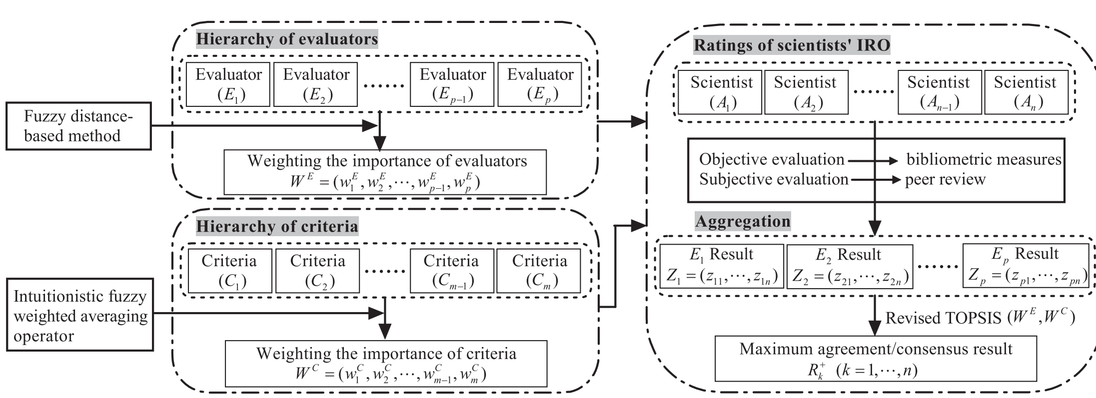


**Fig. 1.** The general framework for the fuzzy MCGDM methodology of IRO evaluation.

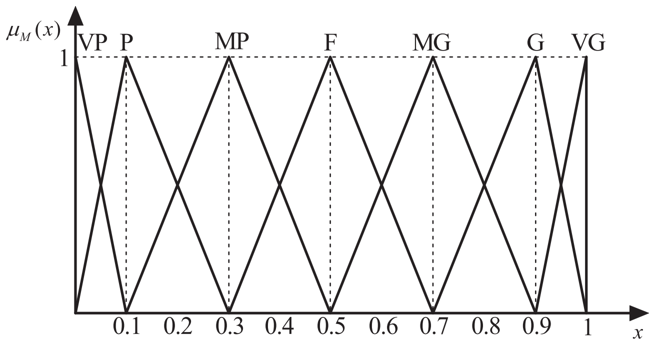


**Fig. 2.** Figure for membership functions of linguistic variables of ratings under the subjective evaluation criteria.

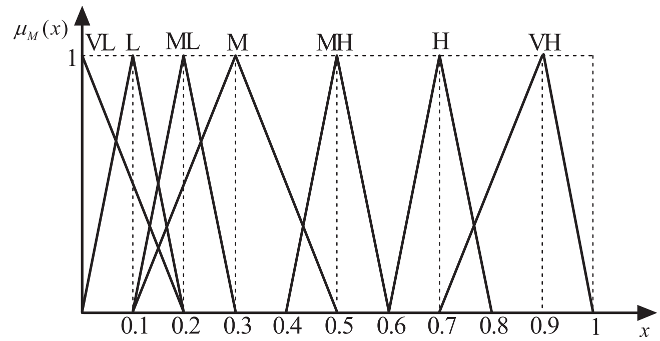


**Fig. 3.** Figure for membership functions of linguistic variables for ratings the importance.

The advantages of employing triangular IFNs in fuzzy modeling and interpreting have been well-justified for MCDM problems [29]. According to Zadeh [37], a fuzzy set $M$ in $X = \{x\}$ is given by $M = \{< x, \mu_M(x) > | x \in X\}$, where $\mu_M : X \to [0,1]$ is the membership function of the fuzzy set $M$; $\mu_M(x) \in [0,1]$ is the membership of $x \in X$ in $M$. The figure for membership functions of linguistic variables for ratings under the subjective evaluation criteria are shown in Fig. 2.

Evaluators make judgments on the importance of all criteria. Similarly, ratings of criteria importance are expressed as triangular IFNs as shown in Table 2 and their membership functions are shown in Fig. 3.

According to [27], for two intuitionistic fuzzy sets $A$ and $B$ in $X = \{x_d, d = 1, \ldots, D\}$, their Euclidean distance is equal to:

$$eIFS(A,B) = \sqrt{\sum_{d=1}^{D}\left[(\mu_A(x_d) - \mu_B(x_d))^2 + (\nu_A(x_d) - \nu_B(x_d))^2 + (\pi_A(x_d) - \pi_B(x_d))^2\right]} \tag{3}$$

where $A = (\mu_A(x_d), \nu_A(x_d), \pi_A(x_d))$, $B = (\mu_B(x_d), \nu_B(x_d), \pi_B(x_d))$.

Let $z_{ikj} = (\mu_{ikj}, \nu_{ikj}, \pi_{ikj})(i = 1,2,\ldots,p; j = 3,\ldots,8; k = 1,2,\ldots,n)$ denote the ratings of all scientists under the subjective evaluation criteria, and $z'_{ij} = (\mu'_{ij}, \nu'_{ij}, \pi'_{ij})(i = 1,2,\ldots,p; j = 1,\ldots,8)$ denote the importance rating of all criteria. These ratings with evaluator weights can be expressed as $\bar{z}_{ikj} = w_i^E z_{ikj} = (w_i^E \mu_{ikj}, w_i^E \nu_{ikj}, w_i^E \pi_{ikj})$ and $\bar{z}'_{ij} = w_i^E z'_{ij} = (w_i^E \mu'_{ij}, w_i^E \nu'_{ij}, w_i^E \pi'_{ij}), (i = 1,2,\ldots,p; j = 3,\ldots,8; k = 1,2,\ldots,n)$. The sum of the Euclidean distance from one evaluator's ratings to another regarding scientists for subjective evaluation criteria can be expressed as:

$$\sum_{i=1}^{p}\sum_{l=1,l\neq i}^{p}\sum_{j=3}^{8}\sum_{k=1}^{n} eIFS(\bar{z}_{ikj}, \bar{z}_{lkj}).$$

Similarly, the sum of the Euclidean distance from one evaluator's ratings to another regarding criteria importance can be expressed as:

$$\sum_{i=1}^{p}\sum_{l=1,l\neq i}^{p}\sum_{j=1}^{8} eIFS(\bar{z}'_{ij}, \bar{z}'_{lj}).$$

To determine the best $w_i^E(i = 1,\ldots,p)$ for the maximum consensus, all ratings with weights of evaluators should move towards one another. This is the principle on the basis of which an aggregated evaluation result is generated. Based on the above analysis, we conduct the optimization model which minimizes the sum of the Euclidean distances between all pairs of evaluation results with weights of evaluators:

$$\min_{w_i^E} D = \sum_{i=1}^{p}\sum_{l=1,l\neq i}^{p}\sum_{j=3}^{8}\sum_{k=1}^{n} eIFS(\bar{z}_{ikj}, \bar{z}_{lkj}) + \sum_{i=1}^{p}\sum_{l=1,l\neq i}^{p}\sum_{j=1}^{8} eIFS(\bar{z}'_{ij}, \bar{z}'_{lj})$$

$$\text{s.t.}\begin{cases} \bar{z}_{ikj} = w_i^E z_{ikj} = (w_i^E \mu_{ikj}, w_i^E \nu_{ikj}, w_i^E \pi_{ikj}), \quad (i = 1,2,\ldots,p; j = 3,\ldots,8; k = 1,2,\ldots,n), \\ \bar{z}_{lkj} = w_i^E z_{lkj} = (w_i^E \mu_{lkj}, w_i^E \nu_{lkj}, w_i^E \pi_{lkj}), \quad (l = 1,2,\ldots,p, l \neq i; j = 3,\ldots,8; k = 1,2,\ldots,n), \\ \bar{z}'_{ij} = w_i^E z'_{ij} = (w_i^E \mu'_{ij}, w_i^E \nu'_{ij}, w_i^E \pi'_{ij}), \quad (i = 1,2,\ldots,p; j = 1,\ldots,8), \\ \bar{z}'_{lj} = w_i^E z'_{lj} = (w_i^E \mu'_{lj}, w_i^E \nu'_{lj}, w_i^E \pi'_{lj}), \quad (l = 1,2,\ldots,p, l \neq i; j = 1,\ldots,8), \\ \sum_{i=1}^{p} w_i^E = 1, \\ w_i^E \geqslant 0, \quad i = 1,\ldots,p. \end{cases} \tag{4}$$

**Table 2**
Linguistic variables and triangular IFNs for rating the importance.

| Linguistic variables | Triangular IFNs |
| --- | --- |
| Very high (VH) | (0.7, 0.9, 1.0) |
| High (H) | (0.6, 0.7, 0.8) |
| Medium high (MH) | (0.4, 0.5, 0.6) |
| Medium (M) | (0.1, 0.3, 0.5) |
| Medium low (ML) | (0.1, 0.2, 0.3) |
| Low (L) | (0.0, 0.1, 0.2) |
| Very low (VL) | (0.0, 0.0, 0.2) |

### *3.3. An intuitionistic fuzzy weighted averaging operator for criteria weight determination*

As discussed in the last section, the criteria importance are rated by evaluators using triangular IFNs, which is expressed by $z'_{ij} = (\mu'_{ij}, \nu'_{ij}, \pi'_{ij})(i = 1,2,\ldots,p; j = 1,\ldots,8)$. In this paper, an intuitionistic fuzzy weighted averaging operator proposed by Xu [33] is utilized to aggregate all evaluator opinions on criteria importance into a group opinion. Considering evaluators' weights $W^E = \{w_i^E, i = 1,\ldots,p\}$ obtained by Eq. (4), the weight of the $j^{th}$ criteria can be obtained by:

$$w_j^C = \frac{\bar{\mu}'_j + \bar{\pi}'_j\left(\frac{\bar{\mu}'_j}{\bar{\mu}'_j+\bar{v}'_j}\right)}{\sum_{j=1}^{8}\left[\bar{\mu}'_j + \bar{\pi}'_j\left(\frac{\bar{\mu}'_j}{\bar{\mu}'_j+\bar{v}'_j}\right)\right]} \tag{5}$$

where

$$\bar{\mu}'_j = \sum_{i=1}^{p} w_i^E \mu'_{ij}, \quad \bar{v}'_j = \sum_{i=1}^{p} w_i^E v'_{ij}, \quad \bar{\pi}'_j = \sum_{i=1}^{p} w_i^E \pi'_{ij}, \text{ and } \sum_{j=1}^{8} w_j^C = 1. \tag{6}$$

From Eqs. (5) and (6), criteria weights can be decided by the linguistic description of the criteria's importance and evaluators' weights.

### 3.4. Aggregation method

A subsequent task is to aggregate evaluation results from different evaluators into an integrated group consensus. Let $X = \Psi(Z_1, Z_2, \ldots, Z_p)$ denote the aggregation of $p$ evaluators' results, where $\Psi(\cdot)$ is an aggregation function. $Z_i (i = 1, \ldots, p)$ is a $n \times 8$ matrix denoting the $i$th evaluator's rating of $n$ scientists's IRO under 8 criteria. There are many aggregation techniques including both linear and nonlinear techniques developed in the MCGDM literature [34]. In this paper we use a common linear additive procedure, so for the subjective evaluation criteria:

$$\tilde{x}_{kj} = \sum_{i=1}^{p} w_i^E Z_i = \sum_{i=1}^{p} w_i^E z_{ikj} = \sum_{i=1}^{p}\left(w_i^E \mu_{ikj}, w_i^E v_{ikj}, w_i^E \pi_{ikj}\right) = (\mu_{kj}, v_{kj}, \pi_{kj})$$
$$\tilde{x}_{kj} \in X, \quad (k = 1, \ldots, n; j = 3, \ldots, 8), \tag{7}$$

where $\tilde{\cdot}$ is a fuzzyness notation. For the objective evaluation criteria $C_1$ and $C_2$, the evaluation results $x_{k1}, x_{k2} (k = 1, 2, \ldots, n)$ are obtained from bibliometric measures which are crisp numbers. Therefore, the fuzzy multi-criteria group evaluation of IRO containing both crisp and fuzzy numbers can be expressed in the matrix format as:

$$X = \begin{bmatrix} x_{11} & x_{12} & \tilde{x}_{13} & \cdots & \tilde{x}_{18} \\ x_{21} & x_{22} & \tilde{x}_{23} & \cdots & \tilde{x}_{28} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \tilde{x}_{n3} & \cdots & \tilde{x}_{n8} \end{bmatrix} \tag{8}$$

To aggregate ratings of scientists for each criterion, we use the TOPSIS concept. Hwang and Yoon [14] presented the technique for order preference by similarity to TOPSIS. TOPSIS takes advantage of the positive-ideal solution (PIS) and the negative-ideal solution (NIS) of multi-attribute problems to rank the plan sets. During the last three decades, many research papers have been published on TOPSIS theories and applications [32,30]. Traditional TOPSIS is only based on crisp evaluation results. In this paper we extent the method to both crisp and fuzzy environment. The procedure of the extended TOPSIS method used in this paper is described as follows:

***Step 1.*** Compute the normalized decision matrix. Vector normalization is applied to calculate $r_{kj}$ and $\tilde{r}_{kj}$.

$$r_{kj} = \frac{x_{kj}}{\sqrt{\sum_{k=1}^{n} x_{kj}^2}}, \quad k = 1, 2, \ldots, n; \quad j = 1, 2, \tag{9}$$

$$\tilde{r}_{kj} = \left(\frac{\mu_{kj}}{d_j^*}, \frac{v_{kj}}{d_j^*}, \frac{\pi_{kj}}{d_j^*}\right), \quad k = 1, 2, \ldots, n; \quad j = 3, \ldots, 8, \tag{10}$$

$$d_j^* = \sqrt{\sum_{k=1}^{n} \pi_{kj}^2}. \tag{11}$$

***Step 2.*** Construct the weighted and normalized evaluation matrix $V$:

$$V = \begin{bmatrix} v_{11} & v_{12} & \tilde{v}_{13} & \cdots & \tilde{v}_{18} \\ v_{21} & v_{22} & \tilde{v}_{23} & \cdots & \tilde{v}_{28} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \tilde{v}_{n3} & \cdots & \tilde{v}_{n8} \end{bmatrix} = \begin{bmatrix} w_1^C r_{11} & w_2^C r_{12} & w_3^C \tilde{r}_{13} & \cdots & w_8^C \tilde{r}_{18} \\ w_1^C r_{21} & w_2^C r_{22} & w_3^C \tilde{r}_{23} & \cdots & w_8^C \tilde{r}_{28} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_1^C r_{n1} & w_2^C r_{n2} & w_3^C \tilde{r}_{n3} & \cdots & w_8^C \tilde{r}_{n8} \end{bmatrix} \tag{12}$$

***Step 3.*** Determine the PIS and NIS. All criteria in this paper are benefit criteria, therefore, the values of PIS ($A^+$) and NIS ($A^-$) are defined as:

$$A^+ = \{\max_k v_{kj} | j = 1, 2, \ldots, 8\} = (v_1^+, v_2^+, \tilde{v}_3^+, \ldots, \tilde{v}_8^+); \tag{13}$$

$$A^- = \{\min_k v_{kj} | j = 1, 2, \ldots, 8\} = (v_1^-, v_2^-, \tilde{v}_3^-, \ldots, \tilde{v}_8^-). \tag{14}$$

$A^+$ and $A^-$ indicate the most and the least preferable IRO of scientists, respectively. For all $k$ and $j = 3, \ldots, 8$, let $\tilde{v}_{kj} = (a_{kj}, b_{kj}, c_{kj})$, $\tilde{v}_j^+ = (a_j^+, b_j^+, c_j^+)$, and $\tilde{v}_j^- = (a_j^-, b_j^-, c_j^-)$ where $a_j^+ = \max_k a_{kj}$, $b_j^+ = \max_k b_{kj}$, $c_j^+ = \max_k c_{kj}$, $a_j^- = \min_k a_{kj}$, $b_j^- = \min_k b_{kj}$, and $c_j^- = \min_k c_{kj}$.

***Step 4.*** Calculate the Euclidean distance. The Euclidean distance between each scientist's IRO and $A^+$ is:

$$S_k^+ = \sqrt{\sum_{j=1}^{j=2}(v_{kj} - v_j^+)^2 + \sum_{j=3}^{j=8}\left[(a_{kj} - a_j^+)^2 + (b_{kj} - b_j^+)^2 + (c_{kj} - c_j^+)^2\right]},$$
$$k = 1, 2, \ldots, n. \tag{15}$$

Similarly, the Euclid distance between each scientist's IRO and $A^-$ is:

$$S_k^- = \sqrt{\sum_{j=1}^{j=2}(v_{kj} - v_j^-)^2 + \sum_{j=3}^{j=8}\left[(a_{kj} - a_j^-)^2 + (b_{kj} - b_j^+)^2 + (a_{kj} - c_j^-)^2\right]},$$
$$k = 1, 2, \ldots, n. \tag{16}$$

***Step 5.*** Calculate the relative closeness of each scientist's IRO to $A^+$.

$$R_k^+ = \frac{S_k^-}{S_k^+ + S_k^-}, \quad 0 < C_k^+ < 1, \quad k = 1, 2, \ldots, n. \tag{17}$$

***Step 6.*** Rank the preference order. By ordering $R_k^+$ in descending order, scientists' IRO can be ranked from the best to the worst. $R_k^+$ can serve as the evaluation score of scientist $A_k$'s research output.

**Table 3**
Related data.

| Scientists | Number of papers published | The adjusted number of papers published | H index |
|---|---|---|---|
| $A_1$ | 60 | 32.63 | 37 |
| $A_2$ | 17 | 12.17 | 12 |
| $A_3$ | 22 | 9.14 | 8 |
| $A_4$ | 10 | 4.00 | 8 |
| $A_5$ | 19 | 6.67 | 6 |
| $A_6$ | 11 | 6.40 | 7 |
| $A_7$ | 7 | 3.83 | 6 |
| $A_8$ | 24 | 10.13 | 3 |
| $A_9$ | 73 | 35.60 | 13 |
| $A_{10}$ | 5 | 2.00 | 1 |
| $A_{11}$ | 2 | 0.67 | 2 |
| $A_{12}$ | 2 | 0.67 | 1 |
| $A_{13}$ | 6 | 2.00 | 2 |
| $A_{14}$ | 3 | 0.83 | 1 |
| $A_{15}$ | 3 | 0.90 | 1 |
| $A_{16}$ | 2 | 0.67 | 2 |
| $A_{17}$ | 10 | 3.15 | 3 |
| $A_{18}$ | 4 | 2.17 | 3 |
| $A_{19}$ | 4 | 2.33 | 1 |
| $A_{20}$ | 3 | 1.30 | 1 |

## 4. Case study

We apply the above described methodology to a practical case to test its feasibility. The publication list and citation data for 20 present members of the Uncertainty Decision-Making Laboratory in Sichuan University and Uncertainty Theory Laboratory in Tsinghua University were collected in 2012 from the Thomson Reuters ISI WoS database. The members include 6 full professors, 5 associate professors and 9 scientists who have been working as senior assistants. Although the database is relatively small, these data represent a sample of researchers from a typical institution, while many other investigations in the literature have concentrated on prominent scientists or rather homogeneous groups of distinguished professors [26,32]. Data are shown in Table 3. Five professors from Drexel university serve as evaluators in this case. They are all active scholars in the uncertainty decision making area and equal in professional profile. It is difficult to tell who is more prestigious and whose opinions are more important in evaluating IRO. The five evaluators were supplemented with 20 scientists' publications and then they gave their ratings under 6 subjective evaluation criteria as shown in Table 4. They also gave ratings about the importance of all 8 criteria as shown in Table 5.

From Tables 4 and 5, we can see that although the five evaluators are similar in scientific background, they still came up with quite different opinions.

### 4.1. Case solution

As discussed above, the evaluator's ratings of the subjective evaluation criteria in Table 4 can be expressed as triangular IFNs $z_{ikj}(i=1,\dots,5;k=1,\dots,20;j=3,\dots,8)$ as shown in Tables 6–11. Similarly, triangular IFNs of evaluator's ratings of criteria importance $\tilde{z}_{ij}(i=1,\dots,5;j=1,\dots,8)$ can be seen in Table 5. According to model (4), we can get evaluators' weights $W^E=(0.197,0.183,0.186,0.204,0.230)$. Because the optimization model minimizes the sum of the Euclidean distances between all pairs of subjective ratings represented by IFNs, the obtained $W^E$ guarantees the maximum agreement/consensus of evaluation results.

According to Eq. (6) and Table 5,

$$\bar{\mu}'_j=\sum_{i=1}^{p}w_i^E\mu'_{ij}=0.197*\begin{pmatrix}0.0\\0.6\\0.1\\0.4\\0.6\\0.6\\0.7\\0.1\end{pmatrix}+0.183*\begin{pmatrix}0.1\\0.6\\0.4\\0.6\\0.6\\0.1\\0.7\\0.0\end{pmatrix}+0.186*\begin{pmatrix}0.1\\0.7\\0.1\\0.1\\0.4\\0.7\\0.6\\0.6\end{pmatrix}+0.204*\begin{pmatrix}0.1\\0.4\\0.6\\0.7\\0.7\\0.4\\0.6\\0.4\end{pmatrix}+0.230*\begin{pmatrix}0.1\\0.6\\0.4\\0.7\\0.6\\0.4\\0.6\\0.4\end{pmatrix}=\begin{pmatrix}0.080\\0.577\\0.325\\0.510\\0.5830.440\\0.637\\0.305\end{pmatrix},$$

$$\bar{\nu}'_j=\sum_{i=1}^{p}w_i^E\nu'_{ij}=0.197*\begin{pmatrix}0.1\\0.7\\0.2\\0.5\\0.7\\0.7\\0.9\\0.3\end{pmatrix}+0.183*\begin{pmatrix}0.2\\0.7\\0.5\\0.7\\0.7\\0.3\\0.9\\0.1\end{pmatrix}+0.186*\begin{pmatrix}0.3\\0.9\\0.3\\0.2\\0.5\\0.9\\0.7\\0.7\end{pmatrix}+0.204*\begin{pmatrix}0.3\\0.5\\0.7\\0.9\\0.9\\0.5\\0.7\\0.5\end{pmatrix}+0.230*\begin{pmatrix}0.3\\0.7\\0.5\\0.9\\0.7\\0.5\\0.7\\0.5\end{pmatrix}=\begin{pmatrix}0.242\\0.696\\0.444\\0.654\\0.703\\0.577\\0.775\\0.424\end{pmatrix},$$

$$\bar{\pi}'_j=\sum_{i=1}^{p}w_i^E\pi'_{ij}=0.197*\begin{pmatrix}0.2\\0.8\\0.3\\0.6\\0.8\\0.8\\1.0\\0.5\end{pmatrix}+0.183*\begin{pmatrix}0.3\\0.8\\0.6\\0.8\\0.8\\0.5\\1.0\\0.2\end{pmatrix}+0.186*\begin{pmatrix}0.5\\1.0\\0.5\\0.3\\0.6\\1.0\\0.8\\0.8\end{pmatrix}+0.204*\begin{pmatrix}0.5\\0.6\\0.8\\1.0\\1.0\\0.6\\0.8\\0.6\end{pmatrix}+0.230*\begin{pmatrix}0.5\\0.8\\0.6\\1.0\\0.8\\0.6\\0.8\\0.6\end{pmatrix}=\begin{pmatrix}0.404\\0.796\\0.562\\0.753\\0.803\\0.695\\0.875\\0.544\end{pmatrix}.$$

Based on Eq. (5), we obtain the weights of criteria $W^C=(0.031,0.163,0.098,0.146,0.164,0.128,0.179,0.091)^T$. Subsequently, by Eqs. (7) and (8) as well as the bibliometric information in Table 3, the evaluation matrix $X$ is obtained:

$$X=\begin{bmatrix}32.63 & 37 & (0.6616,0.8229,0.9225) & \cdots & (0.5486,0.7256,0.8613)\\ 12.17 & 12 & (0.5068,0.6864,0.8456) & \cdots & (0.5654,0.7467,0.891)\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 1.30 & 1 & (0.2517,0.3922,0.5522) & \cdots & (0.4796,0.6613,0.8051)\end{bmatrix}$$

From Eqs. (9)–(12), we get the weighted and normalized evaluation matrix $V$:

$$V = \begin{bmatrix} 0.0191 & 0.1357 & (0.0225, 0.0280, 0.0314) & \cdots & (0.0164, 0.0216, 0.0257) \\ 0.0071 & 0.0440 & (0.0173, 0.0234, 0.0288) & \cdots & (0.0169, 0.0223, 0.0266) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0.0008 & 0.0037 & (0.0086, 0.0134, 0.0188) & \cdots & (0.0143, 0.0197, 0.0240) \end{bmatrix}$$

Therefore, $A^+$ and $A^-$ can be easily obtained:

$$A^+ = \{\max_k v_{kj} | j = 1, 2, \ldots, 8\} = (0.0208, 0.1357, (0.0225, 0.0280, 0.0314), \ldots, (0.0205, 0.0254, 0.0285))^T;$$

$$A^- = \{\min_k v_{kj} | j = 1, 2, \ldots, 8\} = (0.0004, 0.0037, (0.0020, 0.0074, 0.0109), \ldots, (0.0018, 0.0061, 0.0115))^T.$$

**Table 4**
Ratings of scientists under subjective evaluation criteria.

| Criteria | Scientists | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ | Criteria | Scientists | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ | Criteria | Scientists | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_3$ | $A_1$ | F | VG | MG | VG | G | $C_4$ | $A_1$ | VG | G | F | F | MG | $C_5$ | $A_1$ | F | VG | G | G | MG |
| | $A_2$ | MG | MG | MP | VG | MG | | $A_2$ | MG | VG | VG | MG | G | | $A_2$ | F | F | MP | P | P |
| | $A_3$ | G | F | MP | P | F | | $A_3$ | F | MG | VG | F | MP | | $A_3$ | F | MP | MP | P | VP |
| | $A_4$ | P | G | P | F | G | | $A_4$ | F | MG | VG | MG | MP | | $A_4$ | F | VG | MP | G | VG |
| | $A_5$ | VG | MG | VG | VP | MG | | $A_5$ | MG | F | MP | G | MG | | $A_5$ | MP | F | MP | P | VP |
| | $A_6$ | VP | P | VP | VG | VP | | $A_6$ | MP | G | G | MP | F | | $A_6$ | MP | F | VG | G | MG |
| | $A_7$ | VP | G | VG | G | VG | | $A_7$ | VG | G | MP | MG | G | | $A_7$ | P | MP | F | F | F |
| | $A_8$ | F | MP | VP | MP | P | | $A_8$ | MG | MG | VG | MP | MP | | $A_8$ | MP | F | MG | MG | VP |
| | $A_9$ | VP | MG | MP | F | MG | | $A_9$ | F | MG | P | G | MG | | $A_9$ | VP | MP | MP | F | F |
| | $A_{10}$ | MP | MG | MG | MG | VG | | $A_{10}$ | F | MG | F | P | F | | $A_{10}$ | G | MG | F | VG | VG |
| | $A_{11}$ | VG | MG | VP | MP | F | | $A_{11}$ | F | VG | MP | MP | F | | $A_{11}$ | F | MP | MP | F | P |
| | $A_{12}$ | MG | VP | F | MG | MG | | $A_{12}$ | MP | MG | P | MG | VG | | $A_{12}$ | P | F | F | VP | F |
| | $A_{13}$ | MP | MP | P | P | MG | | $A_{13}$ | MP | MG | G | VP | F | | $A_{13}$ | MG | VG | G | G | MP |
| | $A_{14}$ | F | G | MP | MP | MP | | $A_{14}$ | G | MG | F | F | VP | | $A_{14}$ | MP | F | MP | F | P |
| | $A_{15}$ | F | MP | MG | MP | MP | | $A_{15}$ | MP | F | F | VG | MG | | $A_{15}$ | MP | F | F | F | F |
| | $A_{16}$ | MG | VP | VP | F | MP | | $A_{16}$ | G | P | MG | F | MP | | $A_{16}$ | VP | VP | P | P | F |
| | $A_{17}$ | G | VP | MG | VP | G | | $A_{17}$ | F | F | VP | F | P | | $A_{17}$ | F | MG | F | VP | F |
| | $A_{18}$ | F | MG | VP | VP | VP | | $A_{18}$ | MG | MG | F | MP | P | | $A_{18}$ | F | MP | P | F | P |
| | $A_{19}$ | MP | MP | P | MP | P | | $A_{19}$ | G | F | VG | MG | MP | | $A_{19}$ | F | VG | MG | F | VP |
| | $A_{20}$ | G | MG | P | VP | MP | | $A_{20}$ | MG | F | F | F | MP | | $A_{20}$ | VG | G | VP | MG | F |
| $C_6$ | $A_1$ | MP | F | MG | MG | MP | $C_7$ | $A_1$ | MP | G | VG | F | F | $C_8$ | $A_1$ | F | G | MG | F | VG |
| | $A_2$ | F | VG | G | F | MP | | $A_2$ | P | MP | MG | G | MG | | $A_2$ | MG | G | VG | MG | F |
| | $A_3$ | P | MP | F | MG | P | | $A_3$ | F | G | F | F | VP | | $A_3$ | VG | MP | F | MP | P |
| | $A_4$ | MG | G | VG | G | VG | | $A_4$ | P | MP | MP | VP | F | | $A_4$ | P | MG | MG | VG | G |
| | $A_5$ | MP | F | P | VG | MG | | $A_5$ | P | G | VG | MG | VP | | $A_5$ | F | MP | MP | MP | P |
| | $A_6$ | F | F | VG | MG | VP | | $A_6$ | P | P | F | P | VP | | $A_6$ | F | P | MG | G | VG |
| | $A_7$ | MP | VP | VP | F | F | | $A_7$ | MP | G | P | P | F | | $A_7$ | P | MP | MG | G | MP |
| | $A_8$ | F | MP | MP | VP | F | | $A_8$ | F | P | F | VP | F | | $A_8$ | F | F | P | MP | VP |
| | $A_9$ | F | MP | VP | G | MG | | $A_9$ | VG | F | G | MG | MP | | $A_9$ | MP | MG | G | VG | P |
| | $A_{10}$ | P | MP | MP | VP | F | | $A_{10}$ | MP | VG | MG | F | F | | $A_{10}$ | F | P | MG | MG | F |
| | $A_{11}$ | F | P | G | MP | MP | | $A_{11}$ | VP | P | VP | F | MP | | $A_{11}$ | MP | F | MP | P | VP |
| | $A_{12}$ | F | VG | VG | G | MG | | $A_{12}$ | MP | F | F | MP | MG | | $A_{12}$ | MG | VG | VG | G | MG |
| | $A_{13}$ | MP | MG | F | VP | F | | $A_{13}$ | VP | MP | MG | VP | MG | | $A_{13}$ | P | MP | F | MP | F |
| | $A_{14}$ | VG | F | VP | G | VG | | $A_{14}$ | VG | MP | MG | G | VG | | $A_{14}$ | F | MG | F | VP | F |
| | $A_{15}$ | MP | P | P | VP | VP | | $A_{15}$ | MG | P | P | VP | MP | | $A_{15}$ | MG | MP | MG | MP | P |
| | $A_{16}$ | G | MG | F | P | VP | | $A_{16}$ | P | F | F | F | MP | | $A_{16}$ | MP | F | P | F | F |
| | $A_{17}$ | MG | G | MP | VP | F | | $A_{17}$ | F | P | MP | F | P | | $A_{17}$ | VG | VG | G | MG | MP |
| | $A_{18}$ | MG | VG | VG | G | F | | $A_{18}$ | MG | F | F | MP | VP | | $A_{18}$ | P | MP | VP | MP | MP |
| | $A_{19}$ | MG | MP | F | P | VP | | $A_{19}$ | G | VG | VG | F | MP | | $A_{19}$ | F | F | P | MP | P |
| | $A_{20}$ | VG | VG | MP | F | P | | $A_{20}$ | VG | F | MP | VP | MG | | $A_{20}$ | G | VG | MG | F | MP |

**Table 5**
Importance ratings and triangular IFNs of all criteria.

| Criteria | Scientists | | | | |
|---|---|---|---|---|---|
| | $E_1$ $(\mu'_{1j}, \nu'_{1j}, \pi'_{1j})$ | $E_2$ $(\mu'_{2j}, \nu'_{2j}, \pi'_{2j})$ | $E_3$ $(\mu'_{3j}, \nu'_{3j}, \pi'_{3j})$ | $E_4$ $(\mu'_{4j}, \nu'_{4j}, \pi'_{4j})$ | $E_5$ $(\mu'_{5j}, \nu'_{5j}, \pi'_{5j})$ |
| $C_1$ | L (0.0, 0.1, 0.2) | ML (0.1, 0.2, 0.3) | M (0.1, 0.3, 0.5) | M (0.1, 0.3, 0.5) | M (0.1, 0.3, 0.5) |
| $C_2$ | H (0.6, 0.7, 0.8) | H (0.6, 0.7, 0.8) | VH (0.7, 0.9, 1.0) | MH (0.4, 0.5, 0.6) | H (0.6, 0.7, 0.8) |
| $C_3$ | ML (0.1, 0.2, 0.3) | MH (0.4, 0.5, 0.6) | M (0.1, 0.3, 0.5) | H (0.6, 0.7, 0.8) | MH (0.4, 0.5, 0.6) |
| $C_4$ | MH (0.4, 0.5, 0.6) | H (0.6, 0.7, 0.8) | ML (0.1, 0.2, 0.3) | VH (0.7, 0.9, 1.0) | VH (0.7, 0.9, 1.0) |
| $C_5$ | H (0.6, 0.7, 0.8) | H (0.6, 0.7, 0.8) | MH (0.4, 0.5, 0.6) | VH (0.7, 0.9, 1.0) | H (0.6, 0.7, 0.8) |
| $C_6$ | H (0.6, 0.7, 0.8) | M (0.1, 0.3, 0.5) | VH (0.7, 0.9, 1.0) | MH (0.4, 0.5, 0.6) | MH (0.4, 0.5, 0.6) |
| $C_7$ | VH (0.7, 0.9, 1.0) | VH (0.7, 0.9, 1.0) | H (0.6, 0.7, 0.8) | H (0.6, 0.7, 0.8) | H (0.6, 0.7, 0.8) |
| $C_8$ | M (0.1, 0.3, 0.5) | L (0.0, 0.1, 0.2) | H (0.6, 0.7, 0.8) | MH (0.4, 0.5, 0.6) | MH (0.4, 0.5, 0.6) |

**Table 6**
Triangular IFNs of evaluators' ratings under $C_3$.

| Scientists | $z_{1k3}$ | | | $z_{2k3}$ | | | $z_{3k3}$ | | | $z_{4k3}$ | | | $z_{5k3}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_1$ | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 |
| $A_2$ | 0.5 | 0.7 | 0.9 | 0.5 | 0.7 | 0.9 | 0.1 | 0.3 | 0.5 | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 |
| $A_3$ | 0.7 | 0.9 | 1 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 | 0.3 | 0.5 | 0.7 |
| $A_4$ | 0 | 0.1 | 0.3 | 0.7 | 0.9 | 1 | 0 | 0.1 | 0.3 | 0.3 | 0.5 | 0.7 | 0.7 | 0.9 | 1 |
| $A_5$ | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 | 0.9 | 1 | 1 | 0 | 0 | 0.1 | 0.5 | 0.7 | 0.9 |
| $A_6$ | 0 | 0 | 0.1 | 0 | 0.1 | 0.3 | 0 | 0 | 0.1 | 0.9 | 1 | 1 | 0 | 0 | 0.1 |
| $A_7$ | 0 | 0 | 0.1 | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 |
| $A_8$ | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0 | 0 | 0.1 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 |
| $A_9$ | 0 | 0 | 0.1 | 0.5 | 0.7 | 0.9 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.5 | 0.7 | 0.9 |
| $A_{10}$ | 0.1 | 0.3 | 0.5 | 0.5 | 0.7 | 0.9 | 0.5 | 0.7 | 0.9 | 0.5 | 0.7 | 0.9 | 0.9 | 1 | 1 |
| $A_{11}$ | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 | 0 | 0 | 0.1 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 |
| $A_{12}$ | 0.5 | 0.7 | 0.9 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.5 | 0.7 | 0.9 | 0.5 | 0.7 | 0.9 |
| $A_{13}$ | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 |
| $A_{14}$ | 0.3 | 0.5 | 0.7 | 0.7 | 0.9 | 1 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 |
| $A_{15}$ | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0.5 | 0.7 | 0.9 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 |
| $A_{16}$ | 0.5 | 0.7 | 0.9 | 0 | 0 | 0.1 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 |
| $A_{17}$ | 0.7 | 0.9 | 1 | 0 | 0 | 0.1 | 0.5 | 0.7 | 0.9 | 0 | 0 | 0.1 | 0.7 | 0.9 | 1 |
| $A_{18}$ | 0.3 | 0.5 | 0.7 | 0.5 | 0.7 | 0.9 | 0 | 0 | 0.1 | 0 | 0 | 0.1 | 0 | 0 | 0.1 |
| $A_{19}$ | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 |
| $A_{20}$ | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 | 0 | 0.1 | 0.3 | 0 | 0 | 0.1 | 0.1 | 0.3 | 0.5 |

**Table 7**
Triangular IFNs of evaluators' ratings under $C_4$.

| Scientists | $z_{1k4}$ | | | $z_{2k4}$ | | | $z_{3k4}$ | | | $z_{4k4}$ | | | $z_{5k4}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_1$ | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0.5 | 0.7 | 0.9 |
| $A_2$ | 0.5 | 0.7 | 0.9 | 0.9 | 1 | 1 | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 | 0.7 | 0.9 | 1 |
| $A_3$ | 0.3 | 0.5 | 0.7 | 0.5 | 0.7 | 0.9 | 0.9 | 1 | 1 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 |
| $A_4$ | 0.3 | 0.5 | 0.7 | 0.5 | 0.7 | 0.9 | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 | 0.1 | 0.3 | 0.5 |
| $A_5$ | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 |
| $A_6$ | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 0.7 | 0.9 | 1 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 |
| $A_7$ | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 | 0.1 | 0.3 | 0.5 | 0.5 | 0.7 | 0.9 | 0.7 | 0.9 | 1 |
| $A_8$ | 0.5 | 0.7 | 0.9 | 0.5 | 0.7 | 0.9 | 0.9 | 1 | 1 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 |
| $A_9$ | 0.3 | 0.5 | 0.7 | 0.5 | 0.7 | 0.9 | 0 | 0.1 | 0.3 | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 |
| $A_{10}$ | 0.3 | 0.5 | 0.7 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 | 0.3 | 0.5 | 0.7 |
| $A_{11}$ | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 |
| $A_{12}$ | 0.1 | 0.3 | 0.5 | 0.5 | 0.7 | 0.9 | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 0.9 | 1 | 1 |
| $A_{13}$ | 0.1 | 0.3 | 0.5 | 0.5 | 0.7 | 0.9 | 0.7 | 0.9 | 1 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 |
| $A_{14}$ | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0 | 0 | 0.1 |
| $A_{15}$ | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 |
| $A_{16}$ | 0.7 | 0.9 | 1 | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 |
| $A_{17}$ | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 |
| $A_{18}$ | 0.5 | 0.7 | 0.9 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 |
| $A_{19}$ | 0.7 | 0.9 | 1 | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 | 0.1 | 0.3 | 0.5 |
| $A_{20}$ | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 |

**Table 8**
Triangular IFNs of evaluators' ratings under $C_5$.

| Scientists | $z_{1k5}$ | | | $z_{2k5}$ | | | $z_{3k5}$ | | | $z_{4k5}$ | | | $z_{5k5}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_1$ | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 |
| $A_2$ | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 | 0 | 0.1 | 0.3 |
| $A_3$ | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 | 0 | 0 | 0.1 |
| $A_4$ | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 |
| $A_5$ | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 | 0 | 0 | 0.1 |
| $A_6$ | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 |
| $A_7$ | 0 | 0.1 | 0.3 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 |
| $A_8$ | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.5 | 0.7 | 0.9 | 0.5 | 0.7 | 0.9 | 0 | 0 | 0.1 |
| $A_9$ | 0 | 0 | 0.1 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 |
| $A_{10}$ | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 | 0.9 | 1 | 1 |
| $A_{11}$ | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 |
| $A_{12}$ | 0 | 0.1 | 0.3 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 |
| $A_{13}$ | 0.5 | 0.7 | 0.9 | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 | 0.7 | 0.9 | 1 | 0.1 | 0.3 | 0.5 |
| $A_{14}$ | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 |
| $A_{15}$ | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 |
| $A_{16}$ | 0 | 0 | 0.1 | 0 | 0 | 0.1 | 0 | 0.1 | 0.3 | 0 | 0.1 | 0.3 | 0.3 | 0.5 | 0.7 |
| $A_{17}$ | 0.3 | 0.5 | 0.7 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 |
| $A_{18}$ | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 |
| $A_{19}$ | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0 | 0 | 0.1 |
| $A_{20}$ | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 | 0 | 0 | 0.1 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 |

**Table 9**
Triangular IFNs of evaluators' ratings under $C_6$.

| Scientists | $z_{1k6}$ | | | $z_{2k6}$ | | | $z_{3k6}$ | | | $z_{4k6}$ | | | $z_{5k6}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_1$ | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.5 | 0.7 | 0.9 | 0.5 | 0.7 | 0.9 | 0.1 | 0.3 | 0.5 |
| $A_2$ | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 |
| $A_3$ | 0 | 0.1 | 0.3 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.5 | 0.7 | 0.9 | 0 | 0.1 | 0.3 |
| $A_4$ | 0.5 | 0.7 | 0.9 | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 |
| $A_5$ | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 |
| $A_6$ | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 | 0 | 0 | 0.1 |
| $A_7$ | 0.1 | 0.3 | 0.5 | 0 | 0 | 0.1 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 |
| $A_8$ | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 |
| $A_9$ | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0 | 0 | 0.1 | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 |
| $A_{10}$ | 0 | 0.1 | 0.3 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 |
| $A_{11}$ | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 | 0.7 | 0.9 | 1 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 |
| $A_{12}$ | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 |
| $A_{13}$ | 0.1 | 0.3 | 0.5 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 |
| $A_{14}$ | 0.9 | 1 | 1 | 0.3 | 0.5 | 0.7 | 0 | 0 | 0.1 | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 |
| $A_{15}$ | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 | 0 | 0.1 | 0.3 | 0 | 0 | 0.1 | 0 | 0 | 0.1 |
| $A_{16}$ | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 | 0 | 0 | 0.1 |
| $A_{17}$ | 0.5 | 0.7 | 0.9 | 0.7 | 0.9 | 1 | 0.1 | 0.3 | 0.5 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 |
| $A_{18}$ | 0.5 | 0.7 | 0.9 | 0.9 | 1 | 1 | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 | 0.3 | 0.5 | 0.7 |
| $A_{19}$ | 0.5 | 0.7 | 0.9 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 | 0 | 0 | 0.1 |
| $A_{20}$ | 0.9 | 1 | 1 | 0.9 | 1 | 1 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 |

**Table 10**
Triangular IFNs of evaluators' ratings under $C_7$.

| Scientists | $z_{1k7}$ | | | $z_{2k7}$ | | | $z_{3k7}$ | | | $z_{4k7}$ | | | $z_{5k7}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_1$ | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 |
| $A_2$ | 0 | 0.1 | 0.3 | 0.1 | 0.3 | 0.5 | 0.5 | 0.7 | 0.9 | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 |
| $A_3$ | 0.3 | 0.5 | 0.7 | 0.7 | 0.9 | 1 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0 | 0 | 0.1 |
| $A_4$ | 0 | 0.1 | 0.3 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 |
| $A_5$ | 0 | 0.1 | 0.3 | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 | 0 | 0 | 0.1 |
| $A_6$ | 0 | 0.1 | 0.3 | 0 | 0.1 | 0.3 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 | 0 | 0 | 0.1 |
| $A_7$ | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 0 | 0.1 | 0.3 | 0 | 0.1 | 0.3 | 0.3 | 0.5 | 0.7 |
| $A_8$ | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 | 0.3 | 0.5 | 0.7 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 |
| $A_9$ | 0.9 | 1 | 1 | 0.3 | 0.5 | 0.7 | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 | 0.1 | 0.3 | 0.5 |
| $A_{10}$ | 0.1 | 0.3 | 0.5 | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 |
| $A_{11}$ | 0 | 0 | 0.1 | 0 | 0.1 | 0.3 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 |
| $A_{12}$ | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0.5 | 0.7 | 0.9 |
| $A_{13}$ | 0 | 0 | 0.1 | 0.1 | 0.3 | 0.5 | 0.5 | 0.7 | 0.9 | 0 | 0 | 0.1 | 0.5 | 0.7 | 0.9 |
| $A_{14}$ | 0.9 | 1 | 1 | 0.1 | 0.3 | 0.5 | 0.5 | 0.7 | 0.9 | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 |
| $A_{15}$ | 0.5 | 0.7 | 0.9 | 0 | 0.1 | 0.3 | 0 | 0.1 | 0.3 | 0 | 0 | 0.1 | 0.1 | 0.3 | 0.5 |
| $A_{16}$ | 0 | 0.1 | 0.3 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 |
| $A_{17}$ | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 |
| $A_{18}$ | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0 | 0 | 0.1 |
| $A_{19}$ | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 | 0.9 | 1 | 1 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 |
| $A_{20}$ | 0.9 | 1 | 1 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0 | 0 | 0.1 | 0.5 | 0.7 | 0.9 |

**Table 11**
Triangular IFNs of evaluators' ratings under $C_8$.

| Scientists | $z_{1k8}$ | | | $z_{2k8}$ | | | $z_{3k8}$ | | | $z_{4k8}$ | | | $z_{5k8}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_1$ | 0.3 | 0.5 | 0.7 | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0.9 | 1 | 1 |
| $A_2$ | 0.5 | 0.7 | 0.9 | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 |
| $A_3$ | 0.9 | 1 | 1 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 |
| $A_4$ | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 0.5 | 0.7 | 0.9 | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 |
| $A_5$ | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 |
| $A_6$ | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 |
| $A_7$ | 0 | 0.1 | 0.3 | 0.1 | 0.3 | 0.5 | 0.5 | 0.7 | 0.9 | 0.7 | 0.9 | 1 | 0.1 | 0.3 | 0.5 |
| $A_8$ | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 | 0.1 | 0.3 | 0.5 | 0 | 0 | 0.1 |
| $A_9$ | 0.1 | 0.3 | 0.5 | 0.5 | 0.7 | 0.9 | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 | 0 | 0.1 | 0.3 |
| $A_{10}$ | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 |
| $A_{11}$ | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 | 0 | 0 | 0.1 |
| $A_{12}$ | 0.5 | 0.7 | 0.9 | 0.9 | 1 | 1 | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 |
| $A_{13}$ | 0 | 0.1 | 0.3 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 |
| $A_{14}$ | 0.3 | 0.5 | 0.7 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0 | 0 | 0.1 | 0.3 | 0.5 | 0.7 |
| $A_{15}$ | 0.5 | 0.7 | 0.9 | 0.1 | 0.3 | 0.5 | 0.5 | 0.7 | 0.9 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 |
| $A_{16}$ | 0.1 | 0.3 | 0.5 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 |
| $A_{17}$ | 0.9 | 1 | 1 | 0.9 | 1 | 1 | 0.7 | 0.9 | 1 | 0.5 | 0.7 | 0.9 | 0.1 | 0.3 | 0.5 |
| $A_{18}$ | 0 | 0.1 | 0.3 | 0.1 | 0.3 | 0.5 | 0 | 0 | 0.1 | 0.1 | 0.3 | 0.5 | 0.1 | 0.3 | 0.5 |
| $A_{19}$ | 0.3 | 0.5 | 0.7 | 0.3 | 0.5 | 0.7 | 0 | 0.1 | 0.3 | 0.1 | 0.3 | 0.5 | 0 | 0.1 | 0.3 |
| $A_{20}$ | 0.7 | 0.9 | 1 | 0.9 | 1 | 1 | 0.5 | 0.7 | 0.9 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 |

**Table 12**
Evaluation results with different $R_k^+$ and $R_k'^+$ ranks given in bold face.

| Scientists | $S_k^+$ | $S_k^-$ | $R_k^+$ | Rank by $R_k^+$ | Rank by $R_k'^+$ | Rank by $N_p'$ | Rank by $h$ index |
|---|---|---|---|---|---|---|---|
| $A_1$ | 0.0366 | 0.1647 | 0.8181 | 1 | 1 | 2 | 1 |
| $A_2$ | 0.1107 | 0.0890 | 0.4457 | 2 | 2 | 3 | 3 |
| $A_3$ | 0.1369 | 0.0523 | 0.2766 | 14 | 14 | 5 | 4 |
| $A_4$ | 0.1251 | 0.0911 | 0.4212 | 4 | 4 | 8 | 4 |
| $A_5$ | 0.1377 | 0.0636 | 0.3160 | 11 | 11 | 6 | 7 |
| $A_6$ | 0.1392 | 0.0674 | 0.3264 | **10** | **8** | 7 | 6 |
| $A_7$ | 0.1389 | 0.0609 | 0.3048 | **12** | **13** | 9 | 7 |
| $A_8$ | 0.1533 | 0.0408 | 0.2103 | 17 | 17 | 4 | 9 |
| $A_9$ | 0.1092 | 0.0848 | 0.4369 | 3 | 3 | 1 | 2 |
| $A_{10}$ | 0.1469 | 0.0847 | 0.3658 | 5 | 5 | 13 | 15 |
| $A_{11}$ | 0.1616 | 0.0362 | 0.1829 | 20 | 20 | 18 | 12 |
| $A_{12}$ | 0.1482 | 0.0749 | 0.3357 | **7** | **9** | 18 | 15 |
| $A_{13}$ | 0.1501 | 0.0630 | 0.2955 | **13** | **12** | 13 | 12 |
| $A_{14}$ | 0.1475 | 0.0836 | 0.3618 | 6 | 6 | 17 | 15 |
| $A_{15}$ | 0.1639 | 0.0401 | 0.1966 | 18 | 18 | 16 | 15 |
| $A_{16}$ | 0.1608 | 0.0371 | 0.1874 | 19 | 19 | 18 | 12 |
| $A_{17}$ | 0.1505 | 0.0524 | 0.2582 | 16 | 16 | 10 | 9 |
| $A_{18}$ | 0.1524 | 0.0582 | 0.2765 | 15 | 15 | 12 | 9 |
| $A_{19}$ | 0.1502 | 0.0730 | 0.3270 | **9** | **7** | 11 | 15 |
| $A_{20}$ | 0.1451 | 0.0706 | 0.3274 | **8** | **10** | 15 | 15 |

By Eqs. (15)–(17), $S_k^+, S_k^-$ and $R_k^+$ are determined and the 20 scientists' IRO ranked as shown in Table 12.

### 4.2. Discussion

IRO of these 20 scientists are sorted from the best to the worst by the proposed method as shown in Table 12. In this section we discuss the effectiveness of the proposed methodology and compare it with former research.

(1) We recalculate the problem given equal weights to five evaluators, namely $w_i'^E = 0.2, i = 1, \ldots, 5$. In this situation, $W'^C = (0.042, 0.167, 0.095, 0.141, 0.161, 0.127, 0.177, 0.090)$. The rank by $R_k'^+$ column in Table 12 shows the ranks of 20 scientists with evaluators given equal weights. As shown in Table 12, six scientists are ranked differently by $R_k^+$ and $R_k'^+$. The rankings by $R_k^+$ column in Table 12 represents the maximum consensus evaluation results. In this sense, ranking by $R_k^+$ is better than ranking by $R_k'^+$.

(2) Not all scientists' IRO can be ranked by $N_p'$ and the $h$ index because some scientists have the same performance under these indicators. For example, $A_{11}, A_{12}$ and $A_{16}$ are all ranked 18 by $N_p'$, both $A_3$ and $A_4$ are ranked 4 by the $h$ index, and six scientists are ranked 15 by the $h$ index. However, their IRO can be differentiated using the our proposed method. Therefore, the our proposed method has better discrimination performance than the single bibliometric indicators.

(3) It is worth noting that our proposed method comes up with quite different results than the $h$ index, which is one of the most widely used IRO evaluation indicators. For example, $A_3$ is ranked 4 by the $h$ index, but it is ranked 14 by $R_k^+$. This is because although $A_3$ has high performance in impact, impact is only one of eight criteria considered in our proposed method, and it only has a weight of 0.183. Our proposed method considers both objective and subjective evaluations on a total of four aspects, i.e., volume, impact, quality and utility (represented by the eight criteria of IRO). Therefore, the evaluation results from our proposed method are more comprehensive and take more aspects into consideration which can effectively overcome the one-sidedness of a single indicator. Also the proposed method produces quite different results than the indicator $N_p'$. $N_p'$ only measures the research productivity without any consideration of impact or quality. Using a single indicator to measure IRO is unavoidably biased.

(4) There exists some other comprehensive IRO evaluation research. Such as Xu et al. [32] who established an evaluation index system by choosing various bibliometric indicators to comprehensively evaluate IRO. A set of journal evaluation indicators are chose to evaluate papers' quality. However, good journals do not always publish high quality papers. Therefore, employing peer review opinions to evaluate research quality in this paper is more persuasive. Moreover, many IRO research is only based on published research papers, while, the method in this paper has a potential be used in evaluating book chapters, research reports and presentations. Ma et al. [18] established a fuzzy MCGDM process (FMP) model which can aggregate both subjective and objective information under multi-level hierarchies of criteria and evaluators. Based on FMP they further developed a fuzzy MCGDM decision support system (called Decider). The Decider software is more effective in handling information than our proposed method as it can handle information expressed in linguistic terms, boolean values, as well as numeric values to assess and rank a set of alternatives. However, our proposed method is more specific in IRO evaluation as it identifies specific objective and subjective criteria and develops a maximum consensus weighting method for IRO evaluation.

## 5. Conclusion

MCGDM of IRO evaluation is both practically and theoretically important. On the one hand, universities and research departments are faced with demands for greater accountability and the consequences of diminished funding. IRO evaluation today are expected to be both efficient and accountable. These pressures have made MCGDM research of IRO evaluation practically essential. On the other hand, the two long standing key issues in MCGDM, i.e., determination of weights and group consensus have made MCGDM research of IRO evaluation theoretically essential. This paper is a study of maximum consensus MCGDM for IRO evaluation. Considering the inevitable subjective judgments and preferences in IRO evaluation, this paper further extends this problem into fuzzy environment.

Firstly this paper identified a set of objective and subjective evaluation criteria, then developed a fuzzy distance-based method to determine evaluators' weights which minimizes the sum of the Euclidean distances between all pairs of fuzzy evaluation results to enhance group consensus. To determine criteria's weights, an intuitionistic fuzzy weighted averaging operator was used. After that, this paper employed a revised TOPSIS method to aggregate both crisp and fuzzy IRO ratings. Finally, a case study with actual data was conducted to test the feasibility and effectiveness of the proposed method. Compared with the results of arbitrary weights of evaluators, the results from this paper represent the maximum agreement result, therefore, it is more persuasive. Compared with single indicator evaluation, the proposed method takes more aspects into consideration, therefore it can overcome one-sidedness of a single indicator. Compared with other comprehensive IRO evaluation methods, the proposed method also has advantages and applicability.

Due to the complexity of IRO evaluation, the following are areas for our future research: (1) different objective and subjective criteria for different research areas; (2) methods in handling different kinds of information expressed in linguistic terms, boolean values, as well as numeric values to assess and rank IRO; and (3) better understanding and presentation of subjective information.

## Acknowledgements

This research was supported by Chinese Universities Scientific Fund (Grant No. 2010SCU22009) and "985" Program of Sichuan University "Innovative Research Base for Economic Development and Management". Authors are grateful for financial support from the program of China Scholarships Council (No. 201206240126).

## References

[1] D.W. Aksnes, R.E. Taxt, Peer reviews and bibliometric indicators: a comparative study at a Norwegian university, Res. Eval. 13 (2004) 33–41.
[2] K.T. Atanassov, Intuitionistic fuzzy sets, Fuzzy Sets Syst. 20 (1986) 87–96.
[3] L. Bornmann, H.D. Daniel, The citation speed index: a useful bibliometric indicator to add to the h index, J. Inform. 4 (2010) 444–446.
[4] Z. Chunting, How to evaluate a researcher's academic performance? – on the citation-bubble problem and its solution, Sci. Technol. Rev. (2009) 10.
[5] M. Dursun, E.E. Karsak, M.A. Karadayi, A fuzzy multi-criteria group decision making framework for evaluating health-care waste disposal alternatives, Expert Syst. Appl. 38 (2011) 11453–11462.
[6] L. Egghe, Theory and practise of the g-index, Scientometrics 69 (2006) 131–152.
[7] L. Egghe, R. Rousseau, An h-index weighted by citation impact, Inform. Process. Manage. 44 (2008) 770–780.
[8] L. Egghe, R. Guns, R. Rousseau, Measuring co-authors' contribution to an article's visibility, Scientometrics 95 (2013) 55–67.
[9] J.W. Fedderke, The objectivity of national research foundation peer review in South Africa assessed against bibliometric indexes, Scientometrics 97 (2013) 177–206.
[10] C. Fu, S. Yang, An evidential reasoning based consensus model for multiple attribute group decision analysis problems with interval-valued group consensus requirements, Euro. J. Operat. Res. 223 (2012) 167–176.
[11] A. Geuna, B.R. Martin, University research evaluation and funding: an international comparison, Minerva 41 (2003) 277–304.
[12] J.E. Hirsch, An index to quantify an individual's scientific research output, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 16569–16572.
[13] A.W. Harzing, R. van der Wal, A Google Scholar h-index for journals: an alternative metric to measure journal impact in economics and business, J. Am. Soc. Inform. Sci. Technol. 60 (2009) 41–46.
[14] C.L. Hwang, K. Yoon, Multiple Attributes Decision Making Methods and Applications, Springer, BerlinHeiddberg, 1981.
[15] B. Jin, L. Liang, R. Rousseau, L. Egghe, The R-and AR-indices: complementing the h-index, Chin. Sci. Bullet. 52 (2007) 855–863.
[16] S. Lehmann, A.D. Jackson, B.E. Lautrup, A quantitative analysis of indicators of scientific performance, Scientometrics 76 (2008) 369–390.
[17] J. Ma, J. Lu, G. Zhang, Decider: A fuzzy multi-criteria group decision support system, Knowl.-Based Syst. 23 (2010) 23–31.
[18] B.R. Martin, J.E.F. Skea, Academic Research Performance Indicators: An Assessment of the Possibilities, University of Sussex, Science Policy Research Unit, 1992.
[19] C. Musselin, How peer review empowers the academic profession and university managers: changes in relationships between the state, universities and the professoriate, Res. Policy 42 (2013) 1165–1173.
[20] A.J. Nederhof, A.F.J. Van Raan, A bibliometric analysis of six economics research groups: a comparison with peer review, Res. Policy 22 (1993) 353–368.
[21] E.J. Rinia, T.N. Van Leeuwen, H.G. Van Vuren, A.F.J. Van Raan, Influence of interdisciplinarity on peer-review and bibliometric evaluations in physics research, Res. Policy 30 (2001) 357–361.
[22] A.K. Shaw, T.K. Roy, Trapezoidal intuitionistic fuzzy number with some arithmetic operations and its application on reliability evaluation, Int. J. Math. Operat. Res. 5 (2013) 55–73.
[23] M. Schreiber, C.C. Malesios, S. Psarakis, Exploratory factor analysis for the Hirsch index, 17 h-type variants, and some traditional bibliometric indicators, J. Inform. 6 (2012) 347–358.
[24] E. Szmidt, J. Kacprzyk, Distances between intuitionistic fuzzy sets, Fuzzy Sets Syst. 114 (2000) 505–518.
[25] A.F.J. Van Raan, Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises, Scientometrics 36 (1996) 397–420.
[26] B. Vahdani, S.M. Mousavi, H. Hashemi, M. Mousakhani, R. Tavakkoli-Moghaddam, A new compromise solution method for fuzzy group decision-making problems with an application to the contractor selection, Eng. Appl. Artif. Intell. 26 (2013) 779–788.
[27] J.P. Xu, J.Z. Chen, TOPSIS based interactive multi-attributes group decision-making method and its application, J. Syst. Eng. 3 (2008) 276–281.
[28] J.P. Xu, Z.B. Wu, A discrete consensus support model for multiple attribute group decision making, Knowl.-Based Syst. 24 (2011) 1196–1202.
[29] J.P. Xu, Z.M. Li, W.J. Shen, B. Lev, Multi-attribute comprehensive evaluation of individual research output based on published research papers, Knowl.-Based Syst. 43 (2013) 135–142.
[30] Z.S. Xu, Intuitionistic fuzzy aggregation operators, IEEE Trans. Fuzzy Syst. 15 (2007) 1179–1187.
[31] R.R. Yager, Aggregation operators and fuzzy systems modeling, Fuzzy Sets Syst. 67 (1994) 129–145.
[32] J. Ye, Multicriteria group decision-making method using vector similarity measures for trapezoidal intuitionistic fuzzy numbers, Group Dec. Negot. 21 (2012) 519–530.
[33] L. Yu, K.K. Lai, A distance-based group decision-making methodology for multi-person multi-criteria emergency decision support, Dec. Supp. Syst. 51 (2011) 307–315.
[34] L.A. Zadeh, Fuzzy sets, Inform. Control 8 (1965) 338–353.
[35] S.F. Zhang, S.Y. Liu, A GRA-based intuitionistic fuzzy multi-criteria group decision making method for personnel selection, Expert Syst. Appl. 38 (2011) 11401–11405.