Aberystwyth University

Rough-fuzzy rule interpolation

Chen, Chengyuan; MacParthalain, Neil; Li, Ying; Price, Christopher; Quek, Chai; Shen, Qiang

Published in:
Information Sciences

DOI:
10.1016/j.ins.2016.02.036

Publication date:
2016

Citation for published version (APA):
Chen, C., MacParthalain, N., Li, Y., Price, C., Quek, C., & Shen, Q. (2016). Rough-fuzzy rule interpolation. *Information Sciences*, *351*, 1-17. https://doi.org/10.1016/j.ins.2016.02.036

Information Sciences 351 (2016) 1–17

Contents lists available at ScienceDirect

Information Sciences

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

Rough-fuzzy rule interpolation

Chengyuan Chen[a], Neil Mac Parthaláin[b], Ying Li[c], Chris Price[b], Chai Quek[d], Qiang Shen[b,*]

[a] School of Electrical and Information Engineering, Chongqing University of Science and Technology, Chongqing 401331, China
[b] Department of Computer Science, Institute of Mathematics, Physics and Computer Science, Aberystwyth University, Aberystwyth SY23 3DB, UK
[c] School of Computer Science, Northwestern Polytechnical University, Xian 710129, China
[d] School of Computer Engineering, Nanyang Technological University, 637457, Singapore

ARTICLE INFO

Article history:
Received 1 July 2015
Revised 7 January 2016
Accepted 18 February 2016
Available online 27 February 2016

Keywords:
Fuzzy rule interpolation
Rough-fuzzy sets
Transformation-based interpolation

ABSTRACT

Fuzzy rule interpolation forms an important approach for performing inference with systems comprising sparse rule bases. Even when a given observation has no overlap with the antecedent values of any existing rules, fuzzy rule interpolation may still derive a useful conclusion. Unfortunately, very little of the existing work on fuzzy rule interpolation can conjunctively handle more than one form of uncertainty in the rules or observations. In particular, the difficulty in defining the required precise-valued membership functions for the fuzzy sets that are used in conventional fuzzy rule interpolation techniques significantly restricts their application. In this paper, a novel rough-fuzzy approach is proposed in an attempt to address such difficulties. The proposed approach allows the representation, handling and utilisation of different levels of uncertainty in knowledge. This allows transformation-based fuzzy rule interpolation techniques to model and harness additional uncertain information in order to implement an effective fuzzy interpolative reasoning system. Final conclusions are derived by performing rough-fuzzy interpolation over this representation. The effectiveness of the approach is illustrated by a practical application to the prediction of diarrhoeal disease rates in remote villages. It is further evaluated against a range of other benchmark case studies. The experimental results confirm the efficacy of the proposed work.

1. Introduction

The compositional rule of inference [1] offers an effective mechanism for performing fuzzy inference with dense rule bases. Given such a rule base and also an observation that is at least partially covered by that rule base, a conclusion can be inferred from certain rules that intersect with the observation. However, for the cases where a fuzzy rule base contains ‘gaps’ (termed: sparse rule base [2]), if a given observation has no overlap with the antecedent values of any rule, conventional

* Corresponding author.
E-mail addresses: ccy@cqust.edu.cn (C. Chen), ncm@aber.ac.uk (N.M. Parthaláin), lybyp@nwpu.edu.cn (Y. Li), cjp@aber.ac.uk (C. Price), ashcquek@ntu.edu.sg (C. Quek), qqs@aber.ac.uk (Q. Shen).

http://dx.doi.org/10.1016/j.ins.2016.02.036

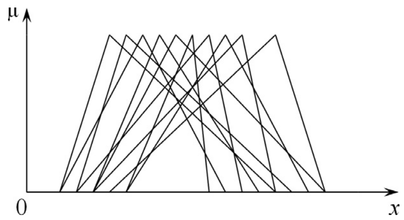

Fig. 1. Different membership functions for a common underlying concept perceived by different people.

fuzzy inference methods cannot derive a conclusion. Fortunately, using fuzzy rule interpolation (FRI) [3,4], certain useful conclusions may still be obtained.

The application of traditional FRI methods may lead to abnormal fuzzy conclusions, however. One particular issue is that the convexity of the derived fuzzy values is not guaranteed [5,6], but convexity is often a crucial requirement for fuzzy inference in order to attain improved interpretability of the results. A number of significant extensions to the original FRI methods have been proposed in literature in an attempt to address this issue, including [7–13]. In particular, the scale and move transformation-based FRI approach (abbreviated to T-FRI hereafter) [14,15] and its generalisation [16,17] can handle interpolation and extrapolation which involve multiple fuzzy rules, where each rule consists of multiple antecedents. Such work also guarantees the uniqueness, as well as the normality and convexity of the interpolated conclusion. This approach has recently been further enhanced with an adaptive mechanism such that appropriate chaining of fuzzy interpolative inferences can be performed [18]. Further development has also been reported that allows the fuzzy rule interpolation and extrapolation to be performed in a backward manner, i.e., from rule consequence to antecedent variables [19].

The aforementioned FRI techniques provide a basic means for dealing with and interpreting uncertainty in rule-based reasoning by exploiting the expressive power of fuzzy sets [20]. However, there is little work in the area of FRI that can handle uncertainty in fuzziness itself. This is because these approaches are implemented using conventional fuzzy set theory as a basis for the underlying representations [21]. Whilst membership functions play an important role in defining fuzzy sets, it is sometimes extremely difficult, if not impossible, to precisely define such membership functions. More generally, there may be different types of uncertainty in fuzzy rule-based systems that need to be captured or modelled [22]: (1) The linguistic variables that are used in the antecedents and consequences of the given rules may be indiscernible. (2) The interpretation of the values of the underlying linguistic variables may be vague, e.g., the same word can mean different things to different people. (3) An element can belong to a fuzzy set with a given degree, but that degree of belonging may itself be uncertain. (4) The obtained rules may be inconsistent when individual views are provided from a group of experts. (5) Observations attainable by inexact knowledge may be noisy and therefore randomly distributed.

Most of these types of uncertainty can be difficult to deal with if crisp membership functions of the fuzzy sets have to be determined. For instance, there are certain extreme weather conditions that would be considered to be *cold* by all people, but other less extreme conditions may still be considered to be *cold* only by certain individuals. The membership functions for different people may therefore be different, depending on their perception, preference, experience, etc. This is shown in Fig. 1. That is, both similarities and differences may exist in defining a given concept. Therefore, the representation of a concept should satisfy the requirements of not only the imprecise description but also both *common* and *individual* perceptions. In this case, using only the membership values of a conventional (type-1) fuzzy set may not be adequate to capture, reflect, and model the given concept. When faced with such higher order uncertainty, which is essentially the uncertainty of evaluation about uncertainty, a general approach would be to simply ignore this higher-level information. However, an obvious drawback of this is that valuable information may be lost about both the concept that is being modelled and the impact that uncertain information may have upon that concept. This, in turn, may lead to unacceptable inference conclusions. Alternative representations are therefore required in order to achieve a better understanding of the concept and manipulate these different levels of uncertainty such that they can be handled appropriately. Thus, it is desirable to develop a new model to represent the membership functions of fuzzy sets, in order to provide a better means of addressing uncertainty in FRI.

The concept of rough sets [23] was originally proposed as a mathematical tool to deal with incomplete or imperfect data and knowledge in information systems. A rough set is itself an approximation of a vague concept by a pair of precise sets, called lower and upper approximations [24,25]. The lower approximation contains all of those objects which *definitely* belong to the set that denotes the given concept, and the upper approximation contains all of those objects which *possibly* belong to that set. As such, rough sets offer a distinct and complementary approach to fuzzy sets in supporting approximate reasoning. Inspired by this observation, it is potentially useful to integrate rough and fuzzy techniques in order to improve the ability to handle uncertainty. This paper proposes such an approach to rough-fuzzy set-based rule interpolation. A specification of rough-fuzzy sets is introduced to describe a particular type of higher order uncertainty, which is characterised by the lower and upper approximation membership functions. This approach facilitates the representation of uncertain fuzzy set membership functions with rough-fuzzy approximations, thereby improving the flexibility of rule interpolation in handling different levels of uncertainty that may be present in sparse rule bases and observations. The work reflects the intuition that the more useful information available to the interpolation process, the better the interpolated results.

The rest of the paper is structured as follows. Section 2 introduces the basic concepts that are used for the development of rough-fuzzy rule interpolation. Section 3 describes the main steps of the proposed rough-fuzzy rule interpolation algorithm. Section 4 presents a realistic problem application that shows the potential for use of the proposed approach for the prediction of diarrhoeal diseases in remote villages. Section 5 provides further examples which illustrate the applicability of the approach and further demonstrate the efficacy of the work. The paper is concluded in Section 6, including suggestions for further development.

2. Rough-fuzzy sets and their representative values

2.1. Rough-fuzzy sets

The starting point for the proposed approach is the ability to represent complicated uncertain knowledge in an effort to perform FRI. When exact membership values are no longer suitable for describing the underlying uncertainty, it is desirable to utilise an alternative higher-order representation. A *rough-fuzzy representation* offers such characteristics, with the first order representation embedded within it. This differs from the concept of conventional rough sets, which is characterised by the lower and upper approximations whose elements are of full memberships [23]. If however, the rough-fuzzy information or data degenerates to the first order representation, the computational mechanism that deals with rough-fuzzy interpolation should also naturally degenerate to the corresponding first order calculus.

Let $I = (\mathbb{U}, \mathbb{A})$ be an information system, where $\mathbb{U}$ is a non-empty set of finite objects (namely, the underlying universe of discourse) and $\mathbb{A}$ is a non-empty finite set of attributes such that $a : \mathbb{U} \to V_a$ for every $a \in \mathbb{A}$ with V_a being the domain that attribute a takes values from. With any $P \subseteq \mathbb{A}$ there is a crisp equivalence relation $IND(P)$ [23]:

$$IND(P) = \{(x, y) \in \mathbb{U}^2 \mid \forall a \in P,\ a(x) = a(y)\} \tag{1}$$

If $(x,\ y) \in IND(P)$, then x and y are indiscernible by attributes from P. The equivalence class with respect to such an indiscernibility relation defined on P is denoted by $[x]_P$, $x \in \mathbb{U}$.

Let $X \subseteq \mathbb{U}$, X be approximated using only the information contained within P by constructing the P-*lower* and P-*upper* approximations of X [23]:

$$\begin{aligned} \underline{P}X &= \{x | [x]_P \subseteq X\} \\ \overline{P}X &= \{x | [x]_P \cap X \neq \emptyset\} \end{aligned} \tag{2}$$

The tuple $\langle \underline{P}X, \overline{P}X \rangle$ is called a rough set.

Definition 2.1. With any $P \subseteq \mathbb{A}$, an alternative equivalence relation $IND(P)$ to the traditional one of Eq. (1) can be defined by

$$IND(P) = \{(x, y) \in \mathbb{U}^2 \mid \forall F_g \in P,\ F_g(x) \in C,\ F_g(y) \in C\} \tag{3}$$

where F_g, $g \in \{1, \ldots, G\}$, are fuzzy sets that jointly define a particular concept C in X, $X \subseteq \mathbb{U}$.

Eq. (3) expresses the equivalence relation between the memberships of x and y to different fuzzy sets of given concept. Using this equivalence relation, the lower and upper approximations for a single C in X can be redefined as follows.

Definition 2.2. Let $IND(P)$ be an equivalence relation on $\mathbb{U}$ and F_g, $g \in \{1, \ldots,\ G\}$, be fuzzy sets in C ($C \in X$), the lower and upper approximations are a pair of fuzzy sets with membership functions defined by the following, respectively:

$$\begin{aligned} \mu_{\underline{P}C}(x \in [x]_P) &= \inf\{\mu_{F_g}(x), g \in \{1, \ldots, G\} | x \in [x]_P\} \\ \mu_{\overline{P}C}(x \in [x]_P) &= \sup\{\mu_{F_g}(x), g \in \{1, \ldots, G\} | x \in [x]_P\} \end{aligned} \tag{4}$$

The tuple $\langle \underline{P}X, \overline{P}X \rangle$ is called a rough-fuzzy (RF) set (which differs from the alternative use of this term in the literature [26] due to the parallel development of these related but different concepts).

Remark 2.1. As discussed previously, similar but different fuzzy sets, which are considered to belong to an equivalence class, may be obtained in describing a given concept. Therefore, the lower and upper approximations of an RF set are defined on the basis of those fuzzy sets which are known to share such an equivalence relation.

Definition 2.3. Let $\mathbb{U}$ be the universe, an RF set $\tilde{A}$ on $\mathbb{U}$ is herein denoted by the lower approximation (LA) $\tilde{A}^L$ and the upper approximation (UA) $\tilde{A}^U$ such that

$$\tilde{A} = \langle x, [\mu^L_{\tilde{A}}(x), \mu^U_{\tilde{A}}(x)] \rangle = \langle \tilde{A}^L, \tilde{A}^U \rangle,\ \ \forall x \in \mathbb{U} \tag{5}$$

where $0 \le \mu^L_{\tilde{A}}(x) \le \mu^U_{\tilde{A}}(x) \le 1$, and the lower and upper approximations are two conventional fuzzy sets, namely, two first order fuzzy sets.

Remark 2.2. The closer the shapes of $\tilde{A}^L$ and $\tilde{A}^U$ are, the less uncertain the information contained within $\tilde{A}$ is. When $\tilde{A}^L$ coincides with $\tilde{A}^U$, the RF set degenerates to a conventional fuzzy set, i.e., $\mu^L_{\tilde{A}}(x) = \mu^U_{\tilde{A}}(x)$, $\forall x \in \mathbb{U}$.

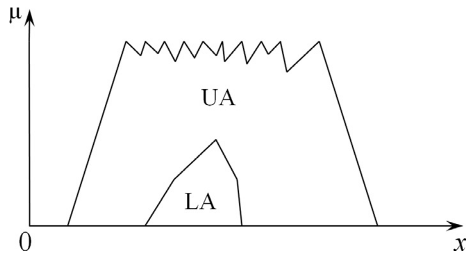

Fig. 2. An RF set corresponding to the situation depicted by Fig. 1.

Reconsider the situation shown in Section 1, where different people may interpret the same concept differently. As reflected in Fig. 1, it is difficult to describe this situation using conventional fuzzy sets. However, RF sets can be adopted to represent this uncertain concept by exploiting the two approximations. In particular, the LA indicates the intersection amongst the regions that are agreed by individuals, while the UA indicates the union of the regions that are given by at least one person, as shown in Fig. 2. RF sets therefore utilise LAs and UAs to express the higher order uncertainty involved in describing a piece of knowledge or data.

2.2. Basic notions

An important concept to introduce is the 'less than' relation between two fuzzy sets [3]. An ordinary set A_1 is said to be *less than* another ordinary fuzzy set A_2, denoted by $A_1 \prec A_2$, if $\forall\alpha \in (0,1]$, the following conditions hold:

$$\inf\{A_{1\alpha}\} < \inf\{A_{2\alpha}\}, \ \sup\{A_{1\alpha}\} < \sup\{A_{2\alpha}\} \tag{6}$$

where $A_{1\alpha}$ and $A_{2\alpha}$ are the α-cut sets of A_1 and A_2, respectively, $\inf\{A_{i\alpha}\}$ is the infimum of $A_{i\alpha}$, and $\sup\{A_{i\alpha}\}$ is the supremum of $A_{i\alpha}$, $i = 1, 2$.

Definition 2.4. An RF set $\tilde{A}_1$ is said to be *less than* another RF set $\tilde{A}_2$, denoted as $\tilde{A}_1 \precsim \tilde{A}_2$, if and only if

$$\tilde{A}_1^L \prec \tilde{A}_2^L, \quad \tilde{A}_1^U \prec \tilde{A}_2^U \tag{7}$$

From this, the notion of neighbouring rules involving RF sets can be defined.

Definition 2.5. Two RF rules

$$R_1 : \text{If } x_1 \text{ is } \tilde{A}_{11}, x_2 \text{ is } \tilde{A}_{12}, \ldots, x_M \text{ is } \tilde{A}_{1M}, \quad \text{then } y \text{ is } \tilde{B}_1$$
$$R_2 : \text{If } x_1 \text{ is } \tilde{A}_{21}, \quad x_2 \text{ is } \tilde{A}_{22}, \ldots, \quad x_M \text{ is } \tilde{A}_{2M}, \quad \text{then } y \text{ is } \tilde{B}_2$$

are said to be *neighbouring rules* if and only if: (1) $\tilde{A}_{1j} \precsim \tilde{A}_{2j}$ or $\tilde{A}_{2j} \precsim \tilde{A}_{1j}$, $j \in \{1, \ldots, M\}$ (where M is the number of antecedent variables in both rules); and (2) there is no individual rule "If x_1 is $\tilde{A}'_1$, x_2 is $\tilde{A}'_2$, …, x_M is $\tilde{A}'_M$, then y is $\tilde{B}'$" such that $\tilde{A}_{1j} \precsim \tilde{A}'_j \precsim \tilde{A}_{2j}$ if $\tilde{A}_{1j} \precsim \tilde{A}_{2j}$, or $\tilde{A}_{2j} \precsim \tilde{A}'_j \precsim \tilde{A}_{1j}$ if $\tilde{A}_{2j} \precsim \tilde{A}_{1j}$, $j \in \{1, \ldots, M\}$.

RF rule interpolation can then be achieved by extending the conventional FRI. In this case, the input and output of an interpolative process are RF sets rather than conventional fuzzy sets.

Definition 2.6. Given an RF rule base and an RF observation, *rough-fuzzy rule interpolation* is a process through which a conclusion from the given observation is obtained by identifying the rules in the rule base which flank the observation and interpolating from those rules.

Note that in the above definition, two rules (e.g., the R_1 and R_2 given previously) are said to flank a given observation [3], say, $O = (\tilde{A}_1^*, \tilde{A}_2^*, \ldots, \tilde{A}_M^*)$, if $\tilde{A}_{1j} \precsim \tilde{A}_j^* \precsim \tilde{A}_{2j}$, or $\tilde{A}_{2j} \precsim \tilde{A}_j^* \precsim \tilde{A}_{1j}$, $j \in \{1, \ldots, M\}$.

2.3. Representative values

In order to support the interpolation of rules involving RF sets following the transformation-based approach [14], the rules which have minimal distances from a given observation need to be selected first. Here, a distance between a given observation and a rule in the rule base is measured on the basis of *representative value* (Rep). The concept of Rep is introduced below.

The Rep value captures important information such as the overall location of an RF set within the definition domain, and is computed and then utilised as the guide to perform subsequent inference during the interpolation process. For simplicity, in this work, it is assumed that only polygonal RF sets are considered; that is, both the lower and the upper approximation are each represented by a polygonal-shaped first order fuzzy set. Note that in the existing T-FRI [15], the Rep(A) of an

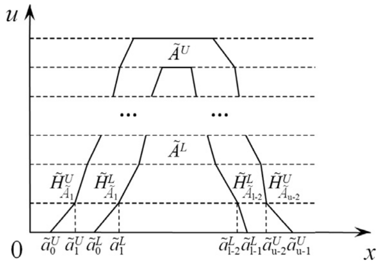

Fig. 3. LA $\tilde{A}^L$ and UA $\tilde{A}^U$ of a polygonal RF set $\tilde{A}$.

ordinary fuzzy set A is defined by the weighted average of the x coordinate values of all odd points a_i, $i \in \{0, \ldots, k-1\}$, such that:

$$\mathrm{Rep}(A) = \sum_{i=0}^{k-1} w_i a_i \tag{8}$$

with $A = (a_0, \ldots, a_{k-1})$ being a polygonal fuzzy set of k odd points, and w_i denoting the weight assigned to the point a_i.

Definition 2.7. Suppose that a polygonal RF set $\tilde{A}$ is given, as shown in Fig. 3, whose lower and upper approximations are: $\tilde{A} = \langle(\tilde{a}^L_0, \ldots, \tilde{a}^L_{l-1}; \tilde{H}^L_{\tilde{A}_1}, \ldots, \tilde{H}^L_{\tilde{A}_{l-2}}), (\tilde{a}^U_0, \ldots, \tilde{a}^U_{u-1}; \tilde{H}^U_{\tilde{A}_1}, \ldots, \tilde{H}^U_{\tilde{A}_{u-2}})\rangle$. The lower and upper Reps $\mathrm{Rep}(\tilde{A}^L)$ and $\mathrm{Rep}(\tilde{A}^U)$ of $\tilde{A}$ are defined by

$$\begin{cases} \mathrm{Rep}(\tilde{A}^L)_x = \sum_{i=0}^{l-1} w^L_i \tilde{a}^L_i \\ \mathrm{Rep}(\tilde{A}^L)_y = \sum_{i=1}^{l-2} w^L_i \tilde{H}^L_{\tilde{A}_i} \end{cases} \quad \begin{cases} \mathrm{Rep}(\tilde{A}^U)_x = \sum_{j=0}^{u-1} w^U_j \tilde{a}^U_j \\ \mathrm{Rep}(\tilde{A}^U)_y = \sum_{j=1}^{u-2} w^U_j \tilde{H}^U_{\tilde{A}_j} \end{cases} \tag{9}$$

where w^V_v ($V \in \{L, U\}$, $v \in \{i, j\}$) is the weight assigned to point $\tilde{a}^V_v$ and its corresponding membership value $\tilde{H}^V_{\tilde{A}_v}$, and x and y denote a certain variable dimension and the corresponding membership distribution, respectively.

In general, different definitions can be adopted for deriving different Rep values. For instance, the simplest case is that all points take the same weight value, i.e., $w^L_i = 1/l$ and $w^U_j = 1/u$. The centre of core can also be used as an alternative. In this case, the Rep may be solely determined by those points with a membership value of 1: $\mathrm{Rep}(\tilde{A}^U)_x = \frac{1}{2}(\tilde{a}^U_{\lceil(l/2)\rceil-1} + \tilde{a}^U_{l-\lceil(l/2)\rceil})$ and $\mathrm{Rep}(\tilde{A}^U)_y = \frac{1}{2}(\tilde{H}^U_{\tilde{A}_{\lceil(l/2)\rceil-1}} + \tilde{H}^U_{\tilde{A}_{l-\lceil(l/2)\rceil}})$. The lower Reps are omitted here, which can be calculated in a similar way involving those points of the maximum membership value. Other alternative definitions can be found in [15]. For a given application, one of these weighting schemes needs to be taken for implementation.

Remark 2.3. In the existing T-FRI, the $\mathrm{Rep}(A)_y$ of a given conventional fuzzy set A is a constant, only the x coordinate is therefore considered as the $\mathrm{Rep}(A)$. However, this is no longer the case in this work due to the introduction of higher order uncertainty, both x and y dimensions must be considered. The calculation of $\mathrm{Rep}(\tilde{A})_y$ follows that used to calculate $\mathrm{Rep}(\tilde{A})_x$ in an effort to maintain consistency.

To distinguish amongst different RF set shapes, the *shape diversity factor f* is also introduced here. The present work defines this concept by following the conventional definition of statistical standard deviation (although this may be defined differently if desired for a particular implementation).

Definition 2.8. The lower and upper shape diversity factors $f^L_{\tilde{A}}$ and $f^U_{\tilde{A}}$ are defined by

$$\begin{cases} f^L_{\tilde{A}} = \sqrt{\frac{\sum_{i=0}^{l-1}(\tilde{a}^L_i - \mathrm{Rep}(\tilde{A}^L)_x)^2}{l}} \\ f^U_{\tilde{A}} = \sqrt{\frac{\sum_{j=0}^{u-1}(\tilde{a}^U_j - \mathrm{Rep}(\tilde{A}^U)_x)^2}{u}} \end{cases} \tag{10}$$

Remark 2.4. A small shape diversity factor implies that the odd points of $\tilde{A}^L$ ($\tilde{A}^U$) tend to be close to those of the lower (upper) Rep. That is, the smaller the shape diversity factor, the smaller the area of the lower (upper) approximation.

To extend the methodology of conventional T-FRI to FRI involving RF sets, a single overall Rep of a given RF set is required. For this, the *weight factor w* of the lower (upper) approximation is defined first, which reflects the relative contribution of the lower (upper) shape diversity in depicting the underlying RF set. The introduction of these lower and upper shape diversity factors helps minimise the possibility that the use of RF sets of different shapes leads to the same overall Rep values.

Definition 2.9. The lower and upper weight factors $w^L_{\tilde{A}}$ and $w^U_{\tilde{A}}$ are defined as the weights of the shape diversity factors, in terms of the areas of the lower and upper approximations, such that:

$$w^V_{\tilde{A}} = \frac{f^V_{\tilde{A}}}{f^L_{\tilde{A}} + f^U_{\tilde{A}}}, \quad V = L, U \tag{11}$$

Remark 2.5. In general, $f^L_{\tilde{A}} + f^U_{\tilde{A}} \neq 0$. If however, $f^L_{\tilde{A}} + f^U_{\tilde{A}} = 0$, i.e., $f^L_{\tilde{A}} = 0$ and $f^U_{\tilde{A}} = 0$, the RF set degenerates to a singleton value, $w^L_{\tilde{A}} = w^U_{\tilde{A}} = 1/2$.

Definition 2.10. The overall Rep of a given RF set $\tilde{A}$ is defined by

$$\text{Rep}(\tilde{A}) = \sum_{V \in \{L,U\}} \left(w^V_{\tilde{A}} \sum_{e \in \{x,y\}} \text{Rep}(\tilde{A}^V)_e \right) \tag{12}$$

where $w^V_{\tilde{A}}$ is the weight assigned to $\text{Rep}(\tilde{A}^V)$ of $\tilde{A}^V$, $V \in \{L, U\}$.

3. Rough-fuzzy rule interpolation

3.1. Selection of closest N rules

As with conventional FRI approaches, in general, multiple rules with multiple antecedents need to be considered in order to obtain an interpolated conclusion. For this, the first step that needs to be considered is to choose the closest N ($N \geq 2$) rules from the rule base with respect to the given observation. A *distance* measure is thus utilised to measure the proximity of the rules by exploiting such Rep values that capture specific information embedded in RF sets.

Without any loss of generality, suppose that there are n RF rules in an RF rule base. A rule R_i, an observation O and the conclusion C are represented by the following, respectively:

$$R_i : \text{If } x_1 \text{ is } \tilde{A}_{i1}, \ldots, x_j \text{ is } \tilde{A}_{ij}, \ldots, x_M \text{ is } \tilde{A}_{iM}, \text{then } y \text{ is } \tilde{B}_i$$
$$O : x_1 \text{ is } \tilde{A}^*_1, \ldots, x_j \text{ is } \tilde{A}^*_j, \ldots, x_M \text{ is } \tilde{A}^*_M$$
$$C : y \text{ is } \tilde{B}^*$$

where $\tilde{A}_{ij}$ denotes the RF set that acts as the value of the jth antecedent of R_i, $\tilde{A}^*_j$ is the observation of the variable x_j, $\tilde{B}^*$ is the desired interpolated conclusion, and $\tilde{B}_i$ denotes the consequent RF set of R_i with $j \in \{1, \ldots, M\}$, with M being the number of antecedent variables.

Definition 3.1. The distance d_{ij} between a pair of RF sets $\tilde{A}_{ij}$ and $\tilde{A}^*_j$ is defined as follows:

$$d_{ij} = d(\tilde{A}_{ij}, \tilde{A}^*_j) = d(\text{Rep}(\tilde{A}_{ij}), \text{Rep}(\tilde{A}^*_j)) \tag{13}$$

where $d(., .)$ is herein computed using the Euclidean distance metric (though any other distance metric may be used as an alternative).

Definition 3.2. The distance d_i between the rule R_i and the observation O is deemed to be the average of the distances between the RF sets of each rule antecedent and the corresponding variable in O:

$$d_i = \sqrt{\sum_{j=1}^{M} {d'_{ij}}^2}, \quad d'_{ij} = \frac{d_{ij}}{\max_j - \min_j} \tag{14}$$

where $\max_j$ and $\min_j$ are the maximum and minimum value in the domain of the variable x_j, $j \in \{1, \ldots, M\}$. The use of normalised distance measure d'_{ij} is to ensure that the resulting distances are compatible with each other over different domains.

Given the above definition, the distances between a given observation and all rules in the rule base can be calculated. The N rules which have minimal distances are chosen as the closest N rules with respect to the given observation. The choice of a larger N will help consider a wider range of neighbouring rules in performing interpolation, thereby more likely to result in global results but requiring significantly more computation. On the contrary, the choice of a smaller N will tend to take into account only neighbouring rules and hence involve less computation time. Since FRI is in general used to derive an approximate result in the first place, in practical application, N can be chosen to be 2. This is the case for conventional rule interpolation also. However, in the following theoretical development to maintain generality, the number of closest rules is set to N ($N \geq 2$) unless otherwise stated.

3.2. Construction of intermediate rule

As with a number of conventional FRI approaches, the approach in this work is developed following the principle of analogical reasoning [27]. First, an artificially created intermediate rule is interpolated such that the antecedent of the intermediate rule is as 'close' to the given observation as possible. Then, a conclusion is worked out from the given observation by firing this generated intermediate rule through a certain analogical reasoning mechanism.

Definition 3.3. Suppose that N closest rules are chosen with respect to a given observation. These rules are represented as R_i, $i \in \{1, \ldots, N\}$, each having M antecedent variables $\tilde{A}_{ij}$, $j \in \{1, \ldots, M\}$, and are used to derive the intermediate rule. Let $w_{\tilde{A}_{ij}}$ denote the weight to which the jth antecedent of the ith closest rule contributes to the emerging intermediate rule, which is defined as the reciprocal of the corresponding distance measure:

$$w_{\tilde{A}_{ij}} = \frac{1}{d_{ij}} = \frac{1}{d(\tilde{A}_{ij}, \tilde{A}^*_j)} \tag{15}$$

where $\tilde{A}^*_j$ denotes the observed RF set of antecedent variable j. The normalised weight $w'_{\tilde{A}_{ij}}$ is then defined by

$$w'_{\tilde{A}_{ij}} = \frac{w_{\tilde{A}_{ij}}}{\sum_{i=1}^{N} w_{\tilde{A}_{ij}}} \tag{16}$$

Remark 3.1. This definition reflects the intuition that the larger the distance is, the less relevant the corresponding attribute is to the observation. In general, $d_{ij} \neq 0$. If however, $d_{ij} = 0$, then $\mathrm{Rep}(\tilde{A}_{ij}) = \mathrm{Rep}(\tilde{A}^*_j)$. In this case, the observation is considered to be identical to the corresponding antecedent of the rule R_i, in terms of their Rep values. Thus, $w_{\tilde{A}_{ij}}$ is set to 1 for the identical cases with the rest set to 0.

The antecedent $\tilde{A}^{IFT}_j$ of the intermediate rule is constructed from the antecedents of the identified closest rules. A process *shift* is then utilised to modify $\tilde{A}^{IFT}_j$ so that the antecedent of the intermediate rule will have the same Rep as $\tilde{A}^*_j$:

$$\tilde{A}'_j = \tilde{A}^{IFT}_j + \delta_{\tilde{A}_j}(\max_j - \min_j), \quad \tilde{A}^{IFT}_j = \sum_{i=1}^{N} w'_{\tilde{A}_{ij}} \tilde{A}_{ij} \tag{17}$$

where $\delta_{\tilde{A}_j}$ is a constant defined by

$$\delta_{\tilde{A}_j} = \frac{\mathrm{Rep}(\tilde{A}^*_j) - \mathrm{Rep}(\tilde{A}^{IFT}_j)}{\max_j - \min_j} \tag{18}$$

The consequence of the intermediate rule $\tilde{B}'$ is calculated by analogy to the computation of the antecedent, such that:

$$\tilde{B}' = \tilde{B}^{IFT} + \delta_{\tilde{B}}(\max - \min), \quad \tilde{B}^{IFT} = \sum_{i=1}^{N} w'_{\tilde{B}_i} \tilde{B}_i \tag{19}$$

where $\tilde{B}^{IFT}$ is the consequence of the intermediate fuzzy rule, max and min are the maximum and minimum values within the domain of the consequent variable, $w'_{\tilde{B}_i}$ and $\delta_{\tilde{B}}$ are the means of $w'_{\tilde{A}_{ij}}$ and $\delta_{\tilde{A}_j}$, $i \in \{1, \ldots, N\}$, $j \in \{1, \ldots, M\}$, respectively, which are defined by

$$w'_{\tilde{B}_i} = \frac{1}{M} \sum_{j=1}^{M} w'_{\tilde{A}_{ij}}, \quad \delta_{\tilde{B}} = \frac{1}{M} \sum_{j=1}^{M} \delta_{\tilde{A}_j} \tag{20}$$

3.3. Interpolation through similarity-constrained transformations

The aforementioned artificially constructed intermediate rule is derived from the chosen closest rules with respect to an observation. It can be used to perform inference without further reference to its originals. Suppose that a certain degree of similarity between the antecedent part of this rule and the observation is established, it is intuitive to require that its consequent part and the eventual conclusion should attain the same similarity degree. That is, for an intermediate rule: "If x_1 is $\tilde{A}'_1$, $\ldots$, x_j is $\tilde{A}'_j$, $\ldots$, x_M is $\tilde{A}'_M$, then y is $\tilde{B}'$", and a given observation $O = (\tilde{A}^*_1, \ldots, \tilde{A}^*_j, \ldots, \tilde{A}^*_M)$, the shape distinguishability between $\tilde{B}'$ and the interpolated consequence $\tilde{B}^*$ is analogous to the combination of the shape distinguishabilities between $\tilde{A}'_j$ and $\tilde{A}^*_j$, $j = 1, 2, \ldots, M$. In order to ensure this, the following three transformations are designed.

Note that all three transformations are separately implemented on each dimension and separately calculated on each of the lower and upper bounds. However, the underlying computational mechanisms are identical. For presentational simplicity, the description of these transformations is given without the subscript j and the superscript L or U.

3.3.1. Scale transformation

Consider the lower (upper) approximation of $\tilde{A}'$ and that of $\tilde{A}^*$, respectively represented as $\tilde{A}' = (\tilde{a}'_0, \ldots, \tilde{a}'_{k-1}; \tilde{H}'_{\tilde{A}_1}, \ldots, \tilde{H}'_{\tilde{A}_{k-2}})$ and $\tilde{A}^* = (\tilde{a}^*_0, \ldots, \tilde{a}^*_{k-1}; \tilde{H}^*_{\tilde{A}_1}, \ldots, \tilde{H}^*_{\tilde{A}_{k-2}})$. The following parameters, termed the scale rates s_p $(p = 0, \ldots, \lfloor (k/2) \rfloor - 1)$ rescale the pth support of $\tilde{A}'$ in order to approximate the corresponding support of $\tilde{A}^*$:

$$s_p = \frac{\tilde{a}^*_{k-p-1} - \tilde{a}^*_p}{\tilde{a}'_{k-p-1} - \tilde{a}'_p} \tag{21}$$

From these scale rates, the following scale ratios $\mathbb{S}_q$ $(q = 1, \ldots, \lfloor (k/2) \rfloor - 1)$ modify the rescaled qth support of $\tilde{A}'$ to further approximate the corresponding support of $\tilde{A}^*$ such that the resulting RF set $\tilde{A}'$ is of the same scale as that of $\tilde{A}^*$:

$$\mathbb{S}_q = \begin{cases} \dfrac{\frac{\tilde{a}^*_{k-q-1} - \tilde{a}^*_q}{\tilde{a}^*_{k-q} - \tilde{a}^*_{q-1}} - \frac{\tilde{a}'_{k-q-1} - \tilde{a}'_q}{\tilde{a}'_{k-q} - \tilde{a}'_{q-1}}}{1 - \frac{\tilde{a}'_{k-q-1} - \tilde{a}'_q}{\tilde{a}'_{k-q} - \tilde{a}'_{q-1}}} & \text{if } s_q \geq s_{q-1} \\ \dfrac{\frac{\tilde{a}^*_{k-q-1} - \tilde{a}^*_q}{\tilde{a}^*_{k-q} - \tilde{a}^*_{q-1}} - \frac{\tilde{a}'_{k-q-1} - \tilde{a}'_q}{\tilde{a}'_{k-q} - \tilde{a}'_{q-1}}}{\frac{\tilde{a}'_{k-q-1} - \tilde{a}'_q}{\tilde{a}'_{k-q} - \tilde{a}'_{q-1}}} & \text{if } s_{q-1} > s_q \end{cases} \tag{22}$$

From this, by imposing the required similarities, the corresponding scale rates s'_p that will help rescale the pth support of $\tilde{B}'$ into the emerging $\tilde{B}^*$ can be obtained such that

$$s'_p = \begin{cases} s_p & \text{if } p = 0 \\ \dfrac{s'_{p-1}(s_p - s_{p-1})\left(\frac{\tilde{b}'_{k-p} - \tilde{b}'_{p-1}}{\tilde{b}'_{k-p-1} - \tilde{b}'_p} - 1\right)}{s_{p-1}\left(\frac{\tilde{a}'_{k-p} - \tilde{a}'_{p-1}}{\tilde{a}'_{k-p-1} - \tilde{a}'_p} - 1\right)} + s'_{p-1} & \text{if } s_p \geq s_{p-1},\ p > 0 \\ \dfrac{s'_{p-1} s_p}{s_{p-1}} & \text{if } s_{p-1} > s_p,\ p > 0 \end{cases} \tag{23}$$

The above shows only the situation where one antecedent variable is considered (for either an LA or an UA). In general, for each antecedent variable j and each approximation V, $V \in \{L, U\}$, such a scale transformation is repeatedly applied to transform $\tilde{A}'^V_j$ to the intermediate terms $\tilde{A}''^V_j$ with s^V_{jp} and $\mathbb{S}^V_{jq}$. $\tilde{B}''^V$ is then generated from $\tilde{B}'^V$ using the aggregated $s^V_{\tilde{B}_p}$ and $\mathbb{S}^V_{\tilde{B}_q}$, where $s^V_{\tilde{B}_p} = \frac{1}{M}\sum_{j=1}^{M} s'^V_{jp}$ and $\mathbb{S}^V_{\tilde{B}_q} = \frac{1}{M}\sum_{j=1}^{M} \mathbb{S}'^V_{jq}$.

3.3.2. Move transformation

The move ratios $\mathbb{M}_r$ $(r = 0, \ldots, \lceil (k/2) \rceil - 2)$ shift the locations of the supports of $\tilde{A}^{(r-1)}$ to that of $\tilde{A}^*$ (where $\tilde{A}^{(r-1)}$ is the term obtained after the $(r-1)$ th sub-move with the initialisation $\tilde{A}^{-1} = \tilde{A}''$):

$$\mathbb{M}_r = \begin{cases} \dfrac{\tilde{a}^*_r - \tilde{a}^{(r-1)}_r}{\min\left\{\frac{\tilde{a}^{(r-1)}_r + \cdots + \tilde{a}^{(r-1)}_{\lceil (k/2) \rceil - 1}}{\lceil (k/2) \rceil - r} - \tilde{a}^{(r-1)}_r, \tilde{a}^{(r-1)}_{k-r} - \tilde{a}^{(r-1)}_{k-r-1}\right\}} & \text{if } \tilde{a}^*_r \geq \tilde{a}^{(r-1)}_r \\ \dfrac{\tilde{a}^*_r - \tilde{a}^{(r-1)}_r}{\min\left\{\tilde{a}^{(r-1)}_{k-r-1} - \frac{\tilde{a}^{(r-1)}_{k-\lceil (k/2) \rceil} + \cdots + \tilde{a}^{(r-1)}_{k-r-1}}{\lceil (k/2) \rceil - r}, \tilde{a}^{(r-1)}_r - \tilde{a}^{(r-1)}_{r-1}\right\}} & \text{if } \tilde{a}^{(r-1)}_r > \tilde{a}^*_r \end{cases} \tag{24}$$

where $\tilde{a}^{(r-1)}_r$ is the new position of $\tilde{a}''_r$ after the $(r-1)$ th sub-move. Initially, when $r = 0$, $\tilde{a}^{(-1)}_0 = \tilde{a}''_0$, $\tilde{a}^{(r-1)}_{k-r} - \tilde{a}^{(r-1)}_{k-r-1}$ and $\tilde{a}^{(r-1)}_r - \tilde{a}^{(r-1)}_{r-1}$ are not considered within the denominators min{., .}.

In general, for each antecedent variable j and each approximation V, $V \in \{L, U\}$, this move transformation is repeatedly applied to obtain $\tilde{A}^{(r)V}_j = \{\tilde{a}^{(r)V}_{j0}, \ldots, \tilde{a}^{(r)V}_{j(k-1)}\}$ from $\tilde{A}^{(r-1)V}_j$ using $\mathbb{M}^V_{jr}$. $\tilde{B}^{(r)V} = \{\tilde{b}^{(r)V}_0, \ldots, \tilde{b}^{(r)V}_{k-1}\}$ is then obtained from $\tilde{B}^{(r-1)V}$ using the aggregated $\mathbb{M}^V_{\tilde{B}_r}$, where $\mathbb{M}^V_{\tilde{B}_r} = \frac{1}{M}\sum_{j=1}^{M} \mathbb{M}^V_{jr}$, resulting in $\tilde{A}^{(\lceil (k/2) \rceil - 2)V}_j = \tilde{A}^{*V}_j$ and $\tilde{B}^{(\lceil (k/2) \rceil - 2)V} = \tilde{B}^{*V}$.

3.3.3. Height transformation

The height rates h_o $(o = 1, \ldots, k-2)$ are utilised to adjust the heights $\tilde{H}'^L_{\tilde{A}_o}$ of $\tilde{A}'^L$ to the heights $\tilde{H}^{*L}_{\tilde{A}_o}$ of $\tilde{A}^{*L}$:

$$h_o = \frac{\tilde{H}^{*L}_{\tilde{A}_o}}{\tilde{H}'^L_{\tilde{A}_o}} \tag{25}$$

where $0 < \tilde{H}^{*L}_{\tilde{A}_o} \leq \tilde{H}^{*U}_{\tilde{A}_o} = 1$ and $0 < \tilde{H}'^L_{\tilde{A}_o} \leq \tilde{H}'^U_{\tilde{A}_o} = 1$. This constraint applies to the interpolated conclusion as well. That is, if the height of $\tilde{B}^{*L}$ is greater than the height of $\tilde{B}^{*U}$ after the height transformation, then $\tilde{H}^{*L}_{\tilde{B}_o} = \tilde{H}^{*U}_{\tilde{B}_o}$.

In general, for each antecedent variable j and each approximation V, $V \in \{L, U\}$, this height transformation is repeatedly applied to transform the heights of $\tilde{A}'^L_j$ to those of $\tilde{A}^{*L}_j$ with h_{jo}. The height of the interpolated conclusion is then obtained using the aggregated $h_{\tilde{B}_o}$, where $h_{\tilde{B}_o} = \frac{1}{M}\sum_{j=1}^{M} h_{jo}$.

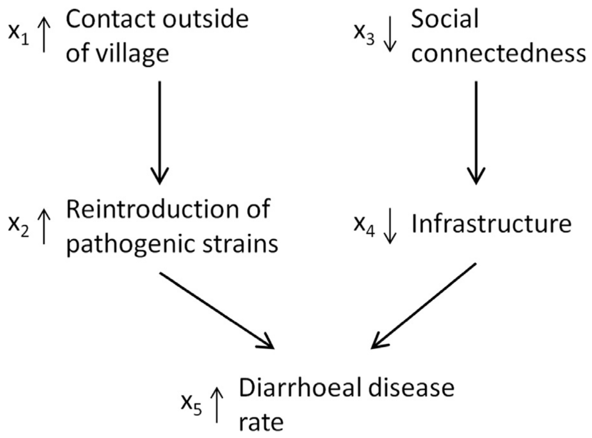

Fig. 4. Causal diagram of a simplified application problem.

Remark 3.2. Scale transformation scales $\tilde{A}'_j$ up or down to $\tilde{A}''_j$, retaining the ratios between left and right slopes, but having different support widths. The closer the scale ratios to 0, the more similar $\tilde{A}'_j$ and $\tilde{A}''_j$. Move transformation shifts $\tilde{A}''_j$ to $\tilde{A}^*_j$ which has the same support width, but having different locations. The closer the move ratios to 0, the more similar $\tilde{A}''_j$ and $\tilde{A}^*_j$. Height transformation adjusts the height of $\tilde{A}'_j$ to that of $\tilde{A}^*_j$ while the other characteristics remain the same. The closer the height rates to 1, the more similar $\tilde{A}'_j$ and $\tilde{A}^*_j$.

Scale, move and height transformations guarantee that the transferred sets have the same type of shapes as that of the original. That is, these three transformations allow the similarity degree between $\tilde{B}'$ and $\tilde{B}^*$ to be determined from those between $\tilde{A}'_j$ and $\tilde{A}^*_j$.

4. Application case study

In this section, a practical problem concerning the prediction of diarrhoeal disease in remote villages is employed in order to demonstrate the potential of the proposed work. It shows how the implemented techniques can help to represent the underlying higher order uncertain information and interpolate a final conclusion.

4.1. Problem overview

Environmental change influences disease burden [28,29]. Intensive studies have been made in an effort to identify logical relationships which lie behind such influences, predicting the consequences of a particular environmental change. This is of significant importance in the assessment of the potential impact of such changes upon the environment and society, e.g., prior to proposing any large-scale infrastructure projects.

One particular application in this area has recently been investigated in [17,18], which is based on the study of [30]. It addresses the issue of measuring how the construction of a new road or railway network in a previously roadless area may affect the epidemiology of infectious diseases in northern coastal Ecuador. A predictive model has been built where many of the factors are not linearly related, but interact with each other in a grid network. Addressing this application problem, an illustrative example is presented here to demonstrate the application of the RF sets and RF rule interpolation. The original problem of [18] is simplified such that all of the factors under consideration are linearly connected. The resulting simplified causal model is shown in Fig. 4.

This causal diagram shows that the diarrhoeal disease rate of a remote village is directly affected by two factors. First, low social connectedness tends to lead to failure in creating adequate water and sanitation infrastructure because the residents are unlikely to know one another well and share social norms [31,32], thereby usually resulting in a high diarrhoeal disease rate. Second, more frequent contact between the residents within a village and those outside tends to increase the rate of introduction of pathogens, thereby also raising the diarrhoeal disease rate.

All factors considered in this example are represented as system variables and each relation between two directly connected factors is represented as a rule associating the relevant variables. In summary, there are five variables in the problem: contact outside of the village, reintroduction of pathogenic strains, social connectedness, hygiene and sanitation infrastructure, and infectious disease rate, denoted as $x_1, \ldots, x_5$, respectively. A selection of the original rules contained in the rule base are given in Table 1, where two rules are sufficient for interpolation. In this model, fuzziness is naturally obtained from the presence of the linguistic terms that describe the real-valued domain variables. Note that different variables are defined upon different domains. To simplify knowledge representation, variable domains are mapped onto the real line and normalised. The illustrative fuzzy sets that represent the normalised linguistic terms expressed by a certain expert are shown

Table 1
Example rules.

	Contact (x_1)	*Reintroduction* (x_2)	*Connectedness* (x_3)	*Infrastructure* (x_4)	*Reintroduction* (x_2)	*Infrastructure* (x_4)	*Rate* (x_5)
R1	L	L	L	LH	VL	MH	L
R2	H	H	MH	H	M	L	MH

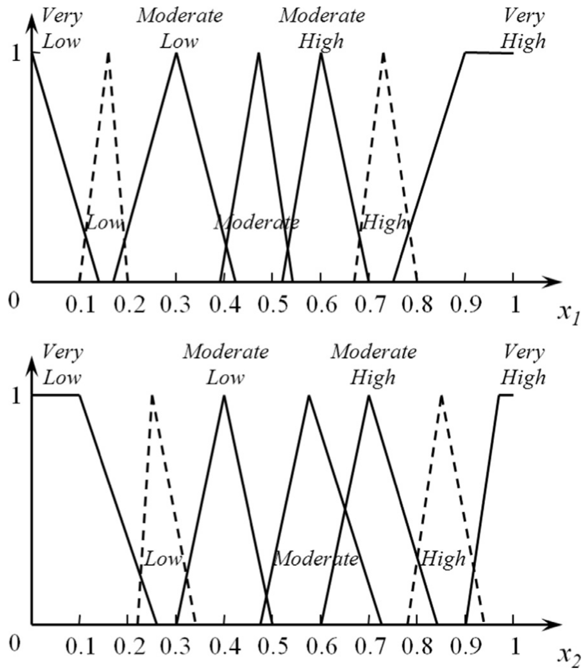

Fig. 5. Definition of the linguistic terms for domain variables from a certain expert.

in Fig. 5, where the triangles with dashed lines indicate the two rules generated from this particular expert's opinion. It is important to note that the original rule base contains substantial gaps, which makes interpolation essential.

In order to evaluate the final disease rate, a group of experts are selected to express their individual view on each factor. However, this procedure relies heavily upon the opinions of experts, who must have a comprehensive and detailed understanding of the underlying problem. Such opinions are often biased and subjective and/or inconsistent between different individuals. Suppose that the opinions from six experts, denoted as $T_1, \ldots, T_6$, in the group are shown in Fig. 6, where subsets of rules (one subset per causal implication): $A \rightarrow B$, $C \rightarrow D$ and $B \wedge D \rightarrow E$ are established by the experts with each supported by two of them. Note that $A = x_1$, $B = x_2$, $C = x_3$, $D = x_4$ and $E = x_5$. This reveals the underlying uncertainty about the opinions from different experts and reinforces the need for RF representation. In particular, opinion from expert T_1 is the one displayed in Fig. 5. For brevity, others are omitted here.

4.2. Experimentation and discussion

4.2.1. Motivation revisited

Given different expert rules and observations, one way to resolve the problem might be to use a conventional FRI approach, say T-FRI to implement the required interpolation separately. Suppose that two pairs of expert rules are contained in a sub-rule base: $A_1 \rightarrow B_1$ and $A_2 \rightarrow B_2$, where $A_{11} \rightarrow B_{11}$ and $A_{21} \rightarrow B_{21}$ are provided by expert T_1, while $A_{12} \rightarrow B_{12}$ and $A_{22} \rightarrow B_{22}$ are provided by expert T_2. Presented with two observations A_1^* and A_2^*, the interpolated result by the use of T-FRI is a set which contains 4 elements. The computation with respect to the remainder of the subsets of rules follows the same procedure, resulting in a consequence set of 32 interpolated results, as listed in Table 2.

Note that the cardinality of the set of interpolated consequent results increases rapidly along with the increase of the cardinality of rule subsets and the number of observations. This results in high computational complexity. As outlined previously, the first step of interpolation requires the computation of the closest rules from a given rule base. A distance measure needs to be employed in order to estimate the proximity between each rule antecedent and the observation. This implies a time complexity of $O(xyz)$, where x is the number of observations to be interpolated, y is the number of antecedent variables, and z is the number of fuzzy rules involved in the relevant rule subset.

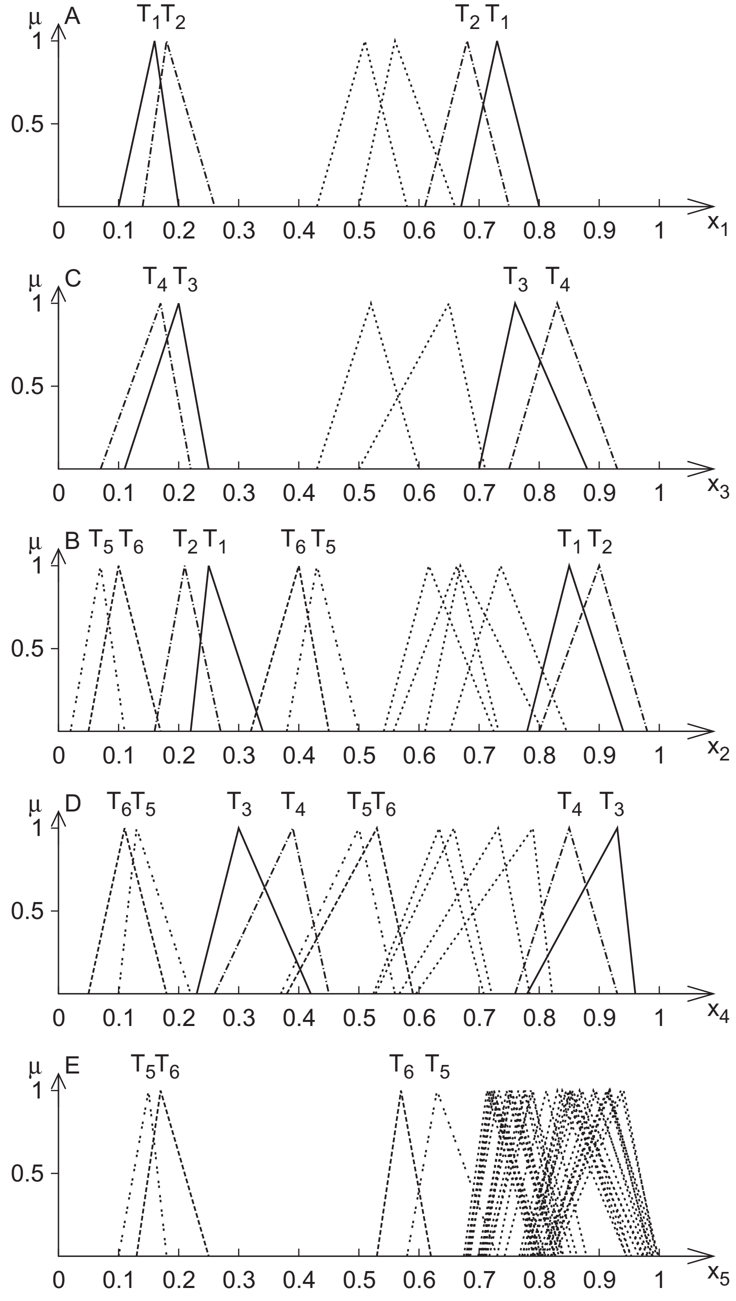

Fig. 6. Interpolated results from conventional FRI.

Table 2
Interpolated results with T-FRI.

	Values		Values		Values		Values
E_1^*	(0.771,0.845,0.953)	E_2^*	(0.805,0.914,0.999)	E_3^*	(0.767,0.830,0.945)	E_4^*	(0.799,0.890,0.989)
E_5^*	(0.775,0.867,0.952)	E_6^*	(0.809,0.936,0.998)	E_7^*	(0.772,0.851,0.945)	E_8^*	(0.802,0.913,0.988)
E_9^*	(0.786,0.852,0.975)	E_{10}^*	(0.821,0.918,0.991)	E_{11}^*	(0.781,0.838,0.966)	E_{12}^*	(0.814,0.894,0.981)
E_{13}^*	(0.792,0.871,0.980)	E_{14}^*	(0.826,0.941,0.995)	E_{15}^*	(0.788,0.856,0.972)	E_{16}^*	(0.819,0.917,0.985)
E_{17}^*	(0.678,0.721,0.798)	E_{18}^*	(0.710,0.775,0.840)	E_{19}^*	(0.675,0.713,0.791)	E_{20}^*	(0.698,0.750,0.826)
E_{21}^*	(0.682,0.733,0.799)	E_{22}^*	(0.712,0.790,0.840)	E_{23}^*	(0.679,0.725,0.792)	E_{24}^*	(0.701,0.765,0.825)
E_{25}^*	(0.685,0.724,0.810)	E_{26}^*	(0.717,0.777,0.854)	E_{27}^*	(0.681,0.717,0.804)	E_{28}^*	(0.705,0.754,0.838)
E_{29}^*	(0.708,0.753,0.833)	E_{30}^*	(0.742,0.812,0.879)	E_{31}^*	(0.704,0.745,0.826)	E_{32}^*	(0.730,0.786,0.863)

Table 3
Relevant RF sets for sub-rule base $A \to B$.

Attribute values	$A_1 = \langle(0.14, 0.17, 0.17, 0.2; 0.75, 0.75), (0.1, 0.16, 0.18, 0.26; 1, 1)\rangle$ $A_2 = \langle(0.67, 0.71, 0.71, 0.75; 0.62, 0.62), (0.61, 0.68, 0.73, 0.8; 1, 1)\rangle$ $B_1 = \langle(0.22, 0.24, 0.24, 0.27; 0.56, 0.56), (0.16, 0.21, 0.25, 0.34; 1, 1)\rangle$ $B_2 = \langle(0.8, 0.87, 0.87, 0.94; 0.74, 0.74), (0.78, 0.85, 0.9, 0.98; 1, 1)\rangle$
Observation	$A^* = \langle(0.5, 0.54, 0.54, 0.58; 0.62, 0.62), (0.43, 0.51, 0.56, 0.66; 1, 1)\rangle$

Table 4
Relevant RF sets for sub-rule base $B \wedge D \to E$.

Rule 1	$B_3 = \langle(0.05, 0.08, 0.08, 0.11; 0.67, 0.67), (0.02, 0.07, 0.10, 0.17; 1, 1)\rangle$ $D_3 = \langle(0.10, 0.12, 0.12, 0.18; 0.8, 0.8), (0.05, 0.11, 0.13, 0.22; 1, 1)\rangle$ $E_1 = \langle(0.13, 0.16, 0.16, 0.18; 0.71, 0.71), (0.1, 0.15, 0.17, 0.25; 1, 1)\rangle$
Rule 2	$B_4 = \langle(0.38, 0.42, 0.42, 0.45; 0.7, 0.7), (0.32, 0.4, 0.43, 0.5; 1, 1)\rangle$ $D_4 = \langle(0.38, 0.51, 0.51, 0.56; 0.86, 0.86), (0.37, 0.50, 0.53, 0.59; 1, 1)\rangle$ $E_2 = \langle(0.58, 0.60, 0.60, 0.62; 0.4, 0.4), (0.53, 0.57, 0.63, 0.73; 1, 1)\rangle$
Observation	$B^* = \langle(0.612, 0.682, 0.682, 0.733; 0.64, 0.64), (0.565, 0.639, 0.697, 0.811; 1, 1)\rangle$ $D^* = \langle(0.582, 0.679, 0.679, 0.708; 0.39, 0.39), (0.531, 0.564, 0.722, 0.808; 1, 1)\rangle$

Suppose that there are m_1 rules in the form of $A \to B$, m_2 rules in $C \to D$, m_3 rules in $B \wedge D \to E$, n_1 observations for the antecedent of the first rule, and n_2 observations for the second. From this, the time complexities for the rule subsets depicting the relations $A \to B$, $C \to D$ and $B \wedge D \to E$ are $O(m_1 n_1)$, $O(m_2 n_2)$ and $O(m_1 n_1 m_2 n_2 m_3)$, respectively. Apart from the computational complexity, this leads to difficulty in determining or interpreting the final result. For example, consider two interpolated results $E_8^* = (0.802, 0.913, 0.988)$ and $E_{25}^* = (0.685, 0.724, 0.810)$. Using the method in [33], the similarity between these two fuzzy sets is 0.002. In this case, it is difficult to make a choice given the almost completely different conclusions. Clearly, it is important to be able to obtain a single consensus opinion which summarises the information contained in disparate or diverging opinions. Fortunately, the proposed RF approach can be applied without suffering from the above difficulty. All such uncertain relations can be captured using RF sets and the conclusion can be derived by RF rule interpolation.

4.2.2. Application of RF-based FRI

An RF rule base and observation can be built on top of those single-expert rule bases and observations using Eq. (4), with examples shown in Fig. 7. That is, all fuzzy values for a single underlying variable are aggregated into an RF set, where the uncertainty is described by the lower and upper approximations. For this example, for simplicity, each variable is associated with two fuzzy values. Such a rule base and observation includes different levels of uncertainty and represents them as RF sets. Also, they are considered in the process of interpolation in order to obtain consensus inference conclusions.

The resultant RF sets for $A \to B$ are listed in Table 3. The interpolation process is described as follows:

1. The lower and upper Reps, shape diversity factors and weight factors are calculated according to Eqs. (9), (10) and (11), respectively.
2. The overall Reps, $\mathrm{Rep}(A_1) = 0.486$, $\mathrm{Rep}(A^*) = 0.842$, $\mathrm{Rep}(A_2) = 1.002$, are calculated from Eq. (12). $A' = \langle(0.505, 0.542, 0.542, 0.579; 0.66, 0.66), (0.452, 0.519, 0.559, 0.632; 1, 1)\rangle$ and $B' = \langle(0.620, 0.674, 0.674, 0.732; 0.68, 0.68), (0.588, 0.651, 0.698, 0.781; 1, 1)\rangle$ are then computed.
3. The scale rates $s^L_{B_0} = 1.084$, $s^U_{B_0} = 1.273$, $s^U_{B_1} = 1.229$, the move ratios $\mathbb{M}^L_{B_0} = -0.179$, $\mathbb{M}^U_{B_0} = 0.093$ and the height rate $h_{B_1} = 0.939$ in the integrated transformation from A' and A^* are calculated with regard to Eqs. (21), (24) and (25), respectively.
4. The scale rates, move ratios and height rate are used to transform B' to the interpolated conclusion $B^* = \langle(0.612, 0.682, 0.682, 0.733; 0.64, 0.64), (0.565, 0.639, 0.697, 0.811; 1, 1)\rangle$, as summarised in Fig. 7.

Since the interpolation for calculating D^* from $C \to D$ is similar to the above process, it is omitted here to avoid repetition. Having generated B^* and D^*, $B \wedge D \to E$ is then utilised to derive the final result. The RF sets involved are listed in Table 4.

Note that both given rules in the form of $B \wedge D \to E$ lie on one side of the observation, the problem of interpolation therefore becomes that of extrapolation. However, the extension of RF rule interpolation to extrapolation is straightforward.

1. For the first antecedent, the distances between B_3, B_4 and the previously interpolated B^* are calculated by Eq. (13). The weights are calculated and normalised using Eqs. (15) and (16), respectively, resulting in the new weights of 0.30 and 0.70. The normalised weights together with the parameter $\delta_B = 0.60$, which is computed by Eq. (18), are then used to generate the required intermediate fuzzy set $B' = \langle(0.634, 0.671, 0.671, 0.701; 0.69, 0.69), (0.584, 0.654, 0.684, 0.754; 1, 1)\rangle$, according to Eq. (17). D' can then be calculated in the same way.

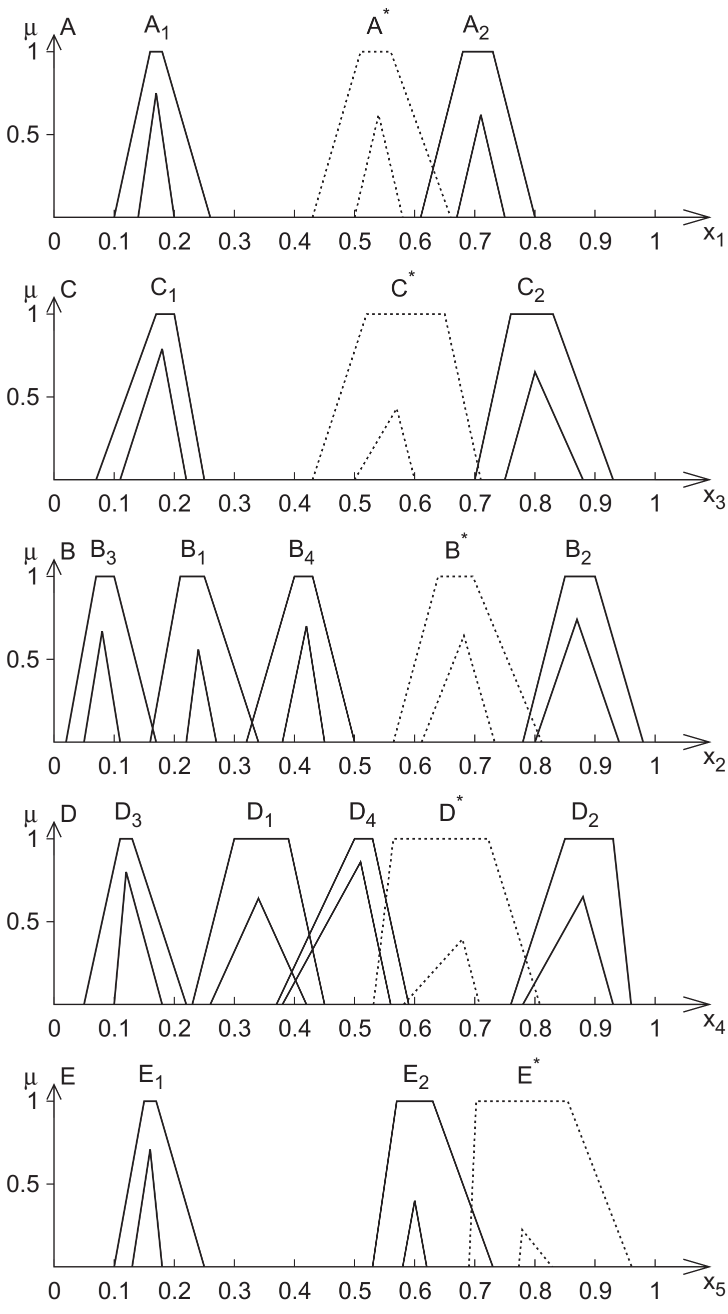

Fig. 7. Interpolated results from RF rule interpolation.

2. To compute the consequent, the averaged weights of 0.26 and 0.74, and $\delta_E = 0.49$ are calculated using Eq. (20). From this, the intermediate output $E' = \langle(0.772, 0.794, 0.794, 0.814; 0.48, 0.48), (0.727, 0.770, 0.819, 0.914;\ 1,\ 1)\rangle$ is obtained using Eq. (19).
3. The average of two support scale rates (1.81 and 0.79) of the LAs is computed according to Eqs. (21), (22) and (23), resulting in $s^L_{E_0} = 1.30$ and forming the scale rate of the aggregated LA. The scale rates of the aggregated UA $s^U_{E_0} = 1.38$ and $s^U_{E_1} = 3.31$ can then be generated following the same procedure. Similarly, the aggregated move ratios $\mathbb{M}^L_{E_0} = 0.81$ and $\mathbb{M}^U_{E_0} = 1.41$ are calculated from two move ratios (0.24 and 1.37) of the LAs and two move ratios (0.60 and 2.22) of the UAs using Eq. (24). These, together with the aggregated height rate, namely, the average $h_{E_1} = 0.46$ from Eq. (25), are employed to transform E', achieving the final result $E^* = \langle(0.773, 0.779, 0.779, 0.829; 0.22, 0.22), (0.690, 0.702, 0.855, 0.961; 1, 1)\rangle$. The interpolated result is again illustrated in Fig. 7.

Table 5
Involved RF sets for Singleton-valued Interpolation.

Attribute values	$\tilde{A}_1 = \langle(3, 3, 3; 1), (3, 3, 3; 1)\rangle$
	$\tilde{A}_2 = \langle(12, 13, 13.5; 0.6), (11, 13, 14; 1)\rangle$
	$\tilde{B}_1 = \langle(4, 4, 4; 1), (4, 4, 4; 1)\rangle$
	$\tilde{B}_2 = \langle(10.5, 11.5, 12; 0.5), (10, 11.5, 13; 1)\rangle$
Observation	$\tilde{A}^* = \langle(6, 7, 8; 0.6), (5, 7, 9; 1)\rangle$

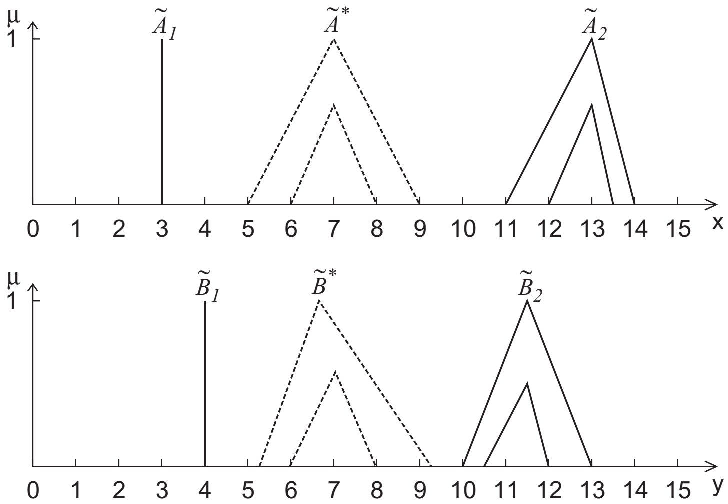

Fig. 8. A single antecedent case with singleton-valued conditions.

Note that this figure reflects the distribution of those results shown in Fig. 6. In particular, the shape of the resultant RF set is similar to the shape distribution of those 32 interpolated sets, whereas the computational complexity of the former is much lower than that of the latter. This can be seen by comparing the calculated time complexities of the former, which are $O(m_1)$, $O(m_2)$ and $O(m_3)$, respectively. It is obvious that the reduction in computation complexity is significant, especially when the number of observations becomes large for a given application. In addition, since a majority of the 32 results are closer to the second rule (see Table 4 and Fig. 6), the resultant RF set is also closer to the consequent of this rule. This demonstrates that the present approach corresponds to the intuition for its development.

5. Further evaluation

5.1. Specific cases

The RF rule interpolation and extrapolation processes have been illustrated in the previous section. However, the particular application does not offer a scenario to test specific cases where rules involve singleton 'fuzzy' values or where RF sets degenerate to conventional fuzzy sets. These cases are empirically checked below.

5.1.1. Singleton-valued interpolation

This case considers one single antecedent variable involving singleton-valued conditions. The relevant RF sets are listed in Table 5.

The interpolated conclusion $\tilde{B}^* = \langle(5.98, 7.04, 7.98; 0.57), (5.27, 6.66, 9.27;\ 1)\rangle$ is calculated following the procedures described in Sections 2 and 3, as shown in Fig. 8. It follows that if certain components involved in the given rules are singleton-valued, the interpolated conclusion remains an RF set given an RF observation. This has a clear intuitive appeal.

5.1.2. Degenerated interpolation

The concept of RF sets extends that of conventional fuzzy sets, while the RF rule interpolation extends from the existing T-FRI. When there is no higher order uncertainty involved, i.e., the LA coincides with the LU, an RF set degenerates to a conventional fuzzy set. If all of the sets under consideration in the implementation of interpolation/extrapolation are conventional fuzzy sets, then the results obtained by the proposed approach are identical to those of T-FRI. The example in Fig. 9 illustrates this.

Consider a case where all of the RF sets concerned degenerate to conventional fuzzy sets, i.e., $\tilde{A}^{*L} = \tilde{A}^{*U}$, $\tilde{A}_k^L = \tilde{A}_k^U$ and $\tilde{B}_k^L = \tilde{B}_k^U$, $k = 1, 2$. Let $\tilde{A}^*$, $\tilde{A}_1$, $\tilde{A}_2$, $\tilde{B}_1$ and $\tilde{B}_2$ be RF sets, where all the terms are listed in Table 6.

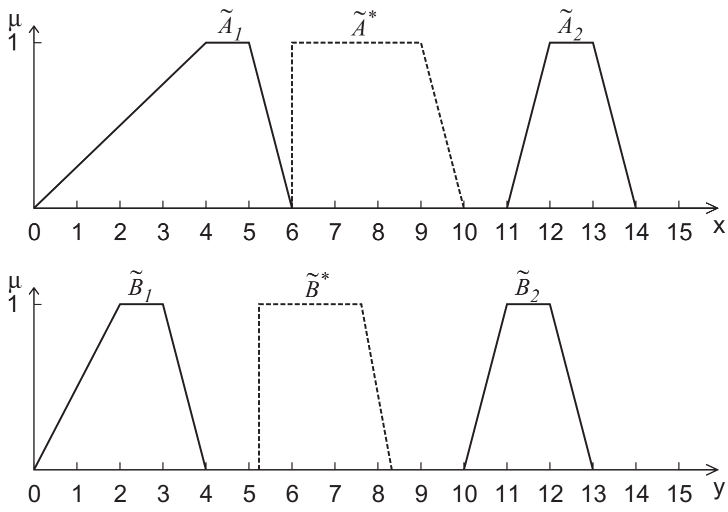

Fig. 9. Interpolation when RF sets degenerate to conventional fuzzy sets.

Table 6
Involved RF sets for Degenerated Interpolation.

Attribute values	$\tilde{A}_1 = \langle(0, 4, 5, 6; 1, 1), (0, 4, 5, 6; 1, 1)\rangle$ $\tilde{A}_2 = \langle(11, 12, 13, 14; 1, 1), (11, 12, 13, 14; 1, 1)\rangle$ $\tilde{B}_1 = \langle(0, 2, 3, 4; 1, 1), (0, 2, 3, 4; 1, 1)\rangle$ $\tilde{B}_2 = \langle(10, 11, 12, 13; 1, 1), (10, 11, 12, 13; 1, 1)\rangle$
Observation	$\tilde{A}^* = \langle(6, 6, 9, 10; 1, 1), (6, 6, 9, 10; 1, 1)\rangle$

Table 7
RMSE% for the benchmark datasets.

	Yacht				*Servo*			
Partitions	Expert 1	Expert 2	Expert 3	RF	Expert 1	Expert 2	Expert 3	RF
2	30.90(∗)	30.46(∗)	30.02(∗)	25.80	26.20(∗)	25.60(∗)	25.00(∗)	23.83
3	19.50(∗)	19.93(∗)	19.64(∗)	19.10	15.13(-)	15.45(∗)	15.74(∗)	15.29
4	18.72(∗)	19.82(∗)	21.20(∗)	17.31	12.59(∗)	12.78(∗)	13.22(∗)	12.19
5	16.53(-)	18.21(∗)	18.52(∗)	16.59	11.60(∗)	11.78(∗)	12.29(∗)	11.36
6	14.72(v)	16.10(-)	17.87(∗)	16.10	11.50(∗)	11.50(∗)	12.07(∗)	11.24
7	15.23(v)	16.35(∗)	16.01(∗)	15.52	11.29(∗)	11.30(∗)	11.59(∗)	11.01
	Concrete				*Housing*			
2	18.10(∗)	18.06(-)	18.03(-)	18.05	17.26(-)	17.22(-)	17.13(v)	17.21
3	16.82(-)	16.94(∗)	17.01(∗)	16.85	16.65(∗)	16.10(∗)	15.54(-)	15.75
4	15.09(∗)	14.84(-)	14.62(v)	14.85	15.15(∗)	14.70(-)	14.44(v)	14.73
5	13.54(v)	14.65(-)	14.62(-)	14.63	14.09(∗)	13.89(-)	13.61(v)	13.86
6	13.47(∗)	13.34(-)	14.66(∗)	13.36	13.51(-)	13.57(-)	13.66(-)	13.56
7	13.14(v)	13.49(∗)	14.42(∗)	13.28	12.98(v)	13.19(-)	13.42(∗)	13.16

Using the proposed approach, the interpolated conclusion $\tilde{B}^* = \langle(5.23, 5.23, 7.61, 8.32;\ 1, 1), (5.23, 5.23, 7.61, 8.32;\ 1, 1)\rangle$ is obtained, as shown in Fig. 9. The details of the calculation are omitted here to avoid repetition. It follows that if all given sets are conventional fuzzy sets, the interpolated result is indeed the same as that achieved using the classical T-FRI.

5.2. Additional application case studies

Further to the examples previously shown, the proposed approach has also been evaluated against further applications for decision support, by adapting UCI-MLR benchmark datasets [34], including *Yacht Hydrodynamics*, *Servo*, *Concrete Compressive Strength*, and *Housing*.

Table 7 shows the results of the averaged root-mean-square error (RMSE) values computed over 10 times 10-fold cross-validation [35,36], in relation to K ($K = 2, \dots, 7$) partitions (i.e., the underlying domain of each variable is divided into 6 fuzzy values) and N ($N = 6$) closest rules that are used to interpolate the conclusions. The results are compared to those achieved by individual T-FRI interpolations that each implement the interpolation with the opinion provided by one particular expert. In order to assess the statistical significance of the obtained results, a Paired *T*-test is used. A significance level of 0.05 is employed in these tables for the achieved accuracies. The RF approach is utilised as a reference, and those results which are better, worse and of no statistical significance are marked with '(v)', '(∗)' and '(-)', respectively.

As reflected in the results, the accuracies from three T-FRI interpolations are unstable. That is, the opinions from an expert may perform well in certain partitions, but badly in others. Theoretically, this is acceptable as one is only an expert in a particular field, namely, the necessary expertise may only be available for a certain concept. However, this leads to difficulty for decision making in practical applications when a consensus of multiple experts or a group-based opinion is required. Fortunately, the performance of the proposed RF approach is generally better than that of T-FRI in isolation, reflecting an important advantage of the proposed approach.

6. Conclusion

This paper has proposed a novel approach for both representing the knowledge involving higher order uncertainty and facilitating rule interpolation with such knowledge. It has introduced the concept of rough-fuzzy (RF) sets, via the use of lower and upper approximation membership functions and presented an algorithm for RF rule interpolation. A number of examples have been provided in the paper in order to illustrate the algorithm's potential, including a realistic problem that predicts diarrhoeal disease rates in roadless villages. These examples demonstrate that the proposed approach is of a natural appeal for interpolation and extrapolation when dealing with different levels of uncertainty that conventional FRI techniques may otherwise ignore. In particular, through RF set-based interpolation, the exploitation of uncertain knowledge across multiple opinions offered by different experts is facilitated, leading to improved results over the use of only expertise offered by an individual expert.

The work offers many areas for further investigation. For example, the opinion of an individual in a group may be distinct from the others, which may result in an empty lower approximation and hence, difficulty for the task of group decision making. However, all of the individual experts should contribute to the outcome, although one outlier should not dominate the overall result. Ordered weighted averaging (OWA) operators [37–39] may be applied to enhance the ability of this approach. Also, certain rules may be very useful or even crucial, but others may be less useful (and may even contain misleading information). A weighting scheme should therefore be assigned to the rules in order to express the belief of usefulness related to each rule. It would be interesting to investigate the performance of the approach by learning the weights from the constructed rule bases [40]. In addition, scaled-up real-world applications would help to further evaluate the full efficacy and potential of this work.

Acknowledgments

The first author is very grateful to Aberystwyth University for providing a PhD scholarship in support of this research. The third and last authors would like to thank the Royal Academy of Engineering, UK, for their support, under Grant 1314RECI025. Thanks also go to the Associate Editor and the reviewers for their constructive comments which have helped improve this work significantly.

References

[1] L.A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, IEEE Trans. Syst. Man Cybern. 3 (1) (1973) 28–44.
[2] K. Wong, D. Tikk, T. Gedeon, L. Kóczy, Fuzzy rule interpolation for multidimensional input spaces with applications: a case study, IEEE Trans. Fuzzy Syst. 13 (6) (2005) 809–819.
[3] L. Kóczy, K. Hirota, Approximate reasoning by linear rule interpolation and general approximation, Int. J. Approx. Reason. 9 (3) (1993) 197–225.
[4] L. Kóczy, K. Hirota, Interpolative reasoning with insufficient evidence in sparse fuzzy rule bases, Inf. Sci. 71 (1) (1993) 169–201.
[5] Y. Shi, M. Mizumoto, Reasoning conditions on Kóczy's interpolative reasoning method in sparse fuzzy rule bases. Part II, Fuzzy Sets Syst. 87 (1) (1997) 47–56.
[6] S. Yan, M. Mizumoto, W. Qiao, Reasoning conditions on Kóczy's interpolative reasoning method in sparse fuzzy rule bases, Fuzzy Sets Syst. 75 (1) (1995) 63–71.
[7] P. Baranyi, L. Kóczy, T. Gedeon, A generalized concept for fuzzy rule interpolation, IEEE Trans. Fuzzy Syst. 12 (6) (2004) 820–837.
[8] Y. Chang, S. Chen, C. Liau, Fuzzy interpolative reasoning for sparse fuzzy-rule-based systems based on the areas of fuzzy sets, IEEE Trans. Fuzzy Syst. 16 (5) (2008) 1285–1301.
[9] S. Chen, Y. Ko, Fuzzy interpolative reasoning for sparse fuzzy rule-based systems based on α-cuts and transformations techniques, IEEE Trans. Fuzzy Syst. 16 (6) (2008) 1626–1648.
[10] S.-M. Chen, Y.-C. Chang, Weighted fuzzy rule interpolation based on GA-based weight-learning techniques, IEEE Trans. Fuzzy Syst. 19 (4) (2011) 729–744.
[11] S.-M. Chen, Y.-K. Ko, Y.-C. Chang, J.-S. Pan, Weighted fuzzy interpolative reasoning based on weighted increment transformation and weighted ratio transformation techniques, IEEE Trans. Fuzzy Syst. 17 (6) (2009) 1412–1427.
[12] W. Hsiao, S. Chen, C. Lee, A new interpolative reasoning method in sparse rule-based systems, Fuzzy Sets Syst. 93 (1) (1998) 17–22.
[13] D. Tikk, P. Baranyi, Comprehensive analysis of a new fuzzy rule interpolation method, IEEE Trans. Fuzzy Syst. 8 (3) (2000) 281–296.
[14] Z. Huang, Q. Shen, Fuzzy interpolative reasoning via scale and move transformations, IEEE Trans. Fuzzy Syst. 14 (2) (2006) 340–359.
[15] Z. Huang, Q. Shen, Fuzzy interpolation and extrapolation: a practical approach, IEEE Trans. Fuzzy Syst. 16 (1) (2008) 13–28.
[16] Q. Shen, L. Yang, Generalisation of scale and move transformation-based fuzzy interpolation, Adv. Comput. Intell. Intell. Inform. 15 (3) (2011) 288–298.
[17] L. Yang, Q. Shen, Closed form fuzzy interpolation, Fuzzy Sets Syst. 225 (2013) 1–22.
[18] L. Yang, Q. Shen, Adaptive fuzzy interpolation, IEEE Trans. Fuzzy Syst. 19 (6) (2011) 1107–1126.
[19] S. Jin, R. Diao, C. Quek, Q. Shen, Backward fuzzy rule interpolation, IEEE Trans. Fuzzy Syst. 22 (6) (2014) 1682–1698.
[20] W. Pedrycz, Fuzzy set technology in knowledge discovery, Fuzzy Sets Syst. 98 (3) (1998) 279–290.
[21] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (3) (1965) 338–353.

[22] X. Fu, Q. Shen, Fuzzy compositional modeling, IEEE Trans. Fuzzy Syst. 18 (4) (2010) 823–840.
[23] Z. Pawlak, Rough sets, Int. J. Comput. Inf. Sci. 11 (5) (1982) 341–356.
[24] Z. Pawlak, Rough sets and intelligent data analysis, Inf. Sci. 147 (1) (2002) 1–12.
[25] Y. Qian, J. Liang, W. Pedrycz, C. Dang, Positive approximation: an accelerator for attribute reduction in rough set theory, Artif. Intell. 174 (9–10) (2010) 597–618.
[26] M. Banerjee, S. Pal, Roughness of a fuzzy set, Inf. Sci. 93 (3) (1996) 235–246.
[27] B. Bouchon-Meunier, L. Valverde, A fuzzy approach to analogical reasoning, Soft Comput. 3 (3) (1999) 141–147.
[28] R.R. Colwell, P.R. Epstein, D. Gubler, N. Maynard, A.J. McMichael, J.A. Patz, R.B. Sack, R. Shope, Climate change and human health, Science 279 (5353) (1998) 968–969.
[29] S.S. Morse, Factors in the emergence of infectious diseases, Emerg. Infect. Dis. 1 (1) (1995) 7–15.
[30] J.N. Eisenberg, W. Cevallos, K. Ponce, K. Levy, S.J. Bates, J.C. Scott, A. Hubbard, N. Vieira, P. Endara, M. Espinel, et al., Environmental change and infectious disease: how new roads affect the transmission of diarrheal pathogens in rural Ecuador, Proc. Natl. Acad. Sci. 103 (51) (2006) 19460–19465.
[31] A. Bebbington, T. Perreault, Social capital, development, and access to resources in highland Ecuador, Econ. Geogr. 75 (4) (1999) 395–418.
[32] C. Grootaert, T. Van Bastelaer, Understanding and Measuring Social Capital: A Multidisciplinary Tool for Practitioners, World Bank, 2002.
[33] C.P. Pappis, N.I. Karacapilidis, A comparative assessment of measures of similarity of fuzzy values, Fuzzy Sets Syst. 56 (2) (1993) 171–174.
[34] K. Bache, M. Lichman, UCI machine learning repository, 2013, URL: http://archive.ics.uci.edu/ml (accessed March 2016).
[35] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection, in: Proceedings of the International Joint Conference on Artficial Intelligence, 1995, pp. 1137–1143.
[36] J. Shao, Linear model selection by cross-validation, J. Am. Stat. Assoc. 88 (422) (1993) 486–494.
[37] T. Boongoen, Q. Shen, Nearest-neighbor guided evaluation of data reliability and its applications, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 40 (6) (2010) 1622–1633.
[38] Z. Xu, Dependent OWA operators, Proceedings of International Conference on Modeling Decisions for Artificial Intelligence, 2006, pp. 172–178.
[39] R.R. Yager, On ordered weighted averaging aggregation operators in multicriteria decisionmaking, IEEE Trans. Syst. Man Cybern. 18 (1) (1988) 183–190.
[40] Y. Yoshinari, W. Pedrycz, K. Hirota, Construction of fuzzy models through clustering techniques, Fuzzy Sets Syst. 54 (2) (1993) 157–165.