Satish S. Salunkhe*, Yashwant Joshi and Ashok Deshpande

Can Fuzzy Relational Calculus Bring Complex Issues in Selection of Examiners into Focus?

DOI 10.1515/jisys-2015-0105

Received September 22, 2015; previously published online December 17, 2015.

Abstract: The examinee and the examiner play pivotal roles in the educational grading system. Students' academic performance evaluation by multiple experts involves epistemic uncertainty, which can be modeled using a fuzzy set theory. How many evaluators/experts are almost similar in their perceptual subjective evaluation of the students answer paper? In other words, how many experts are reliable for a particular evaluation task with a defined possibility level? In this paper, the focus is on *object's features (students' marks)* as a basis in the subjective evaluation process to identify the *degree of similarity* among the domain experts. The case study reveals that 11 out of 20 evaluators are similar in their decision making of students' academic performance with possibility (α -level cut, 0.98). The inter-rater reliability (κ -coefficient) among the selected 11 teachers is 0.41, which signifies a fair/moderate agreement in the evaluation process. This paper proposes an approach that is useful for the selection of experts having similar perceptions in judgment. This paper demonstrates a case study showing how it is useful to educational policy makers in the selection of examiners.

Keywords: Education grading system, examinee and examiner, inter-rater reliability, κ -coefficient, expert's perception, fuzzy sets, fuzzy relational calculus, cosine amplitude method.

1 Introduction

It is crucial to identify unbiased experts with similar perceptions in their judgment. Several researchers have studied this important aspect from various viewpoints: Einhorn [11] argued that a *consensus* between experts is a necessary condition for expertise. Hoffmann [19] used a simple model designed by McGrew [31]. He inferred that consensus in the relevant expert community is an indicator if the following two necessary conditions are fulfilled: *reliability* and *statistical independence*. There is no *epistemic* means to determine the general *reliability* of genuine moral experts [19]. Due to variability in experts' judgments, it is essential to scrutinize multiple experts to obtain the cluster having a similarity in perception with increased *validity* and *reliability* in evaluation tasks [1, 5, 9–12, 15, 18, 28, 29, 34, 35, 41, 42, 44, 47]. Obtaining a single distribution of elicited information that comprises several experts' beliefs is desirable [8, 33]. The correlation between the experts' judgments should be greater, perhaps much greater, when they are judging the same trait than when they are judging different traits [42]. Goldman [17] discussed about how laypersons should evaluate the testimony of experts and decide which of two or more rival experts is most credible. Goldman's definition of identification criteria for experts [16, 17] has been improved by Scholz [40]. Only the expert has epistemic access to the knowledge of the domain D of expertise (Goldman calls it *esoteric knowledge*). By contrast, the layperson only has access to *exoteric knowledge*, i.e. knowledge outside the domain D (Ref. [17], p. 94). The crucial question is: how can a layperson (merely with the help of the exoteric knowledge) identify an expert without having the relevant esoteric knowledge and the cognitive abilities? Ashton [1] suggested an approach

*Corresponding author: **Satish S. Salunkhe**, Mukesh Patel School of Technology Management and Engineering, Narsee Monjee Institute of Management Studies, Mumbai 400 056, India, Tel.: +919869175103, e-mail: satishssalunkhe@gmail.com

Yashwant Joshi: Shri Guru Govind Singh Institute of Engineering and Technology, Swami Ramanand Teerth Marathwada University, Nanded 431 606, India

Ashok Deshpande: Berkeley Initiative Soft Computing-SIG-(EMS), Berkeley, CA 94720-5800, USA

to measure the validity of an experts group while adding a new expert in the group. The studies by Ashton and Ashton [2], Libby and Blashfield [27], Makridakis and Winkler [30], and Winkler and Makridakis [49] follow factorial experimental design for experts selection using group validity. This experimental design has more computational time complexity. Hierarchical clustering can be obtained using fuzzy relational calculus, with a factorial experimental design [24, 25, 43, 48, 51].

After the emergence of the fuzzy set theory in the simple task of looking at relations as fuzzy sets in the universe, as accomplished in a celebrated paper by Zadeh [51], he introduced the concept of fuzzy relation, defined as the notion of equivalence, and gave the concept of fuzzy ordering. Several researchers have made seminal contributions and extended the concept of fuzzy relations. Identification of similar experts/examiners in performance evaluation relates to mostly fuzzy classification based on fuzzy similarity relation. The mathematical concept of fuzzy *equivalence relation* is the basis for fuzzy classification. Although in the literature [3, 4, 6, 7, 20, 22, 26, 32, 38, 45, 46], various approaches for the subjective evaluation of *students' answer script* are proposed using fuzzy sets and logic, there are no references for how to evaluate/classify the expertise of *experts* before they will evaluate the students' answer scripts. How many teachers are almost *similar* in their subjective judgment while evaluating students' answer script? In other words, how many teachers are *reliable* for a particular evaluation task with a defined *possibility level* (α -level cut)? The authors have made an attempt to address this issue using *fuzzy sets* as a basis (part A) with a focus on fuzzy relational calculus to identify similar experts [39]. The study in part A relates to the agreement of teachers based on fuzzy sets at a defined α -level cut using fuzzy similarity measures. It is essential to confirm the similarity among teachers based on actual evaluation of the object (students) as an additional measure.

The paper is organized as follows: Section 2 refers to mathematical preliminaries, while an approach for the selection of teachers for students' evaluation, based on fuzzy sets theoretic operations and fuzzy relational calculus, is described in Section 3. The case study for selecting fuzzy similarity-based examiner/teacher is covered in Section 4. The results and discussion of the case study are discussed in Section 5. The conclusion and future scope for research are integral parts of Section 6.

2 Preliminaries and Notations

This section briefly describes fuzzy relation/fuzzy relational calculus [52] operations used in the paper.

Definition 2.1: Let U be a universe set. A fuzzy set A of U is defined with a membership $\mu_A(x) \rightarrow [0, 1]$, where $\mu_A(x)$, $\forall x \in U$ indicates the degree of x in A [50, 51].

Definition 2.2: Let R be a fuzzy relation on $X \times Y$, i.e. $R = \{((x, y), f_R(x, y)) | (x, y) \in X \times Y\}$, the α -cut matrix R_α is denoted by

$$R_\alpha = \{((x, y), f_R(x, y)) | f_R(x, y) = 1, \text{ if } f_R(x, y) \geq \alpha; f_R(x, y) = 0; \text{ if } f_R(x, y) < \alpha, (x, y) \in X \times Y, \alpha \in [0, 1]\}.$$

Definition 2.3: Let $R \subset X \times Y$ and $S \subset Y \times Z$ be fuzzy relations; the max-min composition $R \circ S$ is defined by

$$R \circ S = \{((x, z), \max_y \{\min_{x,z} \{f_R(x, y), f_S(y, z)\}\}) | x \in X, y \in Y, z \in Z\}.$$

Definition 2.4: A fuzzy relation R on $X \times X$ is called a fuzzy equivalence relation if the following three conditions held [37, 50, 51]:

1. R is reflexive if $f_R(x, x) = 1$, $\forall x \in X$.
2. R is symmetric if $f_R(x, y) = f_R(y, x)$, $\forall x, y \in X$.
3. R is transitive if $R^{(2)} = (R \circ R) \subset R$, or more explicitly

$$f_R(x, z) \geq \max_y \{\min_{x,z} \{f_R(x, y), f_R(y, z)\}\}, \forall x, y, z \in X.$$

Definition 2.5: A fuzzy relation R on $X \times Y$ is called is a fuzzy compatible or tolerance or proximity relation if it satisfies reflexive and symmetric conditions.

Definition 2.6: The transitive closure, R_T , of a fuzzy relation R is defined as the relation that is transitive, contain R , and has the smallest possible membership grades.

Definition 2.7: Let R be a fuzzy compatibility relation on a finite universal set X with $|X| = n$, then the max-min transitive closure of R in the relation is defined as the relation $R^{(n-1)}$ [21, 25].

Algorithm A: Find the transitive closure R_T of fuzzy compatibility relation [25].

- Step 1: Calculate $R^{(2)}$ if $R^{(2)} \subset R$ or $R^{(2)} = R$, then transitive closure $R_T = R$ and stop. Otherwise, $k = 2$, go to step 2.
- Step 2: If $2^k \geq n - 1$, then $R_T = R^{(n-1)}$ and stop. Otherwise, calculate $R^{(2^k)} = R^{(2^{k-1})} \circ R^{(2^{k-1})}$, if $R^{(2^k)} = R^{(2^{k-1})}$, then transitive closure $R_T = R^{(2^k)}$ and stop. Otherwise, go to step 3.
- Step 3: $k = k + 1$, go to step 2.

Definition 2.8: A partition of S means a family of disjoint subsets, say $\{S_1, S_2, \dots, S_n\}$, such that the union of these subsets coincides with the entire set S . In other words, $S_1 \cup S_2 \cup \dots \cup S_n$ and $S_i \cap S_j = \emptyset$, $\forall i \neq j$.

Definition 2.9: The normalized Euclidean distance [23, 26]:

$$d(A, B) = \sqrt{\frac{\sum_{i=1}^n |m_A(x_i) - m_B(x_i)|^2}{n}}, \quad (1)$$

where n is the cardinality of the universe of discourse, A and B are fuzzy sets, and d is a distance measure.

Definition 2.10: Fleiss κ [13, 14] is a statistical measure of inter-rater reliability. κ can be defined as

$$\kappa = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}. \quad (2)$$

The factor $1 - \bar{P}_e$ gives the degree of agreement that is attainable above chance, and $\bar{P} - \bar{P}_e$ gives the degree of agreement actually achieved above chance. If the raters agree strongly, then $\kappa = 1$. If there is no agreement among the raters, then $\kappa < 0$.

Definition 2.11: The cosine amplitude method is manipulated on a collection of n data samples. The collected data set is represented as $X = \{x_1, x_2, \dots, x_n\}$. Each of the n samples is represented as a vector with an m dimension, $x_i = \{x_{i1}, x_{i2}, \dots, x_{im}\}$. The position of each datum in space is represented by m feature values. The relation value r_{ij} reflects a similarity relationship between x_i and x_j data. For n data samples, the size of the relation matrix will be $n \times n$. The relation matrix always obeys the rules of being reflexive and symmetric, and so it is a tolerance relation. All r_{ij} values are always in the interval of $[0, 1]$ in this method, and they are calculated through the following equation:

$$r_{ij} = \frac{\left| \sum_{k=1}^m x_{ik} x_{jk} \right|}{\sqrt{\left(\sum_{k=1}^m x_{ik}^2 \right) \left(\sum_{k=1}^m x_{jk}^2 \right)}}, \quad i, j = 1, 2, \dots, n. \quad (3)$$

If x_i and x_j are very similar to each other, r_{ij} becomes close to one. Unlike this situation, if they are very dissimilar to each other, r_{ij} becomes close to zero.

3 The Proposed Approach

Fuzzy sets as a basis (Part A) have been used to identify similar experts by the authors as a first step in the experts selection process [39]. The study in part A relates to agreement among experts at a defined α -cut level in construction of fuzzy sets. However, it is also equally important to identify the pairwise similarity among experts when they actually evaluate the features of the object. This has prompted the authors to apply one of the facets of other fuzzy relational calculus, which is explained below. Further screening of experts can be done using “Objects Features as a Basis (Part B),” which is proposed below.

3.1 Object’s Characteristics (e.g. Students’ Academic Performance) as a Basis (Part B)

The steps to follow using cosine amplitude similarity calculation among n experts for each object O_k are summarized below.

Step 1 – For each object O_k , extract a column feature vector $V_{jk} = \{m(Q_{jk1}), m(Q_{jk2}), \dots, m(Q_{jkn})\}$, which is a collection of all feature values (i.e. attributes) Q_i , from each expert E_j , where $1 \leq i \leq n$.

Step 2 – Each feature may have a different weight, w_i . Normalize each feature values $Q_{jki} \in V_j$ with respect to the weightage of feature w_i to obtain the normalized vector of an object O_k , where $1 \leq i \leq n$.

$$m'(Q_{jki}) = m(Q_{jki}) / w_i. \quad (4)$$

$$V'_{jk} = \{m'(Q_{jk1}), m'(Q_{jk2}), \dots, m'(Q_{jkn})\}. \quad (5)$$

Step 3 – Normalize the column feature vector V'_{jk} to obtain V''_{jk} ,

$$m''(Q_{jki}) = \frac{m'(Q_{jki})}{\sum_{i=1}^n m'(Q_{jki})}, \quad (6)$$

$$V''_{jk} = \{m''(Q_{jk1}), m''(Q_{jk2}), \dots, m''(Q_{jkn})\}. \quad (7)$$

Step 4 – Apply the similarity measure technique on V''_{jk} , i.e. on $\langle V''_{1k}, V''_{2k}, \dots, V''_{nk} \rangle$, to get the fuzzy tolerance relation similarity matrix $[S_k]_{n \times n}$. This similarity matrix reflects pairwise similarity index between n experts in the context of an object O_k .

Step 5 – Using max-min composition and transitive closure, transform the fuzzy tolerance relation $[S_k]_{n \times n}$ obtained in step 4 to fuzzy equivalence relation $[S'_k]_{n \times n}$.

Step 6 – For each α -cut value (i.e. possibility) of fuzzy equivalence relation $[S'_k]_{n \times n}$, find the clusters (i.e. partitions) of experts.

Step 7 – Select the cluster having the highest cardinality of experts. If there are more than one clusters having equal or marginal equal numbers of experts, then select all those clusters. We presuppose that the confidence in the decision making of similar experts selection could strengthen when a group comprising the highest number of experts is selected at a defined α -cut (possibility) level.

Step 8 – Repeat steps 1–7 for each object O_k . Calculate the frequency distribution of each expert based on its occurrence in selected clusters obtained in step 7 for all k objects at a reference α -cut (e.g. $\alpha = 0.98$).

Step 9 – Organizational policies may then decide a certain cutoff limit on the rate of experts’ occurrence in selected clusters. Select top n experts who are above the cutoff limit (e.g. above 80%).

4 Case Study

The case study relates the selection of examiners/teachers for evaluating the answer script of students having a similar evaluation perception. The examination answer script samples in subject *Marathi* language were

obtained from 237 secondary school students from three different institutions in Mumbai, India, during academic year 2013–2014. Each student wrote a 10–12-page solution with respect to 12 subjective questions. The questions selected in the question paper can be used to assess students' writing ability. Answers were evaluated using a 10-point rubric for our study from the Secondary School (SSC) Board, Maharashtra. Also, 20 subject matter experts (teachers) from different schools were identified for the answer scripts evaluation. All 20 evaluators belonging to the "Marathi" department of the SSC board evaluated every 237 answer scripts. This experiment helped us observe the nature of consistency among all experts in the evaluation process. Experts were trained before any evaluation began with respect to the assessment rubric scale and evaluation process. Before the evaluation process begins, each evaluator expressed their evaluation judgment based on their perception for each linguistic variable. Five satisfaction levels, g , to award marks scored for the 12 questions, are selected, like *very poor*, *poor*, *average*, *good*, and *very good*. Then, all the 20 teachers constructed fuzzy sets for each linguistic variable. The evaluation process for the pilot study occurred in May 2014. All 20 evaluators met together in the central assessment examination room in Marathi Vidyalaya, Mumbai (India), school campus for evaluation. Each teacher was given an answer script to grade. Each teacher awarded a score to every question in numeric value based on the weightage of the question, as shown in Table 1. All teachers reported the students' performance on a separate result sheet. The evaluators made no marking on the answer script. In the evaluation process, the students' identities were concealed using masks and identified using unique alphanumeric codes to preserve anonymity. Let

- $(Student)_k$ denote the k th student, where $1 \leq k \leq 237$;
- T_j denote the j th teacher, where $1 \leq j \leq 20$;
- Q_i denote the i th question, where $1 \leq i \leq 12$;
- $m(Q_{ijn})$ denote the marks awarded to question n , by teacher j , of student k ; and
- W_i denote the weightage of question i , as shown in Table 1.

The typical computations to follow using the cosine amplitude similarity measure among 20 experts for $(Student)_{12}$ are summarized below. The typical evaluation score sheet of student roll number 12 evaluated by 20 experts is shown in Table 2, which is normalized as shown in Table 3. A column-wise normalized data sheet is shown in Table 4. These normalized data of students' marks corresponding to each teacher column is a vector that is used in pairwise cosine amplitude similarity measure computations among all teacher vectors. The similarity index value is in the range $[0, 1]$, where 0 indicates absolute dissimilarity and 1 indicates absolute similarity among teachers in the context of evaluation judgment.

5 Results and Discussion

The procedure detailed in Section 3 was followed, and the results obtained are discussed below:

The relation in Table 3 is reflexive, symmetric but is not transitive and is a fuzzy tolerance relation between 20 experts for student 12. The fuzzy tolerance relation shown in Table 5 is transformed to a fuzzy equivalence relation, as shown in Table 6.

It is necessary to convert the fuzzy equivalence relation presented in Table 6 to a crisp value known as the defuzzification process, as shown in Table 7 at possibility α -value 0.98. Figure 1 shows the partitioning of 20 teachers at various intervals of α -cut for student S_{12} . Cluster $\{1,2,3,4,5,7,8,9,10,11,12,17,18,19\}$ having the maximum number of experts as shown in Table 8 at possibility α -value 0.98 for student ID 12 is selected.

The detailed computational procedure given in Section 3 is performed on all 237 students in order to identify similar experts at the defined α -cut level ($\alpha = 0.98$). The frequency distribution of occurrence for each

Table 1: Weightage of Each Question.

Question no.	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12
Marks (W_i)	4	3	3	5	4	3	3	4	4	8	10	4

Table 2: Data Sheet for Student ID 12 Evaluated by 20 Experts.

Question no.	Teacher 01	Teacher 02	Teacher 03	Teacher 04	Teacher 05	Teacher 06	Teacher 07	Teacher 08	Teacher 09	Teacher 10	Teacher 11	Teacher 12	Teacher 13	Teacher 14	Teacher 15	Teacher 16	Teacher 17	Teacher 18	Teacher 19	Teacher 20
Q1	3	3	2.5	3	3	3	3	3	2	4	2.5	3	4	4	4	1.5	2.5	3	3	2.5
Q2	2	2.5	2	3	2.5	2	1.5	2.5	3	3	1.5	2	3	3	3	3	2	2	3	2
Q3	2	2	1.5	2.5	3	3	2	3	3	3	2	2	3	3	3	1.5	1.5	2.5	3	2.5
Q4	3	3	3	3	3	3	2	4	3.5	4	2.5	3	4	4	4	3	2	3.5	3	3
Q5	3	2.5	2	2.5	2.5	1	1	3.5	3	3	1.5	2	4	4	4	2	2	2	2	3
Q6	2.5	2	2	3	2	1.5	2.5	2	2	2.5	2	2	2	2	2	3	1.5	2.5	2	2.5
Q7	2.5	2.5	2	2	2.5	2	1.5	2	2	2.5	1	1.5	2	2	2	1	1.5	3	2.5	2.5
Q8	3.5	3	1.5	2.5	3	3	2	2.5	2	3	1.5	2	3	3	3	2	2.5	4	3	2
Q9	3.5	2.5	3	3	3	2	2	2	2	3	2	2.5	3	3	3	1.5	2	2.5	2	2
Q10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Q11	5.5	5	6	7	7	6	6.5	4	5	7	4	5	6	6	5	4	4	7.5	5	5
Q12	3	2.5	3	3.5	3.5	4	3	4	3	3.5	2.5	2	1.5	1	1	3	2	3.5	2.5	1.5

Table 3: Normalized Data Sheet of Student ID 12.

Question no.	Teacher 01	Teacher 02	Teacher 03	Teacher 04	Teacher 05	Teacher 06	Teacher 07	Teacher 08	Teacher 09	Teacher 10	Teacher 11	Teacher 12	Teacher 13	Teacher 14	Teacher 15	Teacher 16	Teacher 17	Teacher 18	Teacher 19	Teacher 20
Q1	7.5	7.5	6.25	7.5	7.5	7.5	7.5	7.5	5	10	6.25	7.5	10	10	10	3.75	6.25	7.5	7.5	6.25
Q2	6.667	8.3333	6.6667	10	8.3333	6.6667	5	8.3333	10	10	5	6.6667	10	10	10	10	6.6667	6.6667	10	6.6667
Q3	6.6667	6.6667	5	8.3333	10	10	6.6667	10	10	10	6.6667	6.6667	10	10	10	5	5	8.3333	10	8.3333
Q4	6	6	6	6	6	6	4	8	7	8	5	6	8	8	8	6	4	7	6	6
Q5	7.5	6.25	5	6.25	6.25	2.5	2.5	8.75	7.5	7.5	3.75	5	10	10	10	5	5	5	5	7.5
Q6	8.3333	6.6667	6.6667	10	6.6667	5	8.3333	6.6667	6.6667	8.3333	6.6667	6.6667	6.6667	6.6667	6.6667	10	5	8.3333	6.6667	8.3333
Q7	8.3333	8.3333	6.6667	6.6667	8.3333	6.6667	5	6.6667	6.6667	8.3333	3.3333	5	6.6667	6.6667	6.6667	3.3333	5	10	8.3333	8.3333
Q8	8.75	7.5	3.75	6.25	7.5	7.5	5	6.25	5	7.5	3.75	5	7.5	7.5	7.5	5	6.25	10	7.5	5
Q9	8.75	6.25	7.5	7.5	7.5	5	5	5	5	7.5	5	6.25	7.5	7.5	7.5	3.75	5	6.25	5	5
Q10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Q11	5.5	5	6	7	7	6	6.5	4	5	7	4	5	6	6	5	4	4	7.5	5	5
Q12	7.5	6.25	7.5	8.75	8.75	10	7.5	10	7.5	8.75	6.25	5	3.75	2.5	2.5	7.5	5	8.75	6.25	3.75

Table 4: Column-Wise Normalized Data Sheet of Student ID 12.

Question no.	Teacher 01	Teacher 02	Teacher 03	Teacher 04	Teacher 05	Teacher 06	Teacher 07	Teacher 08	Teacher 09	Teacher 10	Teacher 11	Teacher 12	Teacher 13	Teacher 14	Teacher 15	Teacher 16	Teacher 17	Teacher 18	Teacher 19	Teacher 20
Q1	0.092	0.1003	0.0933	0.089	0.0895	0.103	0.119	0.0924	0.0664	0.1076	0.1123	0.1158	0.1162	0.1179	0.1193	0.0592	0.1093	0.0879	0.0971	0.0891
Q2	0.0818	0.1115	0.0995	0.1187	0.0994	0.0915	0.0794	0.1027	0.1327	0.1076	0.0898	0.103	0.1162	0.1179	0.1193	0.1579	0.1166	0.0781	0.1294	0.095
Q3	0.0818	0.0892	0.0746	0.0989	0.1193	0.1373	0.1058	0.1232	0.1327	0.1076	0.1198	0.103	0.1162	0.1179	0.1193	0.0789	0.0875	0.0977	0.1294	0.1188
Q4	0.0736	0.0803	0.0896	0.0712	0.0716	0.0824	0.0635	0.0986	0.0929	0.0861	0.0898	0.0927	0.0929	0.0943	0.0954	0.0947	0.07	0.082	0.0777	0.0855
Q5	0.092	0.0836	0.0746	0.0742	0.0746	0.0343	0.0397	0.1078	0.0996	0.0807	0.0674	0.0772	0.1162	0.1179	0.1193	0.0789	0.0875	0.0586	0.0647	0.1069
Q6	0.1022	0.0892	0.0995	0.1187	0.0795	0.0686	0.1323	0.0821	0.0885	0.0897	0.1198	0.103	0.0774	0.0786	0.0795	0.1579	0.0875	0.0977	0.0863	0.1188
Q7	0.1022	0.1115	0.0995	0.0791	0.0994	0.0915	0.0794	0.0821	0.0885	0.0897	0.0599	0.0772	0.0774	0.0786	0.0795	0.0526	0.0875	0.1172	0.1079	0.1188
Q8	0.1074	0.1003	0.056	0.0742	0.0895	0.103	0.0794	0.077	0.0664	0.0807	0.0674	0.0772	0.0871	0.0884	0.0895	0.0789	0.1093	0.1172	0.0971	0.0713
Q9	0.1074	0.0836	0.1119	0.089	0.0895	0.0686	0.0794	0.0616	0.0664	0.0807	0.0898	0.0965	0.0871	0.0884	0.0895	0.0592	0.0875	0.0732	0.0647	0.0713
Q10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Q11	0.0675	0.0669	0.0896	0.0831	0.0835	0.0824	0.1032	0.0493	0.0664	0.0753	0.0719	0.0772	0.0697	0.0707	0.0596	0.0632	0.07	0.0879	0.0647	0.0713
Q12	0.092	0.0836	0.1119	0.1039	0.1044	0.1373	0.119	0.1232	0.0996	0.0942	0.1123	0.0772	0.0436	0.0295	0.0298	0.1184	0.0875	0.1025	0.0809	0.0534

Table 5: Fuzzy Tolerance Relation among 20 Experts.

Student 12	Teacher 01	Teacher 02	Teacher 03	Teacher 04	Teacher 05	Teacher 06	Teacher 07	Teacher 08	Teacher 09	Teacher 10	Teacher 11	Teacher 12	Teacher 13	Teacher 14	Teacher 15	Teacher 16	Teacher 17	Teacher 18	Teacher 19	Teacher 20
Teacher 01	1	0.9889	0.976	0.9753	0.9806	0.9397	0.9526	0.9571	0.9471	0.9802	0.9618	0.9784	0.9563	0.9474	0.9464	0.9166	0.9869	0.9813	0.9597	0.9654
Teacher 02	0.9889	1	0.9737	0.9799	0.9869	0.9522	0.9492	0.9683	0.9678	0.9919	0.9605	0.9843	0.9721	0.9646	0.9638	0.9279	0.995	0.9826	0.9858	0.9753
Teacher 03	0.976	0.9737	1	0.9837	0.9762	0.9408	0.9625	0.9534	0.9532	0.9801	0.9713	0.9792	0.9376	0.9256	0.9218	0.9298	0.968	0.9628	0.9481	0.9504
Teacher 04	0.9753	0.9799	0.9837	1	0.9841	0.9533	0.9759	0.9667	0.9748	0.9891	0.9852	0.9866	0.953	0.9423	0.9395	0.9657	0.9816	0.968	0.9733	0.9642
Teacher 05	0.9806	0.9869	0.9762	0.9841	1	0.9789	0.9652	0.9757	0.9766	0.993	0.9744	0.9815	0.9602	0.9494	0.9465	0.9186	0.9842	0.9841	0.9849	0.9653
Teacher 06	0.9397	0.9522	0.9408	0.9533	0.9789	1	0.9626	0.9574	0.945	0.9678	0.9613	0.9494	0.9092	0.8938	0.891	0.8895	0.9512	0.9719	0.9665	0.9122
Teacher 07	0.9526	0.9492	0.9625	0.9759	0.9652	0.9626	1	0.9324	0.9239	0.9667	0.9814	0.9669	0.9071	0.8942	0.8888	0.9202	0.9519	0.9673	0.9449	0.9307
Teacher 08	0.9571	0.9683	0.9534	0.9667	0.9757	0.9574	0.9324	1	0.9853	0.9822	0.9703	0.9658	0.9562	0.9428	0.9442	0.9325	0.966	0.9523	0.9679	0.9563
Teacher 09	0.9471	0.9678	0.9532	0.9748	0.9766	0.945	0.9239	0.9853	1	0.9798	0.9601	0.9649	0.961	0.9512	0.9506	0.948	0.9619	0.9453	0.9789	0.9675
Teacher 10	0.9802	0.9919	0.9801	0.9891	0.993	0.9678	0.9667	0.9822	0.9798	1	0.9839	0.9943	0.9764	0.9672	0.9655	0.9363	0.9906	0.977	0.9868	0.9744
Teacher 11	0.9618	0.9605	0.9713	0.9852	0.9744	0.9613	0.9814	0.9703	0.9601	0.9839	1	0.9866	0.9463	0.9345	0.933	0.9416	0.9644	0.9567	0.9583	0.9509
Teacher 12	0.9784	0.9843	0.9792	0.9866	0.9815	0.9494	0.9669	0.9658	0.9649	0.9943	0.9866	1	0.9796	0.9731	0.9712	0.9326	0.9855	0.9658	0.9742	0.9726
Teacher 13	0.9563	0.9721	0.9376	0.953	0.9602	0.9092	0.9071	0.9562	0.961	0.9764	0.9463	0.9796	1	0.9989	0.9983	0.8951	0.9746	0.9323	0.9657	0.9722
Teacher 14	0.9474	0.9646	0.9256	0.9423	0.9494	0.8938	0.8942	0.9428	0.9512	0.9672	0.9345	0.9731	0.9989	1	0.9993	0.8829	0.9665	0.9218	0.9588	0.9695
Teacher 15	0.9464	0.9638	0.9218	0.9395	0.9465	0.891	0.8888	0.9442	0.9506	0.9655	0.933	0.9712	0.9983	0.9993	1	0.8825	0.9654	0.9182	0.9585	0.9684
Teacher 16	0.9166	0.9279	0.9298	0.9657	0.9186	0.8895	0.9202	0.9325	0.948	0.9363	0.9416	0.9326	0.8951	0.8829	0.8825	1	0.9319	0.909	0.9254	0.9131
Teacher 17	0.9869	0.995	0.968	0.9816	0.9842	0.9512	0.9519	0.966	0.9619	0.9906	0.9644	0.9855	0.9746	0.9665	0.9654	0.9319	1	0.9743	0.9801	0.9615
Teacher 18	0.9813	0.9826	0.9628	0.968	0.9841	0.9719	0.9673	0.9523	0.9453	0.977	0.9567	0.9658	0.9323	0.9218	0.9182	0.909	0.9743	1	0.9713	0.9563
Teacher 19	0.9597	0.9858	0.9481	0.9733	0.9849	0.9665	0.9449	0.9679	0.9789	0.9868	0.9583	0.9742	0.9657	0.9588	0.9585	0.9254	0.9801	0.9713	1	0.9695
Teacher 20	0.9654	0.9753	0.9504	0.9642	0.9653	0.9122	0.9307	0.9563	0.9675	0.9744	0.9509	0.9726	0.9722	0.9695	0.9684	0.9131	0.9615	0.9563	0.9695	1

Table 6: Fuzzy Equivalence Relation among 20 Experts Using Max-Min Composition.

Student	12	Teacher	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20
Teacher 01	1	0.9889	0.9837	0.9889	0.9889	0.9789	0.9814	0.9822	0.9822	0.9822	0.9822	0.9889	0.9866	0.9889	0.9796	0.9796	0.9796	0.9657	0.9889	0.9841	0.9868	0.9753
Teacher 02	0.9889	1	0.9837	0.9891	0.9919	0.9789	0.9814	0.9822	0.9822	0.9822	0.9822	0.9919	0.9866	0.9919	0.9796	0.9796	0.9796	0.9657	0.995	0.9841	0.9868	0.9753
Teacher 03	0.9837	0.9837	1	0.9837	0.9837	0.9789	0.9814	0.9822	0.9822	0.9822	0.9822	0.9837	0.9837	0.9837	0.9796	0.9796	0.9796	0.9657	0.9837	0.9837	0.9837	0.9753
Teacher 04	0.9889	0.9891	0.9837	1	0.9891	0.9789	0.9814	0.9822	0.9822	0.9822	0.9822	0.9891	0.9866	0.9891	0.9796	0.9796	0.9796	0.9657	0.9891	0.9841	0.9868	0.9753
Teacher 05	0.9889	0.9919	0.9837	0.9891	1	0.9789	0.9814	0.9822	0.9822	0.9822	0.9822	0.993	0.9866	0.993	0.9796	0.9796	0.9796	0.9657	0.9919	0.9841	0.9868	0.9753
Teacher 06	0.9789	0.9789	0.9789	0.9789	0.9789	1	0.9789	0.9789	0.9789	0.9789	0.9789	0.9789	0.9866	0.9789	0.9789	0.9789	0.9789	0.9657	0.9789	0.9789	0.9789	0.9753
Teacher 07	0.9814	0.9814	0.9814	0.9814	0.9814	0.9789	1	0.9814	0.9814	0.9814	0.9814	0.9814	0.9814	0.9814	0.9796	0.9796	0.9796	0.9657	0.9814	0.9814	0.9814	0.9753
Teacher 08	0.9822	0.9822	0.9822	0.9822	0.9822	0.9789	0.9814	1	0.9853	0.9853	0.9853	0.9822	0.9822	0.9822	0.9796	0.9796	0.9796	0.9657	0.9822	0.9822	0.9822	0.9753
Teacher 09	0.9822	0.9822	0.9822	0.9822	0.9822	0.9789	0.9814	0.9822	0.9822	0.9822	0.9822	0.9822	0.9822	0.9822	0.9796	0.9796	0.9796	0.9657	0.9822	0.9822	0.9822	0.9753
Teacher 10	0.9889	0.9919	0.9837	0.9891	0.993	0.9789	0.9814	0.9822	0.9822	0.9822	0.9822	1	0.9866	0.9943	0.9796	0.9796	0.9796	0.9657	0.9919	0.9841	0.9868	0.9753
Teacher 11	0.9866	0.9866	0.9837	0.9865	0.9866	0.9789	0.9814	0.9822	0.9822	0.9822	0.9822	0.9866	1	0.9866	0.9796	0.9796	0.9796	0.9657	0.9866	0.9841	0.9866	0.9753
Teacher 12	0.9889	0.9919	0.9837	0.9891	0.993	0.9789	0.9814	0.9822	0.9822	0.9822	0.9822	0.9943	0.9866	1	0.9796	0.9796	0.9796	0.9657	0.9919	0.9841	0.9868	0.9753
Teacher 13	0.9796	0.9796	0.9796	0.9796	0.9796	0.9789	0.9796	0.9796	0.9796	0.9796	0.9796	0.9796	0.9796	0.9796	1	0.9989	0.9989	0.9657	0.9796	0.9796	0.9796	0.9753
Teacher 14	0.9796	0.9796	0.9796	0.9796	0.9796	0.9789	0.9796	0.9796	0.9796	0.9796	0.9796	0.9796	0.9796	0.9796	0.9796	1	0.9993	0.9657	0.9796	0.9796	0.9796	0.9753
Teacher 15	0.9796	0.9796	0.9796	0.9796	0.9796	0.9789	0.9796	0.9796	0.9796	0.9796	0.9796	0.9796	0.9796	0.9796	0.9796	0.9796	1	0.9657	0.9796	0.9796	0.9796	0.9753
Teacher 16	0.9657	0.9657	0.9657	0.9657	0.9657	0.9657	0.9657	0.9657	0.9657	0.9657	0.9657	0.9657	0.9657	0.9657	0.9657	0.9657	0.9657	1	0.9657	0.9657	0.9657	0.9657
Teacher 17	0.9889	0.995	0.9837	0.9891	0.9919	0.9789	0.9814	0.9822	0.9822	0.9822	0.9822	0.9919	0.9866	0.9919	0.9796	0.9796	0.9796	0.9657	1	0.9841	0.9868	0.9753
Teacher 18	0.9841	0.9841	0.9837	0.9841	0.9841	0.9789	0.9814	0.9822	0.9822	0.9822	0.9822	0.9841	0.9841	0.9841	0.9796	0.9796	0.9796	0.9657	0.9841	1	0.9841	0.9753
Teacher 19	0.9868	0.9868	0.9837	0.9868	0.9868	0.9789	0.9814	0.9822	0.9822	0.9822	0.9822	0.9868	0.9866	0.9868	0.9796	0.9796	0.9796	0.9657	0.9868	0.9841	1	0.9753
Teacher 20	0.9753	0.9753	0.9753	0.9753	0.9753	0.9753	0.9753	0.9753	0.9753	0.9753	0.9753	0.9753	0.9753	0.9753	0.9753	0.9753	0.9753	0.9657	0.9753	0.9753	0.9753	1

Figure 1: Dendrogram on Fuzzy Equivalence Relation $[S'_{12}]_{20 \times 20}$ for Partitioning of Teachers on Different Intervals of α -Cut Using the Cosine-Amplitude Method.

Table 8: The Partitioning of Teachers on Different Intervals of α -Cut 0.98 for Student ID 12.

Clustering method	Possibility α -value	No. of clusters	Obtained partitions of all 20 experts {clusters with elements}, different clusters are separated by ';' for Student ID-12
Cosine amplitude	0.9814	5	$\{\{1,2,3,4,5,7,8,9,10,11,12,17,18,19\}; \{6\}; \{13,14,15\}; \{16\}; \{20\}\}$

20 experts in context with all 237 students is shown in Table 9. The final ranking of all 20 experts at α -cut ($\alpha = 0.98$) using the cosine-amplitude method for all 237 students, ranging from highest to lowest, is given below. A histogram of all teachers is shown in Figure 2.

$$\langle T_{10} < T_1 < T_2 < T_8 < T_{15} < T_{14} < T_{13} < T_{11} < T_{18} < T_{12} < T_3 < T_{17} < T_4 < T_{20} < T_5 < T_7 < T_{16} < T_{19} < T_6 < T_9 \rangle$$

Educational *administrators/policy makers* select 11 out of 20 experts using a benchmark of 70% over the experts' occurrence frequency for all 237 students.

$$\langle T_{10}, T_1, T_2, T_8, T_{15}, T_{14}, T_{13}, T_{11}, T_{18}, T_{12}, T_3 \rangle$$

are selected for evaluating the students' academic performance in using the *Marathi* subject in Maharashtra State Board of Secondary School Certificate Exam, India.

The inter-rater reliability (Fleiss κ coefficient) among all 20 experts, selected 11 experts, and rejected 9 experts are computed for the dataset of 237 students' evaluation score sheet, as shown in Table 10. The Fleiss κ coefficient for the selected 11 experts is 0.41, indicating *moderate* agreement among all 11 experts according to Fleiss' guidelines to interpret the κ statistics. It can be also inferred that for all the 20 teachers as well as rejected 9 teachers, the computed κ coefficient is lower than that of the selected 11 teachers. This additional information will help the decision maker for the final selection of 11 experts. In our view, we can consider the perception of all the 11 experts for aggregating into multiexpert knowledgebase to obtain a fair result of evaluation. Eleven out of 20 experts are selected for future evaluation of the students' performance of Marathi subject in Maharashtra State Secondary Examination Board, Mumbai, India.

In the recent past, researchers have used various statistical and fuzzy logic-based methods in computing similarity between the experts with no general agreement. In summary, development of the state of the

Table 9: Ranking of 20 Expert's at α -Cut 0.98 Using Cosine-Amplitude Method for 237 Students.

Teacher ID	No. of times expert's occurrence in selected clusters out of 237 times	% of expert's occurrence in selected cluster at α -Cut 0.98
T10	219	92.41
T01	218	91.98
T02	216	91.14
T08	205	86.5
T15	197	83.12
T14	196	82.7
T13	196	82.7
T11	191	80.59
T18	182	76.79
T12	179	75.53
T03	168	70.89
T17	155	69.62
T04	162	68.35
T20	161	67.93
T05	158	66.67
T07	153	64.56
T16	141	59.49
T19	135	56.96
T06	98	41.35
T09	75	31.65

Figure 2: Histogram of 20 Experts for Belonging to Partition Having Maximum Cardinality at α -Cut ($\alpha = 0.98$) for All 237 Students.**Table 10:** Summary of Fleiss κ Statistics for Agreement between All 20 Experts vs. Selected 11 Experts vs. Rejected Nine Experts.

Fleiss κ statistics	Data set of all 20 teachers	Data set of selected 11 teachers	Data set of rejected 9 teachers
P_BAR	0.475211	0.4936517	0.4823996
P_e	0.141801	0.1413933	0.145761
κ Coefficient	0.3884996	0.4102675	0.3940801

art, especially in expert selection, is a distant dream. In the absence of a standard universal approach, it was considered appropriate to use the authenticated method of κ coefficient for the validation of similar experts having high inter-rater agreement. The κ coefficient is invariably used in medical imaging and other related

fields. The expert identification problem to justifiably distinguish between a *reliable* and an *unreliable* expert by *lay people* is highly debated in recent social epistemology [19]. The method for the selection of similar experts presented in this paper by a layperson with only esoteric knowledge can be used in several areas of science and technology in general, and selection of similar examiners in the education system in particular. The software implementation of this approach is done in MATLAB R2008a and Microsoft Access 2010.

6 Concluding Remarks

The fair and unbiased evaluation of students' academic performance depends on the reliability in experts' judgments. A combination of the *epistemic uncertainty* in experts' subjective judgment and *aleatory uncertainty* in object's features is the basis of the formalism presented in this paper. The authors have outline the formalism for the selection of experts based on their subjective judgment using fuzzy relational calculus. The identification of similar experts/examiners in performance evaluation relates to mostly fuzzy classification based on fuzzy similarity relation. The authors propose to extend this concept to large data sets in order to ensure its credibility. The method is somewhat like hierarchical clustering in multivariate data analysis. Are experts similar in their thinking? This is one of the questions that need to be answered while implementing a fuzzy inference system. The authors have looked into this issue and will incorporate the knowledgebase of the identified 11 experts in the fuzzy expert system to infer the performance of students in linguistic terms with a degree of certainty.

Bibliography

- [1] R. H. Ashton, Combining the judgments of experts: how many and which ones?, *Organ. Behav. Hum. Dec.* **38** (1986), 405–414.
- [2] A. H. Ashton and R. H. Ashton, Aggregating subjective forecasts: some empirical results, *Manage. Sci.* **31** (1985), 1499–1508.
- [3] S. M. Bai and S. M. Chen, Automatically constructing grade membership functions for student's evaluation for fuzzy grading systems, in: *Proceeding World Automation Congress (WAC)*, Budapest, Hungary, 2006.
- [4] R. Biswas, An application of fuzzy sets in student's evaluation, *Fuzzy Set. Syst.* **74** (1995), 187–194.
- [5] J. S. Carroll and J. W. Payne, The psychology of parole decision processes: a joint application of attribution theory and information-processing psychology, in: J. S. Carroll and J. W. Payne, eds., *Cognition and Social Psychology*, pp. 13–32, Erlbaum, Hillsdale, NJ, 1976.
- [6] S. M. Chen and C. H. Lee, New methods for student's evaluating using fuzzy set, in: *Proceedings of International Conference on Artificial Intelligence*, 1996.
- [7] S.-M. Chen and H.-Y. Wang, Evaluating student's answer script based on interval valued fuzzy grade sheets, *Expert Syst. Appl.* **36** (2009), 9839–9846.
- [8] R. T. Clemens and R. L. Winkler, Combining probability distributions from experts in risk analysis, *Risk Anal.* **19** (1999), 187–203.
- [9] A. A. DeSmet, D. G. Fryback and J. R. Thornbury, A second look at the utility of radiographic skull examination for trauma, *Am J. Radiol.* **132** (1978), 95–99.
- [10] E. Ebbesen and V. Konecni, Decision making and information integration in the courts: the setting of bail, *J. Pers. Soc. Psychol.* **32** (1975), 805–821.
- [11] H. Einhorn, Expert judgment: some necessary conditions and an example, *J. Appl. Psychol.* **59** (1974), 562–571.
- [12] K. A. Ericsson and N. Charness, Expert performance – its structure and acquisition, *Am. Psychol.* **49** (1994), 725–747.
- [13] J. L. Fleiss, *Statistical methods for rates and proportions*, 2nd ed., John Wiley, New York, pp. 38–46, 1981.
- [14] J. L. Fleiss and J. Cohen, The equivalence of weighted κ and the intraclass correlation coefficient as measures of reliability, *Educ. Psychol. Meas.* **33** (1973), 613–619.
- [15] J. E. Foss, W. R. Wright and R. H. Coles, Testing the accuracy of field textures, *Soil Sci. Soc. Am. Pro.* **39** (1975), 800–802.
- [16] A. I. Goldman, *Knowledge in Social World*, Oxford, pp. 276–271, 1999.
- [17] A. I. Goldman, Experts: which ones should you trust?, *Philos. Phenomen. Res.* **63** (2001), 85–110.
- [18] M. R. Grier, Decision making about patient care, *Nurs. Res.* **25** (1976), 105–110.

- [19] M. Hoffmann, How to identify moral experts? An application of Goldman's criteria for expert identification to domain of morality, *Analyse und Kritik* **34** (2012), 299–313.
- [20] C.-H. Hsies, Evaluating students answer scripts with fuzzy arithmetic, *Tamsui Oxford J. Management Sci.* **12** (1996), 15–25.
- [21] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic Theory and Application*, Prentice Hall PTR, Upper Saddle River, NJ, 1995.
- [22] O. Kosheleva, How to make sure that the grading scheme encourages students to learn all the material: fuzzy-motivated solution and its justification, world conference on soft computing, San Francisco State University, 2011.
- [23] B. Kosko, Fuzzy entropy and conditioning, *Inform. Sciences* **40** (1986), 165–174.
- [24] H. S. Lee, Automatic clustering of business processes in business systems, *Eur. J. Oper. Res.* **114** (1999), 354–362.
- [25] H. S. Lee, An optimal algorithm for computing the max-min transitive closure of a fuzzy similarity matrix, *Fuzzy Set. Syst.* **123** (2001), 129–136.
- [26] T. K. Li and S.-M. Chen, A new method for students learning achievement evaluation by automatically generating the weights of attributes with fuzzy reasoning capability, in: *Proceedings of the Eighth International Conference on Machine Learning and Cybernetics*, IEEE, Baoding, 2009.
- [27] R. Libby and R. K. Blashfield, Performance of a composite as a function of the number of judges, *Organ. Behav. Hum. Perform.* **21** (1978), 121–129.
- [28] S. Lichtenstein and B. Fischhoff, Do those who know more also know more about how much they know?, *Organ Behav. Hum. Perf.* **20** (1977), 159–183.
- [29] S. Lichtenstein, B. Fischhoff and L. D. Phillips, Calibration of probabilities: the state of the art to 1980, in: D. Kahneman, P. Slovic and A. Tversky, eds., *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge, UK, 1982.
- [30] S. Makridakis and R. L. Winkler, Averages of forecasts: Some empirical results, *Manage. Sci.* **29** (1983), 987–996.
- [31] T. McGrew, How foundationalists do crossword puzzles, *Philos. Stud.* **96** (1999), 333–350.
- [32] J. R. Nolan, An expert fuzzy classification system for supporting the grading of student writing samples, *Expert Syst. Appl.* **15** (1998), 59–68.
- [33] A. O'Hagan, Eliciting expert beliefs in substantial practical applications, *The Statistician* **47** (1998), 21–35.
- [34] S. Oskamp, The relationship of clinical experience and training methods to several criteria of clinical prediction, *Psychol. Monogr.* **76** (1962), 1–27.
- [35] S. Oskamp, Overconfidence in case-study judgments, *J. Consult. Psychol.* **29** (1965), 261–265.
- [36] S. Raha, N. R. Pal and K. S. Ray, Similarity based approximate reasoning: methodology and application, *IEEE T. Syst. Man. Cy. A.* **32** (2002), 541–547.
- [37] T. J. Ross, *Fuzzy Logic with engineering applications*, 3rd ed. John Wiley and Sons Ltd., UK, 2010.
- [38] I. Saleh and S. Kim, A fuzzy system for evaluating students learning achievement, *Expert. Syst. Appl.* **36** (2009), 6236–6243.
- [39] S. S. Salunkhe, Y. V. Joshi and A. Deshpande, Fuzzy similarity measures as a basis for the selection of examiners, *J. Fuzzy Set Valued Anal.* **2015** (2015), 1–14.
- [40] O. R. Scholz, Experts: What they are and how we recognize them – A discussion of Alvin Goldman's view, *Grazer Philosophische Studien* **79** (2009), 187–205.
- [41] J. Shanteau, Competence in experts: the role of task characteristics, *Organ Behav. Hum. Dec.* **53** (1992), 252–266.
- [42] M. R. Steenbergen and G. Marks, Evaluating expert judgments, *Eur. J. Polit. Res.* **46** (2007), 347–366.
- [43] S. Tamura, S. Higuchi and K. Tanaka, Pattern classification based on fuzzy relations, *IEEE T. Syst. Man Cyb.* **SMC-1** (1971), 61–66.
- [44] D. Trumbo, C. Adams, M. Milner and L. Schipper, Reliability and accuracy in the inspection of hard red winter wheat, *Cereal Sci. Today* **7** (1962), 62–71.
- [45] H.-Y. Wang and S.-M. Chen, New methods for evaluating the answer scripts of students using fuzzy sets, IEA/AIE 2006, LNAI 4031, 2006.
- [46] H.-Y. Wang and S.-M. Chen, Evaluating students answer scripts using vague values, *Appl. Intell.* **28** (2008), 183–193.
- [47] P. Williams, The use of confidence factors in forecasting, *B. Am. Meteorol. Soc.* **32** (1951), 279–281.
- [48] R. L. Winkler and R. T. Clemen, Experts vs. multiple methods: combining correlation assessments, *Decision Analysis* **1** (2004), 167–176.
- [49] R. L. Winkler and S. Makridakis, The combination of forecasts, *J. R. Stat. Soc., Series A*, 146, Pt. 2, (1983) 150–157.
- [50] L. A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning, *Inform. Sciences* **8** (1965), 199–249.
- [51] L. A. Zadeh, Similarity relations and fuzzy orderings, *Inform. Sciences* **3** (1971), 177–200.
- [52] H. J. Zimmermann, *Fuzzy set theory and its application*, 4th ed. Springer, USA, 2001.