IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received January 26, 2022, accepted February 22, 2022, date of publication March 8, 2022, date of current version March 17, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3157616

A Novel Fuzzy Monotone Relationship Method With Its Application on Inclusion Degree

JIN LIANG AND **ZUO PENG ZHANG**
School of Software, South China Normal University, Foshan, Guangdong 528225, China
Corresponding author: Jin Liang (466689586@qq.com)

ABSTRACT This paper introduces a new fuzzy monotone relationship and its associated method, which are applied to feature selection and correlation analysis. Specifically, after the concept of a fuzzy monotone is introduced, this paper first defines a new fuzzy monotone relationship between inputs and output. Second, a fuzzy inclusive monotone model is constructed on inclusion degree through several proved propositions, together with presenting a fuzzy inclusive monotone decision membership function. Third, a new algorithm is developed according to the proposed model for feature selection or correlation analysis. Compared with several methods, the proposed algorithm has been validated on several data sets. The results indicate that the proposed algorithm is effective for the selection of numeric attributes, and the correlation analysis. The novel fuzzy monotone relationship and the method are validated through theoretic proof and experimental results.

INDEX TERMS Fuzzy relations, fuzzy monotone, inclusion degree, feature selection and correlation analysis, information system.

I. INTRODUCTION

For a multi-input and output system, it is important to determine which inputs are important to the output. If the transfer function for a system is given, it is easy to specify the relationship between its inputs and output. In real applications, such inputs and output are, however, complex with nonlinear relationships. When an output attribute is discrete, many classification methods are available for dealing with the problem. When an output attribute is numeric or continuous, some models such as regression are used instead. However, this is not always effective, particularly for the complex nonlinear relationship of inputs and output. On the other hand, it is not always effective to identify the reduced number of important input attributes with respect to numeric output attributes. After Zadeh [1] introduced fuzzy sets, some fuzzy sets methods have been introduced to deal with some nonlinear relationships. Although some papers [2], [3] introduced that fuzzy rough set methods can be directly applied to continuous data, they basically refer to the input continuous data, when the output data is continuous, they always have the continuous output data discretized and classified first, and transform the problem to the classification problem, however, this process would result in information loss and errors [3]–[5]. In information systems, the main goal of the attribute reduction classification methods is to remove redundant information, so that a correct decision can quickly be made while preserving or even improving the classification ability [6]. However, it is unclear whether the continuous quantity variations of numeric decision attributes are dependent on the reduced condition attributes or not in these methods. The monotonicity is always used to describe the quantity variation relationship between the continuous input and output ends, and it is also an important property between the input and the output. The fuzzy and monotone technologies are sometimes combined together in some applications. Edward [7] introduced the fuzzy monotone function, together with logic control applications. Many papers [8]–[21] have discussed the monotonicity property in the fuzzy inference system (FIS) including Mamdiani, TSK and etc. Some papers [8]–[19] discussed various useful mathematical conditions to satisfy the monotonicity property for different fuzzy inference systems' models; Some paper [21] discussed the data driven monotone fuzzy system. And the monotonicity is also discussed as an important property in aggregation functions too [20]. These works on fuzzy with monotone focus mainly on constructing an input and output model or introducing the fuzzy monotone function by using the fuzzy logic and language, not focus on

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wang.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

the data analysis for feature selection or correlative analysis. Here we mainly focus on the data analysis combining with monotone and fuzzy technologies between the continuous input and output ends. Since traditional monotone and other some relationships are not directly suitable to deal with nonlinear data well, and few relationships can directly be applied to select a reduced number of important input attributes or find correlative input attributes very well from the continuous output data because of complexity nonlinear relationship. This paper will introduce a new fuzzy monotone relationship, which differs from the traditional monotone one and is suitable to deal with nonlinear data. Based on this new concept, we develop a method that is able to select a reduced number of important input attributes from the continuous output data, and research the correlative analysis for partial dependence relationship between the continuous input and output ends. And the new method based on the novel fuzzy monotone relationship can describe the continuous quantity variation relationship between the continuous input and output ends.

The rest of this paper is organized as follows. In Section II, we review some concepts of decision tables, inclusion degree, and correlative fuzzy concept. In Section III, we present a novel fuzzy monotone relationship model, discuss the monotone mapping relationship, and define the fuzzy inclusive monotone dependence relationship. We then construct the fuzzy inclusive monotone model between inputs and an output after some propositions and proofs. In Section IV, a new algorithm is presented according to the fuzzy inclusive monotone model. In Section V, the new algorithm is applied to several continuous data sets to validate its effect, and compared with the other methods. The advantages of the method are then discussed. Section V concludes this paper.

II. PRELIMINARY

In this section, we review several relevant basic concepts and preliminaries.

A. DECISION TABLE

An information system is defined as $S = (U, A, V, F)$, where $U = \{e_1, e_2, \ldots, e_n\}$ is a finite set of objects called the universe (n is the number of objects), A a finite set of attributes, $V = \bigcup_{a \in A} V_a$ is the range of A, and $f : U \times A \to V$ is an information function that assigns an information value to each object, i.e., $\forall a \in A, e' \in U, f(e', a) \in V_a$. If an attributes in A can be divided into two sets of condition attributes C and decision attributes D, i.e., if $A = C \cup D$ and $C \cap D = \emptyset$, then the information system S is called a decision system or decision table [22]–[33]. Here we further suppose $C = \{C_1, C_2, \ldots, C_m\}$ and $D = \{D_1, D_2, \ldots, D_k\}$, with each element having their attribute value sets $C_{iv} = \{x_{i1}, x_{i2}, \ldots, x_{in}\}$ and $D_{jv} = \{y_{j1}, y_{j2}, \ldots, y_{jn}\}$, $1 \le i \le m$, $1 \le j \le k$, where C_{iv} is a value set of condition attribute C_i and D_{jv} for decision attribute D_j. For $\forall e_i \in U$, we have $f(e_i, C_i) = x_{ii} \in V_{C_i}$ and $f(e_i, D_j) = y_{ji} \in V_{D_j}$.

B. INCLUSION DEGREE CONCEPT

The concept of an inclusion degree comes from the including measure among sets. If $I(B/A)$ denotes the degree for set A included in set B, then the following properties hold [34], [35]:

1) $0 \le I(B/A) \le 1$;
2) If $A \subseteq B$ then $I(B/A) = 1$; and
3) If $B_1 \subseteq B_2$ then $I(B_1/A) \le I(B_2/A)$ holds.

For $\forall A, B \in U$, where U is a set, we define $I(B/A) = \frac{|A \cap B|}{|B|}$ as an inclusion degree. In a case of $B = \emptyset$, we set $I(B/A) = 1$.

C. CORRELATIVE FUZZY CONCEPT

We suppose that F is a fuzzy set on U. In terms of a relevant crisp universal set, each fuzzy set is defined by a function that is similar to the characteristic function of crisp sets. This function is called a membership function. To define a fuzzy set F on a given universal set U, we assign each member x of U a real number in the unit interval [0,1] by using a membership function. This number is viewed as the degree of membership x in A. Denoting a membership function of a fuzzy set F as μ_F, we have $\mu_F : U \to [0, 1]$. For each $x \in U$, the value $\mu_F(x)$ is the degree of membership x in F, with the form of $\mu_F : x \to \mu_F(x)$. In an alternative notation, the membership function denoted by F has the same form $F : U \to [0, 1]$. Clearly, $F(x)$ is the degree of membership of x in F [36][37].

In this paper, we mainly review the intersection characteristic of fuzzy sets. Given two fuzzy sets F_1 and F_2 defined on the same universal set U, their standard intersection $F_1 \bigwedge F_2$ is defined by the equation $(F_1 \bigwedge F_2)(x) = \min(F_1(x), F_2(x))$ for each $x \in U$, where min is the minimum operator. The intersection operation can be extended to a finite number of fuzzy sets, so a concept of the degree of membership in a fuzzy set can be introduced. A membership function on U is denoted as μ. For each $x \in U$, $\mu(x)$ is the degree of membership x in U, $\mu(U)$ the intersection of $\mu(x)$, and the form is $\mu(U) = \bigwedge_{x \in U} \mu(x)$. Similarly, $\mu(F) = \bigwedge_{x \in F} \mu(x)$ is the degree of membership in the fuzzy set F [36].

III. FUZZY MONOTONE MODEL BASED ON INCLUSION DEGREE

A. FUZZY MONOTONE RELATIONSHIP

The existing monotone relationship is described as follows: With A_1 and B_1 as two sets, $\le_{A_1}$ and $\le_{B_1}$ are two separate linear ordering relationships, and $f : A_1 \to B_1$, for any $x_1, x_2 \in A_1$. If $x_1 \le_{A_1} x_2$ and $f(x_1) \le_{B_1} f(x_2)$, we then call f the monotonically increasing. In contrast, if $x_1 \le_{A_1} x_2$ and $f(x_2) \le_{B_1} f(x_1)$, then f is called the monotonically decreasing. Note that A_1 refers to a value set of a certain condition attribute or input attribute, B_1 a value set of a certain decision attribute or output attribute, and f the monotone relationship in a decision table or information table. However, the existing traditional monotone relationship is just used to describe the linear correlation phenomenon, and too strictly to describe nonlinear correlation phenomenon, which is very popular in our life. And the sentence patterns "the more ... the more... " and "the more... the less... " are always used to describe the correlation phenomenon. In order to describe the nonlinear correlation and deal with the correlative data well, an idea of the fuzzy monotone was first presented in [37], which is different from the existing monotone relationship. In this paper, we present a new fuzzy monotone relationship according to the idea.

There exists a monotone quantity-dependent relationship between an input and an output according to the law of conservation of energy. An increase in the value of an input, for example, will lead to the increase in values of some outputs. Note that the relationship between the input and output is different from the existing monotone relationship. This new monotone relationship cannot be described by the existing monotone relationship directly. As we know, the disturbance influence of an attribute's value is in a certain range. Based on this fact, our observation is as follows: more than half output values in an interval range A' appear to be larger than or equal to more than half output values in the other interval range B', when more than half input values in A' appear to be larger than or equal to more than half input values in B'. We call this phenomenon as a fuzzy monotonically increasing between an input and an output. Similarly, if more than half output values in A' appear to be larger than or equal to more than half output values in B' when more than half input values in A' appear to be smaller than or equal to more than half input values in B', we call the relationship as a fuzzy monotonically decreasing between the input and output.Why we select "half values" to describe the phenomenon, because less than half values as limit is uncertain to describe the phenomenon, such as the sentence " more than 40 percentage values in A' are larger than or equal to more than 40 percentage values in B'", there exist some contradiction situations in this description, for example, possibly there exists "more than 45 percentage values in A' are larger than or equal to more than 45 percentage values in B' ", however, this possibly means that "55 percentage values in A' are less than or equal to 55 percentage values in B'", then it is uncertain to describe whether is the "the more... the more... " or "the more ... the less... ". When the comparison limit is over half, the description range becomes narrow compared with the limit as half, and the situation can certainly be described by the limit as half too. Then it is reasonable to select the half as the comparison limit. Obviously, the new concept of a fuzzy monotone relationship is different from that of the existing monotone relationship. In the following, we provide an example to illustrate this new relationship.

Example 1: Suppose a decision table given in TABLE 1:

TABLE 1. A numeric decision table example.

U	C1	C2	C3	D
e1	1.2	2.4	4.0	1.7
e2	1.5	0.6	3.7	1.9
e3	1.45	1.5	4.1	1.93
e4	1.47	1.1	3.8	1.92
e5	1.7	0.5	4.1	2.1
e6	1.65	1.8	3.9	2.0
e7	1.71	0.9	4.3	2.2
e8	1.75	1.7	4.4	2.15
e9	1.73	0.2	4.2	2.3

From TABLE 1, we can see that there is no monotone relationship between decision attribute D and condition attributes $C1$ or $C2$. But when we partition the nine samples into three intervals, namely $U = \{\{e1, e2, e3\}, \{e4, e5, e6\}, \{e7, e8, e9\}\}$, $C1 = \{\{1.2, 1.5, 1.45\}, \{1.47, 1.7, 1.65\}, \{1.71, 1.75, 1.73\}\}$, $C2 = \{\{2.4, 0.6, 1.5\}, \{1.1, 0.5, 1.8\}, \{0.9, 1.7, 0.2\}\}$, $D = \{\{1.7, 1.9, 1.93\}, \{1.92, 2.1, 2.0\}, \{2.2, 2.15, 2.3\}\}$, and denote the first interval as $IV1$, the second interval as $IV2$ and the third interval as $IV3$, then most values in the interval $IV2$ of $C1$ are bigger than those values in the interval $IV1$ of $C1$. Similarly, most values of the interval $IV3$ are bigger than those of the interval $IV2$ in $C1$. Accordingly, most values in the interval $IV2$ of D are bigger than those values in the interval $IV1$ of D, and so are for the intervals $IV3$ and $IV2$ in D. Thus, we call that condition attribute $C1$ has a fuzzy monotonically increasing relationship with decision attribute D. But apparently the condition attribute $C2$ has no such a relationship with the attribute D.

In order to identify a fuzzy monotone relationship between any two sets E and F, we define a fuzzy monotone relationship as follows.

Definition 1: For any two partial order sets E and F, suppose $E = \{x_1, x_2, \ldots, x_n\}$, $F = \{y_1, y_2, \ldots, y_n\}$, where n is an integer. There is a one to one mapping f, $f : E \mapsto F$. We order E in ascending order to produce a new set of $E^a = \{x_1^a, x_2^a, \ldots, x_n^a\}$. Through the mapping $f : E^a \mapsto F^a$, we obtain the corresponding $F^a = \{y_1^a, y_2^a, \ldots, y_n^a\}$. If there exists such a partition l of E^a with $E^{pa} = E_1^a \bigcup E_2^a \bigcup \ldots \bigcup E_l^a$ and $E^{pa} = \{E_1^a, E_2^a, \ldots, E_l^a\}$, where E^{pa} is the l partition of E^a, $2 \le l \le n$, for any i, j and $1 \le i < j \le l$, then we have that the values of elements in E_j^a are all bigger than or equal to those in E_i^a and $E_i^a \bigcap E_j^a = \phi$, through the mapping f, a partition of F^a noted F^{pa} forms, and $F^{pa} = F_1^a \bigcup F_2^a \bigcup \ldots \bigcup F_l^a$ and $F^{pa} = \{F_1^a, F_2^a, \ldots, F_l^a\}$. There are corresponding $F_i^a, F_j^a \in F^{pa}$, suppose $\forall y_t \in F_i^a$, $\forall y_s \in F_j^a$, two sets $S_{\ge} = \{y_s, y_t | y_s \ge y_t\}$ and $S_{\le} = \{y_s, y_t | y_s \le y_t\}$ are defined. For $S_{\ge}$, two fuzzy membership functions $\mu_{S_{\ge}}(y_s) = |y_s|/|F_j^a|$ and $\mu_{S_{\ge}}(y_t) = |y_t|/|F_i^a|$ are defined; for $S_{\le}$, two fuzzy membership functions $\mu_{S_{\le}}(y_s) = |y_s|/|F_j^a|$ and $\mu_{S_{\le}}(y_t) = |y_t|/|F_i^a|$ are defined, where $|y_s|$, $|y_t|$, $|F_i^a|$, and $|F_j^a|$ respectively represent the cardinal number of y_s, y_t, F_i^a, and F_j^a for $S_{\ge}$ or $S_{\le}$. If certainly having

$$\mu_{S_{\ge}}(y_s) \ge 0.5 \quad and \quad \mu_{S_{\ge}}(y_t) \ge 0.5 \tag{1}$$

then the relationship between the sets of E and F is a fuzzy monotonically increasing relationship under the partition l and mapping f. Otherwise, if we have

$$\mu_{S_{\le}}(y_s) \ge 0.5 \quad and \quad \mu_{S_{\le}}(y_t) \ge 0.5 \tag{2}$$

then the relationship between the set E and F is a fuzzy monotonically decreasing relationship under the partition l

and the mapping f. Then (1) represents that the percentage of some elements in F_i^a is bigger than or equal to 0.5, and those elements whose values are smaller or equal to some y_s in F_j^a, whose percentage in F_j^a is also bigger than or equal to 0.5. In other words, (1) represents the percentage of some elements in F_j^a that is bigger than or equal to 0.5, and these elements' values are bigger than or equal to most elements' values in F_i^a, similarly to (2).

Some properties of Definition 1 are discussed as follows:

1) The number 0.5 is a suitable limit in Definition 1
When we set the limit lower than 0.5, such as 0.4, then there are some uncertain situations, $\mu_{S_\geq}(y_s) \geq 0.4$ and $\mu_{S_\geq}(y_t) \geq 0.4$ cannot deduce that the relationship is certainly to be fuzzy monotone increasing, for example, when $\mu_{S_\geq}(y_s) = 0.45$ and $\mu_{S_\geq}(y_t) = 0.45$, then $\mu_{S_\leq}(y_s) \geq 0.55$ and $\mu_{S_\leq}(y_t) \geq 0.55$, which means the relationship is fuzzy monotone decreasing not fuzzy monotone increasing. When we set the limit higher than 0.5, the result of judging the relationship can also be judged by the limit 0.5.
2) The character of strong or weak of the relationship
The bigger the values of $\mu_{S_\geq}(y_s)$ and $\mu_{S_\geq}(y_t)$ are, the stronger of the fuzzy monotone increasing is; the bigger the values of $\mu_{S_\leq}(y_t)$ and $\mu_{S_\leq}(y_t)$ are, the stronger of the fuzzy monotone decreasing is. However, the strong or weak character of the relationship is not only related to the values of fuzzy membership function, but also related to the range of partition l. The discussion is as follows.
3) A supposition of the value variation of fuzzy membership function according to the partition
We suppose that the effect of disturbance is limit, and the effect of disturbance is relatively weaker in the large range partition than in the small range partition, then the values of fuzzy monotone membership function are relatively bigger in the large range partition than in the small range partition in the general tendency.
4) The tendency of the value of fuzzy membership function according to the partition
We cannot deduce whether there exists fuzzy monotone relationship or not between the input and output according to only once partition. According to item 3, if the values of fuzzy membership function are increasing in the general tendency with the partition range from small to big, then we can judge that there exists the fuzzy monotone relationship between the input and output. The values of fuzzy membership function may be fluctuation in the increasing course of general tendency when the input is nonlinear with the output.

The differences between two kinds of these relationships can be summarized as follows: A monotone relationship is used to describe every value pairs between the input and output in linear order. A fuzzy monotone relationship, on the other hand, is used to characterize the element monotone relationship between different subset values of input and output sets, which are separated into some sets. Thus, a novel fuzzy monotone dependent relationship is introduced here. Because the fuzzy monotonically decreasing is similar to the fuzzy monotonically increasing, sometimes we just discuss the fuzzy monotonically increasing relationship in the following.

B. MONOTONE AND MAPPING

There certainly exists a one-to-one mapping among objects, condition attributes' values, and decision attributes' values in a decision table. If any object $e_s \in U$, then there certainly exist relevant $x_{is} \in C_{iv}$ and $y_{js} \in D_{jv}$. The monotone relationship between C_i and D_j is described by the below definitions.

Definition 2: For any $e_k, e_l \in U$, if there exists $y_{jk} \geq y_{jl} \Rightarrow x_{ik} \geq x_{il}$, then the relationship between decision attribute D_j and condition attribute C_i is called monotonically increasing dependence.

Proposition 1: For any $e_k, e_l \in U$, if $y_{jk} \geq y_{jl} \Rightarrow x_{ik} \geq x_{il}$, then $x_{ik} \geq x_{il} \Rightarrow y_{jk} \geq y_{jl}$.

Proof: If $x_{ik} \geq x_{il} \Rightarrow y_{jk} \geq y_{jl}$ does not come into existence, then there exists $x_{ik} \geq x_{il}$ and $y_{jk} \leq y_{jl}$, because $y_{jk} \leq y_{jl} \Rightarrow x_{ik} \leq x_{il}$ leads to a contradiction with $x_{ik} \geq x_{il}$. Thus, the proposition holds. ■

According to the above supposition, not only the value but also the order should be taken into consideration. Thus, there certainly exists mapping $f_1 : D_{jv} \mapsto C_{iv}$. For any $k \in \{1, 2, \ldots, n\}$, $y_{jk} \in D_{jv}, x_{ik} \in C_{iv}$, there is $f_1(y_{jk}) = x_{ik}$. Certainly, there also exists inverse mapping $f_1^{-1} : C_{iv} \mapsto D_{jv}$. Then we have $f_1^{-1}(x_{ik}) = y_{jk}$. Similarly, there exists mapping $g : U \mapsto \{D_{jv}, C_{iv}\}$. Then we have $g(e_k) = \{y_{jk}, x_{ik}\}$ and existing inverse mapping $g^{-1}(y_{jk}, x_{ik}) = e_k$. The relationship among objects, condition attributes' values, and decision attributes' values is apparently a one-to-one mapping relationship in a decision table, because the order is also taken into account except for the value.

C. FUZZY MONOTONE DEPENDENCE RELATIONSHIP IN A DECISION TABLE

We define a fuzzy monotone dependence relationship in a decision table according to intervals as follows:

Definition 3: For any condition attribute $C_i \in C$ and decision attribute $D_j \in D$ in a decision table, suppose that with n objects in U, there exist a set of C_i values $C_{iv} = \{x_{i1}, x_{i2}, \ldots, x_{in}\}$ and a set of D_j values $D_{jv} = \{y_{j1}, y_{j2}, \ldots, y_{jn}\}$. If there is such a partition p that C_{iv} and D_{jv} satisfy Definition 1, the relationship between condition attribute C_i and decision attribute D_j is then a fuzzy monotone dependence relationship with respect to the partition p.

After reordering D_{jv} and C_{iv} separately in ascending order, we can obtain two sets of new data values: $D'_{jv} = \{y'_{j1}, y'_{j2}, \ldots, y'_{jn}\}$ and $C'_{iv} = \{x'_{i1}, x'_{i2}, \ldots, x'_{in}\}$. Similarly, a new set data of values $C^d_{iv} = \{x^d_{i1}, x^d_{i2}, \ldots, x^d_{in}\}$ is formed by reordering C_{iv} in descending order. Using the above mapping g^{-1}, we form three new sets of reordered objects:

$U_{D_j} = \{e_{D1}, e_{D2}, \ldots, e_{Dn}\}$, $U_{C_i} = \{e_{C1}, e_{C2}, \ldots, e_{Cn}\}$ and $U^d_{C_i} = \{e^d_{C1}, e^d_{C2}, \ldots, e^d_{Cn}\}$, where we have $g^{-1} : D'_{jv} \mapsto U_{D_j}$, $g^{-1} : C'_{iv} \mapsto U_{C_i}$ and $g^{-1} : C^d_{iv} \mapsto U^d_{C_i}$. There is a partial order relationship among elements in U_{C_i}, U_{D_j} and $U^d_{C_i}$. If the partial order relationship is removed from U_{C_i}, U_{D_j}, $U^d_{C_i}$ and U sets, then these sets would become equal sets. For searching for the inclusion degree in the corresponding intervals of U_{C_i}, U_{D_j} and $U^d_{C_i}$, we discuss the fuzzy inclusive monotone relationship in the following.

D. FUZZY INCLUSIVE MONOTONE RELATIONSHIP MODEL AND MEMBERSHIP FUNCTION

We partition U_{C_i}, U_{D_j} and $U^d_{C_i}$ into intervals according to the element number $k(1 \leq k \leq \llcorner n/2 \lrcorner)$ for each interval. Then there are $p = \ulcorner n/k \urcorner$ intervals after this partition. As a result, let $U'_{D_j} = U_{D1} \bigcup \ldots \bigcup U_{Dp}$ and $U'_{D_j} = \{U_{D1}, U_{D2}, \ldots, U_{Dp}\}$ corresponding to U_{D_j}, where $U_{Dv} = \{e_{D(k*(v-1)+1),\ldots,e_{D(k*v)}}\}$, $1 \leq v < p$ and $U_{Dp} = \{e_{D(k*(p-1)+1),\ldots,e_{Dn}}\}$. let $U'_{C_i} = U_{C1} \bigcup \ldots \bigcup U_{Cp}$ and $U'_{C_i} = \{U_{C1}, U_{C2}, \ldots, U_{Cp}\}$ corresponding to U_{C_i}, where we have $U_{Cv} = \{e_{C(k*(v-1)+1)}, \ldots, e_{C(k*v)}\}$, $1 \leq v < p$ and $U_{Cp} = \{e_{C(k*(p-1)+1)}, \ldots, e_{Cn}\}$. let $U^{dp}_{C_i} = U^d_{C1} \bigcup \ldots \bigcup U^d_{Cp}$ and $U^{dp}_{C_i} = \{U^d_{C1}, U^d_{C2}, \ldots, U^d_{Cp}\}$ corresponding to $U^d_{C_i}$, where we have $U^d_{Cv} = \{e^d_{C(k*(v-1)+1)}, \ldots, e^d_{C(k*v)}\}$, $1 \leq v < p$ and $U^d_{Cp} = \{e^d_{C(k*(p-1)+1)}, \ldots, e^d_{Cn}\}$. Then a partial order relationship exists among these partitioned intervals. We denote the partition as the p partition. Here, we define an inclusion degree $\mu(A, B) = \frac{|A \bigcap B|}{|A|} = I(B/A), A, B \subseteq U$, where $|A|$ is the cardinal number of A. Let $A = \phi$ then we have $I(B/A) = 1$. Then when we use the inclusion degree to conclude a fuzzy monotone relationship. A fuzzy monotone relationship is also called as a fuzzy inclusive monotone relationship here. Apparently, a fuzzy inclusive monotone relationship is certainly a fuzzy monotone relationship.

According to the above D_{jv}, C_{iv}, D'_{jv}, C'_{iv}, C^d_{iv}, mapping g and g^{-1}, U_{D_j}, U_{C_i}, $U^d_{C_i}$, U'_{C_i}, U'_{D_j} and $U^{dp}_{C_i}$, we present two propositions as follows:

Proposition 2: After the p partition for U_{C_i}, U_{D_j} and $U^d_{C_i}$, for any two $U_{Dq}, U_{Dr} \in U'_{D_j}, q < r$, there are two $U_{Ch}, U_{Cl} \in U'_{C_i}, h < l$, and two $U^d_{Cs}, U^d_{Ct} \in U^{dp}_{C_i}, s < t$. If $\mu(U_{Dq}, U_{Ch}) \geq 0.5$ and $\mu(U_{Dr}, U_{Cl}) \geq 0.5$, then the relationship between decision attribute D_j and condition attribute C_i is a fuzzy inclusive monotonically increasing relationship with respect to the p partition. Otherwise, if $(\mu(U_{Dq}, U_{Cl}) \geq 0.5 \; \textit{and} \; \mu(U_{Dr}, U_{Ch}) \geq 0.5)$ or $(\mu(U_{Dq}, U^d_{Cs}) \geq 0.5 \; \textit{and} \; \mu(U_{Dr}, U^d_{Ct}) \geq 0.5)$, then the relationship between decision attribute D_j and condition attribute C_i is a fuzzy inclusive monotonically decreasing relationship with respect to the p partition.

Proof: Apparently, U_{Dq} and U_{Dr} are included in U. When $\mu(U_{Dr}, U_{Cl}) \geq 0.5$ and $\mu(U_{Dq}, U_{Ch}) \geq 0.5$, then the intervals of U_{Dr} and U_{Cl} have most common objects with respect to the partition of the decision attribute and the condition attribute, which are similar to U_{Dq} and U_{Ch}. Because $q < r$, $h < l$, and sets of D'_{jv} and C'_{iv} are in ascending order, through the mapping g, D_{jv} and C_{iv} are suited to (1) in Definition 1. Then the relationship between D_{jv} and C_{iv} is proved to be a fuzzy inclusive monotonically increasing dependence relationship. Similarly, when $(\mu(U_{Dq}, U_{Cl}) \geq 0.5 \; \textit{and} \; \mu(U_{Dr}, U_{Ch}) \geq 0.5)$ or $(\mu(U_{Dq}, U^d_{Cs}) \geq 0.5 \; \textit{and} \; \mu(U_{Dr}, U^d_{Ct}) \geq 0.5)$, D_{jv} and C_{iv} satisfy (2) in Definition 1. Then the relationship between D_{jv} and C_{iv} is a fuzzy inclusive monotonically decreasing dependence relationship. ∎

Propostion 3: After the p partition for U_{D_j}, U_{C_i} and $U^d_{C_i}$, for any x, y and z, where $1 \leq x \leq p$, $1 \leq y \leq p$, $1 \leq z \leq p$, $U_{Dx} \in U'_{D_j}$, $U_{Cy} \in U'_{C_i}$, $U^d_{Cz} \in U^{dp}_{C_i}$. If and only if $x = y$, and there exists $\mu(U_{Dx}, U_{Cy}) \geq 0.5$, then the relationship between the decision attribute D_j and the condition attribute C_i is a fuzzy inclusive monotonically increasing relationship with respect to the p partition. Otherwise, if and only if $x = z$, and there exists $\mu(U_{Dx}, U^d_{Cz}) \geq 0.5$, then the relationship between the decision attribute D_j and the condition attribute C_i is a fuzzy inclusive monotonically decreasing relationship with respect to the p partition.

Proof: First, we prove the condition for the fuzzy inclusive monotonically increasing relationship, for any x', which is the suffix of a set in U'_{D_j} and $x' \neq x$, if $\mu(U_{Dx}, U_{Cy}) \geq 0.5$, then we can conclude that $\mu(U_{Dx'}, U_{Cy}) \leq 0.5$. Similarly, for any y', which is the suffix of a set in U'_{C_i} and $y' \neq y$, there certainly exists $\mu(U_{Dx}, U_{Cy'}) \leq 0.5$. According to Proposition 2, if $x = y$, we have $\mu(U_{Dx}, U_{Cy}) \geq 0.5$. Then the conclusion is proved. Thus, we just need to search for the case of $x \neq y$ and $\mu(U_{Dx}, U_{Cy}) \geq 0.5$. Let $x < y$, for any x_1, where $1 \leq x_1 < x$. There exists y_1, where $1 \leq y_1 < y$ and $\mu(U_{Dx_1}, U_{Cy_1}) \geq 0.5$. Otherwise, the relationship is apparently not one of fuzzy inclusive monotonically increasing dependence. This is because in the case of $y < y_1$ and $\mu(U_{Dx_1}, U_{Cy_1}) \geq 0.5$, the conditions $x_1 < x$, $y < y_1$, $\mu(U_{Dx}, U_{Cy}) \geq 0.5$, and $\mu(U_{Dx_1}, U_{Cy_1}) \geq 0.5$ would lead to contradicting the fuzzy inclusive monotonically increasing relationship in Proposition 2. Then there certainly exists a y_2 satisfying $\mu(U_{Dx_1}, U_{Cy_2}) \leq 0.5$ for any x_1, where $1 \leq y_2 < y$, because $x < y$, according to the pigeonhole principle of combinatorics. Thus, there is only one possibility of being a fuzzy inclusive monotonically increasing relationship between D_j and C_i; that is, there exists a certain x'' where $x'' > x$ and $\mu(U_{Dx''}, U_{Cy_2}) \geq 0.5$. However, if $\mu(U_{Dx''}, U_{Cy_2}) \geq 0.5$, because of $x'' > x$, $y_2 < y$, and $\mu(U_{Dx}, U_{Cy}) \geq 0.5$, it would lead to the contradiction of the fuzzy inclusive monotonically increasing relationship in Proposition 2. Then the relationship between decision attribute D_j and condition attribute C_i is not a fuzzy inclusive monotonically increasing relationship according to Proposition 2. Similarly, in the case of $x > y$, the relationship between the decision attribute and condition attribute is also not a fuzzy inclusive monotonically increasing relationship according to Proposition 2. Thus, if and only if $x = y$, there exists $\mu(U_{Dx}, U_{Cy}) \geq 0.5$, then the relationship between decision attribute D_j and condition attribute

C_i is a fuzzy inclusive monotonically increasing relationship with respect to the p partition. Similarly, we can prove that if and only if $x = z$, there exists $\mu(U_{Dx}, U^d_{Cz}) \geq 0.5$, and then the relationship between the decision attribute D_j and the condition attribute C_i is a fuzzy inclusive monotonically decreasing relationship with respect to the p partition. The proposition holds. ■

According to Proposition 3, we can write the fuzzy inclusive monotonically decision membership function (FIMDMF) between decision attribute D_j and condition attribute C_i as follows:

$$\underset{p}{\mu}(D_j, C_i) = \begin{cases} \bigwedge_{v=1}^{p} \mu(U_{Dv}, U_{Cv}), & \mu(U_{Dv}, U_{Cv}) \neq 0, \quad \text{FI} \\ \bigwedge_{v=1}^{p} \mu(U_{Dv}, U^d_{Cv}), & \mu(U_{Dv}, U^d_{Cv}) \neq 0, \quad \text{FD} \\ 0, & \text{others}, \quad \text{NM} \end{cases} \tag{3}$$

where the FI represents the fuzzy inclusive monotonically increasing in (3),the FD represents the fuzzy inclusive monotonically decreasing in (3), and the NM represents none of the monotone. Thus (3) denotes that the relationship between D_j and C_i is a fuzzy inclusive monotonically increasing or decreasing relationship or none of the fuzzy inclusive monotone relationship.

Proposition 4: A monotone relationship is a special case of its corresponding fuzzy inclusive monotone relationship.

Proof: In the case of $p = n$ and $\underset{p}{\mu}(D_j, C_i) = \bigwedge_{v=1}^{p} \mu(U_{Dv}, U_{Cv}) = 1$, the relationship between decision attribute D_j and condition attribute C_i apparently is a monotonically increasing relationship. In the case of $p = n$ and $\underset{p}{\mu}(D_j, C_i) = \bigwedge_{v=1}^{p} \mu(U_{Dv}, U^d_{Cv}) = 1$, the relationship between decision attribute D_j and condition attribute C_i apparently is a monotonically decreasing relationship. ■

E. DISCUSSIONS ON CORRELATIVE PARAMETERS OF THE FUZZY INCLUSIVE MONOTONE MODEL

In this section, we discuss how the element number k of one interval and interval partition number p to influence the degree of fuzzy inclusive monotone relationship. If the relationship between decision attribute D_j and condition attribute C_i is certainly a fuzzy inclusive monotone, the range of intervals will become relatively smaller and the influence of disturbances will become relatively stronger with the smaller k and larger p. The value of FIMDMF will then become relatively smaller. Thus, if the value of FIMDMF is larger than or equal to 0.5 in case k is between 1 and $\llcorner n/2 \lrcorner$, and the value of FIMDMF tends to become larger with the increasing k, then the relationship between decision attribute D_j and condition attribute C_i is possibly a fuzzy inclusive monotone relationship. If the value of FIMDMF is larger than or equal to 0.5 in case k is around $\llcorner n/2 \lrcorner$, then the degree of the fuzzy inclusive monotone relationship is, however, relatively weak and easily disturbed. On the other hand, if the value of FIMDMF is larger than 0.5 in case k is far away from $\llcorner n/2 \lrcorner$, then the degree of the fuzzy inclusive monotone relationship is strong and not easily disturbed. In addition, it is obvious that the larger the value of FIMDMF, the stronger the degree of the fuzzy inclusive monotone relationship is. Because the *pth* interval is a remainder interval and the element number of the interval is mostly less than k, when happened, we mainly consider the values of FIMDMF of intervals' number from 1 to $p - 1$.

IV. FIMDMF ALGORITHM

According to the proposed fuzzy inclusive monotone model, an algorithm called FIMDMF is presented to determine whether decision attribute D_j is fuzzy inclusive monotone with condition attribute C_i or not. If does, the attribute C_i will be included in the reduced fuzzy inclusive monotonically increasing or decreasing attribute sets according to the attribute D_j. The algorithm details as follows:

Algorithm 1 The FIMDMF Algorithm

Output: a set of fuzzy inclusive increasing or decreasing reduced condition attributes with respect to the decision attribute.

Initialized: *fuzzy_increasing* = {}; *fuzzy_decreasing* = {}.
for each $C_i \in C$ **do**
  Form three new sets D'_{jv}, C'_{iv} and C^d_{iv} by ranking the decision attribute value set D_{jv} in ascending order, and the condition attribute value set C_{iv} in ascending and descending order, respectively.
  Rank objects in the U set according to the decision attribute value set D'_{jv} and the condition attribute value set C'_{iv} and C^d_{iv}, respectively. Three reorder objects sets U_{D_j}, U_{C_i} and $U^d_{C_i}$ are then formed.
  for $k \leftarrow 1$ *to* $\llcorner n/2 \lrcorner$ *step* 1 **do**
    if $(n \;\; mod \;\; k) = 0$ **then** ▷ mod is calculating residue
      $p \leftarrow n/k$
    else
      $p \leftarrow \ulcorner n/k \urcorner - 1$
    end if
    $U\_C_i[k] \leftarrow \bigwedge_{v=1}^{p} \mu(U_{Dv}, U_{Cv})$
    if $U\_C_i[k] \geq 0.5$ **then**
      $monotone[k] \leftarrow 1$
    else if $\bigwedge_{v=1}^{p} \mu(U_{Dv}, U^d_{Cv}) \geq 0.5$ **then**
      $U\_C_i[k] \leftarrow \bigwedge_{v=1}^{p} \mu(U_{Dv}, U^d_{Cv})$
      $monotone[k] \leftarrow 2$
    else
      $U\_C_i[k] \leftarrow 0$
      $monotone[k] \leftarrow 0$
    end if

21: **end for**
22: $count1 \leftarrow 0$
23: $count2 \leftarrow 0$
24: **for** $k \leftarrow \llcorner n/2 \lrcorner$ *to* $\llcorner n * st \lrcorner$ *step* -1 **do** ▷ st is the percent between 1% and 50%, and assigned a value according to the situation
25: **if** $U\_C_i[k] = 0$ **then**
26: break
27: **else if** $monotone[k] = 1$ *and* $U\_C_i[k] \geq sn$ **then** ▷ sn is a number bigger than 0.5, and assigned a value according to the situation
28: $count1 \leftarrow count1 + 1$
29: **else if** $monotone[k] = 2$ *and* $U\_C_i[k] \geq sn$ **then**
30: $count2 \leftarrow count2 + 1$
31: **end if**
32: **end for**
33: **if** $count1 \geq sc$ **then** ▷ sc is a number between 1 and ($\llcorner n/2 \lrcorner - \llcorner n * st \lrcorner$), and assigned a value according to the situation
34: $fuzzy\_increasing \leftarrow fuzzy\_increasing \bigcup C_i$
35: **else if** $count2 \geq sc$ **then**
36: $fuzzy\_decreasing \leftarrow fuzzy\_decreasing \bigcup C_i$
37: **end if**
38: **end for**

The time complexity of the algorithm is mainly in the number 2 line circulation and the number 5 line circulation. The circulation number in the 2 line is m and the circulation number in the 5 line is $\lfloor n/2 \rfloor$. Thus, the time complex of the algorithm is about $O(m * n/2)$.

Example 2: We use the example given in TABLE 1 to explain our FIMDMF algorithm. After ranking the value sets of $C1, C2, C3$ and D in ascending order, respectively, we can get $U_{C1} = \{e1, e3, e4, e2, e6, e5, e7, e9, e8\}$, $U_{C2} = \{e9, e5, e2, e7, e4, e3, e8, e6, e1\}$, $U_{C3} = \{e2, e4, e6, e1, e3, e5, e9, e7, e8\}$, and $U_D = \{e1, e2, e4, e3, e6, e5, e8, e7, e9\}$. We just select the element number k from 3 to 4. when $k = 3$, and $p = 3$, then the $C1, C2, C3$ and D is partitioned into $U_{C1} = \{\{e1, e3, e4\}, \{e2, e6, e5\}, \{e7, e9, e8\}\}$, $U_{C2} = \{\{e9, e5, e2\}, \{e7, e4, e3\}, \{e8, e6, e1\}\}$, $U_{C3} = \{\{e2, e4, e6\}, \{e1, e3, e5\}, \{e9, e7, e8\}\}$, and $U_D = \{\{e1, e2, e4\}, \{e3, e6, e5\}, \{e8, e7, e9\}\}$, respectively. We simply express them as $U_{C1} = \{U_{C11}, U_{C12}, U_{C13}\}$, $U_{C2} = \{U_{C21}, U_{C22}, U_{C23}\}$, $U_{C3} = \{U_{C31}, U_{C32}, U_{C33}\}$, and $U_D = \{U_{D1}, U_{D2}, U_{D3}\}$. Apparently, we have $\mu(U_{D1}, U_{C11}) = \frac{2}{3}$, $\mu(U_{D2}, U_{C12}) = \frac{2}{3}$, $\mu(U_{D3}, U_{C13}) = 1$, $\mu(U_{D1}, U_{C21}) = \frac{1}{3}$, $\mu(U_{D2}, U_{C22}) = \frac{1}{3}$, $\mu(U_{D3}, U_{C23}) = \frac{1}{3}$, $\mu(U_{D1}, U_{C31}) = \frac{2}{3}$, $\mu(U_{D2}, U_{C32}) = \frac{2}{3}$ and $\mu(U_{D3}, U_{C33}) = 1$. We then have $\underset{p}{\mu}(U_D, U_{C1}) = \bigwedge_{v=1}^{3} \mu(U_{Dv}, U_{C1v}) = \frac{2}{3} \geq 0.5$, $\underset{p}{\mu}(U_D, U_{C2}) = \bigwedge_{v=1}^{3} \mu(U_{Dv}, U_{C2v}) = \frac{1}{3} < 0.5$ and $\underset{p}{\mu}(U_D, U_{C3}) = \bigwedge_{v=1}^{3} \mu(U_{Dv}, U_{C3v}) = \frac{2}{3} \geq 0.5$. When $k = 4$, and $p = 3$, then the partition result for $C1, C2, C3$ and D is that $U_{C1} = \{\{e1, e3, e4, e2\}, \{e6, e5, e7, e9\}, \{e8\}\}$, $U_{C2} = \{\{e9, e5, e2, e7\}, \{e4, e3, e8, e6\}, \{e1\}\}$, $U_{C3} = \{\{e2, e4, e6, e1\}, \{e3, e5, e9, e7\}, \{e8\}\}$, and $U_D = \{\{e1, e2, e4, e3\}, \{e6, e5, e8, e7\}, \{e9\}\}$, similarly we can compute $\underset{p-1}{\mu}(U_D, U_{C1}) = \bigwedge_{v=1}^{2} \mu(U_{Dv}, U_{C1v}) = \frac{3}{4} \geq 0.5$, $\underset{p-1}{\mu}(U_D, U_{C2}) = \bigwedge_{v=1}^{2} \mu(U_{Dv}, U_{C2v}) = \frac{1}{4} < 0.5$ and $\underset{p-1}{\mu}(U_D, U_{C3}) = \bigwedge_{v=1}^{2} \mu(U_{Dv}, U_{C3v}) = 0.5 \geq 0.5$, then a decision filtration rule is designed as when $k \geq 3$, the values of FIMDMF must be bigger than or equal to 0.5 for fuzzy monotonically increasing. According to the rule, $C1$ and $C3$ are included in the reduced condition attribute set for fuzzy monotonically increasing, while $C2$ is not. The degree of $C1$ fuzzy monotonically increasing with D is apparently stronger than that of $C3$.

By using this algorithm, we can easily find out which fuzzy monotonically increasing or decreasing of condition attributes are important for the certain decision attribute and compute the association degree of those condition attributes. We can then determine those important condition attributes for the decision attribute and obtain a reduction set of condition attributes for the decision attribute.In addition, we can analyse the partial correlative dependence between the input and output ends by using our proposed method.

V. EXPERIMENTS

The methods of dominance and fuzzy neighborhood rough set often deal with the continuous data sets in the OIS [24]–[31]. In this section, we compared our method named *FIMDMF* with them. The method of the variable-precision-dominance-based rough set (*VPDRS*) [25]–[31] is suitable for dealing with continuous data in ordered information systems. The fuzzy neighborhood rough method (*FNRS*) [24] deals with the continuous input data under a precondition of classified output decision data. Then the continuous data at both inputs and output were directly done in the experiments on the *VPDRS* method, and the experiment was done on the *FNRS* method after having the output decision data set classified through the dominance method. First, we describe the core part of *VPDRS* as follows:

Definition 4: Let $S = (U, At \bigcup D, V, F)$ be a decision information table. For $\forall R \subseteq At$, the β-dependency degree of R with respect to D is defined as follows [27], [28], [31]:

$$DEP_{R^{\succeq}_{\beta}}(D) = \frac{1}{r} \sum_{D_j \in U/D^{\succeq}} \frac{|\underline{R}^{\succeq}_{\beta}(D_j)|}{|U|}$$

where $U/D^{\succeq} = \{D_1, D_2, \ldots, D_r\}$ consists of decision classes induced by a dominance relationship, for $\forall a \in At - R$, the β-dominance significance measure of a in At is defined as

$$Sig^{\succeq}_{\beta}(a, R, D, U) = DEP_{(R \bigcup a)^{\succeq}_{\beta}}(D) - DEP_{R^{\succeq}_{\beta}}(D)$$

Second, the core part of FNRS was described as follows:

Definition 5: Let $S = (U, At \bigcup D, V, F)$ be a decision information table. Given a neighborhood radius λ, and $B \subseteq AT$, and the fuzzy decision $\{\widetilde{D}_1, \widetilde{D}_2, \ldots, \widetilde{D}_r\}$ induced by D. N_B is the fuzzy similarity relation on U induced by B. Then the variable precision lower and upper approximations of D with respect to B are defined as follows, respectively [24], [27], [28], [31].

$$\underline{N}_B^{\lambda,\alpha}(D) = \{\underline{N}_B^{\lambda,\alpha}(\widetilde{D}_1), \underline{N}_B^{\lambda,\alpha}(\widetilde{D}_2), \ldots, \underline{N}_B^{\lambda,\alpha}(\widetilde{D}_r)\}$$
$$\overline{N}_B^{\lambda,\beta}(D) = \{\overline{N}_B^{\lambda,\beta}(\widetilde{D}_1), \overline{N}_B^{\lambda,\beta}(\widetilde{D}_2), \ldots, \overline{N}_B^{\lambda,\beta}(\widetilde{D}_r)\}$$

where we have

$$\underline{N}_B^{\lambda,\alpha}(\widetilde{D}_i) = \{x_i \in D_i | I([x_i]_B^\lambda, \widetilde{D}_i) \geq \alpha\}, \quad 0.5 \leq \alpha \leq 1$$
$$\overline{N}_B^{\lambda,\beta}(\widetilde{D}_i) = \{x_i \in D_i | I([x_i]_B^\lambda, \widetilde{D}_i) \geq \beta\}, \quad 0 \leq \beta \leq 0.5$$

and $\underline{N}_B^{\lambda,\alpha}(D) = \bigcup_{i=1}^r \underline{N}_B^{\lambda,\alpha}(\widetilde{D}_i)$, the variable precision dependency of D on B is defined as

$$\sigma_B^{\lambda,\alpha}(D) = \frac{|\underline{N}_B^{\lambda,\alpha}(D)|}{|U|}$$

for $\forall a \in At - B$, the significance of a with respect to B in At is defined as

$$SIG^{\lambda,\alpha}(a, B, D) = \sigma_{B\bigcup\{a\}}^{\lambda,\alpha}(D) - \sigma_B^{\lambda,\alpha}(D)$$

More details of VPDRS and FNRS are referred to [24]–[31]. Because the optimized feature selection was proved to be a NP-hard problem through rough sets methods [23], [33], the heuristic algorithms were adopted in almost all feature selection or attribute reduction methods based on rough set theory. So did the VPDRS and FNRS methods. The heuristic algorithms in VPDRS and FNRS have several steps. The first step is to set the R and B to respective null set according to Definitions 4 and 5. The second step iteratively searches for the significance of each a in $At - R$ and $At - B$, and selects the a with the biggest significance value into R and B separately. The third step is that the iteration is ended for VPDRS until the significance of each a in $At - R$ is equal to zero, and for FNRS until the significance of each a in $At - B$ is smaller than or equal to zero. Finally, the R and B become the reduction set for VPDRS and FNRS, respectively. Matlab was adopted to do the experiments.

For the attribute reduction, we conducted the experiments to validate our FIMDMF method, and compare it with VPDRS method, FNRS method, and the method denoted as FMF in [32], respectively. The decision filter rule used in FIMDMF for the experiments is that the FIMDMF values of the input condition attribute a are not equal to zero, and at least one-third FIMDMF values are bigger than 0.545 when the element number k of partition intervals is between $\lfloor n/2 \rfloor$ and $\lfloor n/2 \rfloor - \lfloor 0.02 \times n \rfloor$, where the a is selected to the corresponding reduction set. In order to verify the effectiveness of the methods, not only the runtime, but also the error rate are used for experimental metrics. The error rate is defined as follows: suppose the core attributes reduction set is denoted as $core$, the attribute reduction set for one method is denoted as ars and the error rate as er, then the er is expressed as $er = \frac{|\{core-ars\}\bigcup\{ars-core\}|}{|core|}$, where $|\{ars - core\} \bigcup \{core - ars\}|$ is the cardinal number of the symmetric difference between ars and $core$, and $|core|$ the cardinal number of $core$. The core attributes reduction set is gained from the known documents. The lower the error rate is, the better the method is. For correlation analysis, we compare our method with the Spearman and Pearson methods [38]–[40]. Because these two methods are popular in correlation ananlysis.

TABLE 2. FIMDMF result.

Runtime	Reduction set	er
0.083133	C1,C3	0

TABLE 3. FMF result.

Runtime	Reduction set	er
0.2459	C1,C3	0

TABLE 4. VPDRS results.

Parameter β	Runtime	Reduction set	er
0.5	0.24508	C1,C2,C3	0.5
0.6	0.24357	C1,C2,C3	0.5
0.7	0.23053	C1,C2,C3	0.5
0.8	0.22734	C1,C2,C3	0.5
0.9	0.22056	C1,C2	1
1.0	0.24982	C1,C2	1

A. COMPARISONS AGAINST THE DATA IN TABLE 1

In order to illustrate the variation relationship between condition attributes and decision attributes, we transform the data in TABLE 1 by using $y = (x-min(a))/(max(a)-min(a))$, where x, $max(a)$ and $min(a)$ are respectively a value, maximum value, and minimum value in the attribute a values' set. Fig.1, Fig.2 and Fig.3 show the respective relationship between condition attributes C1, C2, C3 and decision attribute D. From the three figures, we can easily find that the variation tendency of C1 and C3 apparently is similar to that of D except for C2. Then the selected core condition attributes are C1 and C3.

The experiment results on compared algorithms are reported in TABLE 2 to TABLE 5, respectively. From the results, FIMDMF performs in the shortest running time with no errors. However, the VPDRS and FNRS methods have the higher error rates. The FIMDMF and FMF methods are more effective in this case.

For correlation analysis, the experiment results of Peason and Spearman methods are reported in TABLEs 6 and 7. From TABLE 6, only the input C1 is correlated with D by the Pearson method because the p-values of C2 and C3 are bigger than 0.05. From TABLE 7, both C1 and C3 are considered to be correlated with D because their r-values are bigger than 0.7 when their p-values are smaller 0.05. So the FIMDMF, FMF and Spearman methods have the same correct result in this example.

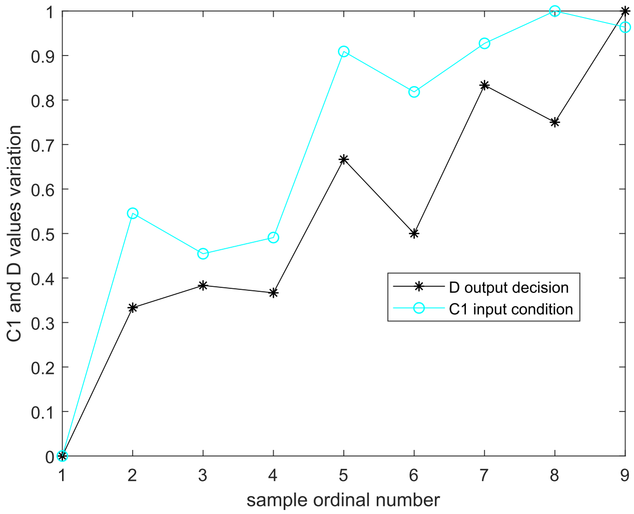

FIGURE 1. Variation tendency between C1 and D.

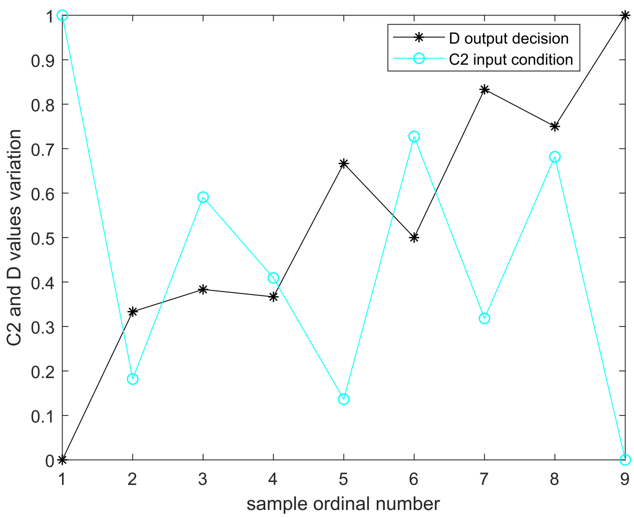

FIGURE 2. Variation tendency between C2 and D.

TABLE 5. FNRS results.

Parameter ρ, λ, α	Runtime	Reduction set	er
$\rho = 0.5, \lambda = 0.1, \alpha = 0.5$	0.2038	null	1
$\rho = 0.5, \lambda = 0.1, \alpha = 0.6$	0.017229	null	1
$\vdots$	$\vdots$	null	1
$\rho = 0.5, \lambda = 0.1, \alpha = 1$	0.18954	null	1
$\rho = 0.5, \lambda = 0.2, \alpha = 0.5$	0.042635	null	1
$\rho = 0.5, \lambda = 0.2, \alpha = 0.6$	0.015982	null	1
$\vdots$	$\vdots$	null	1
$\rho = 0.5, \lambda = 0.2, \alpha = 1$	0.014835	null	1
$\rho = 0.5, \lambda = 0.3, \alpha = 0.5$	0.01587	null	1
$\rho = 0.5, \lambda = 0.3, \alpha = 0.6$	0.015019	null	1
$\vdots$	$\vdots$	null	1
$\rho = 0.5, \lambda = 0.3, \alpha = 1$	0.015893	null	1
$\rho = 0.5, \lambda = 0.4, \alpha = 0.5$	0.066693	null	1
$\rho = 0.5, \lambda = 0.4, \alpha = 0.6$	0.014425	null	1
$\vdots$	$\vdots$	null	1
$\rho = 0.5, \lambda = 0.4, \alpha = 1$	0.014308	null	1
$\rho = 0.5, \lambda = 0.5, \alpha = 0.5$	0.21261	null	1
$\rho = 0.5, \lambda = 0.5, \alpha = 0.6$	0.017545	null	1
$\vdots$	$\vdots$	null	1
$\rho = 0.5, \lambda = 0.5, \alpha = 1$	0.0158	null	1

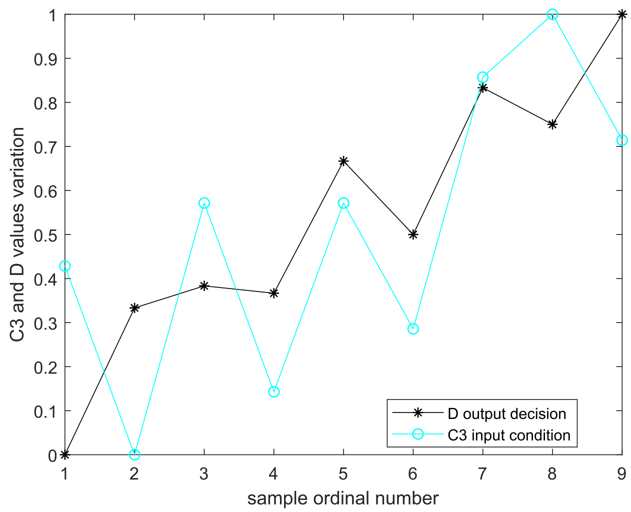

FIGURE 3. Variation tendency between C3 and D.

TABLE 6. The results of Pearson between the inputs and output D.

Input variables	r-values	p-values
C1	0.9312	2.6242e-04(< 0.01)
C2	-0.6344	0.0665(> 0.05)
C3	0.6461	0.0601(> 0.05)

B. COMPARISONS AGAINST UCI WATER TREATMENT DATA

In this experiment, we use the UCI Machine Learning Repository wastewater treatment data set [41]. We removed the incomplete data to obtain a complete decision table. After pre-processing, the dataset in this experiment consists of 380 data objects with 38 attributes. The 38 attributes contain 22 input condition attributes and 7 output decision attributes. The 24th attribute called DBO-S (output biological demand of oxygen) is an important one in the wastewater treatment. Therefore, the DBO-S attribute acts as a decision attribute to determine important input condition attributes from 22 input condition attributes. The details of the 22 condition attributes and decision attribute DBO-S in UCI water treatment data are listed in TABLE 8

Because the maximum of the DBO-S samples is three times more than that of the other DBO-S samples, we excluded the DBO-S samples with maximum values. Then we conducted the experiments on the rest of 379 samples. From the reported results of water treatment [42], [43], we know that the important core input attributes for the DBO-S output include biological demand of oxygen, chemical demand of oxygen, suspended solids and sediments. These core inputs to the secondary settler are more important than to the primary settler. As such, we denote the set of the core input attributes as the core $= \{4, 5, 6, 8, 11, 12, 14, 17, 18, 19, 21\}$.

TABLE 7. The results of Spearman between the inputs and output D.

Input variables	r-values	p-values
C1	0.8833	$0.0031(< 0.01)$
C2	-0.4833	$0.1938(> 0.05)$
C3	0.7866	$0.0153(< 0.05)$

TABLE 8. UCI water treatment data information.

Number	Attribute Name	Representation
1	Q-E	input flow to plant
2	ZN-E	input Zinc to plant
3	PH-E	input pH to plant
4	DBO-E	input Biological demand of oxygen to plant
5	DQO-E	input chemical demand of oxygen to plant
6	SS-E	input suspended solids to plant
7	SSV-E	input volatile suspended solids to plant
8	SED-E	input sediments to plant
9	COND-E	input conductivity to plant
10	PH-P	input pH to primary settler
11	DBO-P	input Biological demand of oxygen to primary settler
12	SS-P	input suspended solids to primary settler
13	SSV-P	input volatile suspended solids to primary settler
14	SED-P	input sediments to primary settler
15	COND-P	input conductivity to primary settler
16	PH-D	input pH to secondary settler
17	DBO-D	input Biological demand of oxygen to secondary settler
18	DQO-D	input chemical demand of oxygen to secondary settler
19	SS-D	input suspended solids to secondary settler
20	SSV-D	input volatile suspended solids to secondary settler
21	SED-D	input sediments to secondary settler
22	COND-D	input conductivity to secondary settler
24	DBO-S	output Biological demand for oxygen

TABLE 9. FIMDMF result for DBO-S.

Runtime	Fuzzy monotone increasing	Fuzzy monotone decreasing	er
1.43	4,5,6,8,11,12,13,14,17,18,19,21,22	1	0.2727

TABLE 10. FMF result for DBO-S.

Runtime	Fuzzy monotone increasing	Fuzzy monotone decreasing	er
0.37476	4,5,6,11,14,17,18,19,21	none	0.1818

TABLE 11. VPDRS results for DBO-S.

Parameter β	Runtime	Reduction set	er
0.5	1646.063	All attributes(no reduction)	1
0.6	1660.3301	All attributes(no reduction)	1
0.7	1585.7009	All attributes(no reduction)	1
0.8	1153.27	2,4,10,12,13,14,17,18,19	0.727
0.9	1470.1775	1,2,3,4,5,6,8,9,13,14,16,17,18,19,20,21	0.909
1.0	1495.8724	1,2,4,5,6,7,8,9,10,12,13,14,16,17,18,19,20,21	0.818

TABLE 12. FNRS results for DBO-S.

Parameter ρ, λ, α	Runtime	Reduction set	er
$\rho = 0.5, \lambda = 0.1, \alpha = 0.5$	15.3634	null	1
$\rho = 0.5, \lambda = 0.1, \alpha = 0.6$	15.1373	null	1
$\vdots$	$\vdots$	null	1
$\rho = 0.5, \lambda = 0.1, \alpha = 1$	15.3279	null	1
$\rho = 0.5, \lambda = 0.2, \alpha = 0.5$	15.229	null	1
$\rho = 0.5, \lambda = 0.2, \alpha = 0.6$	15.518	null	1
$\vdots$	$\vdots$	null	1
$\rho = 0.5, \lambda = 0.2, \alpha = 1$	15.3065	null	1
$\rho = 0.5, \lambda = 0.3, \alpha = 0.5$	15.4535	null	1
$\rho = 0.5, \lambda = 0.3, \alpha = 0.6$	15.622	null	1
$\vdots$	$\vdots$	null	1
$\rho = 0.5, \lambda = 0.3, \alpha = 1$	15.2499	null	1
$\rho = 0.5, \lambda = 0.4, \alpha = 0.5$	15.6637	null	1
$\rho = 0.5, \lambda = 0.4, \alpha = 0.6$	15.3464	null	1
$\vdots$	$\vdots$	null	1
$\rho = 0.5, \lambda = 0.4, \alpha = 1$	15.4793	null	1
$\rho = 0.5, \lambda = 0.5, \alpha = 0.5$	15.5351	null	1
$\rho = 0.5, \lambda = 0.5, \alpha = 0.6$	15.4225	null	1
$\vdots$	$\vdots$	null	1
$\rho = 0.5, \lambda = 0.5, \alpha = 1$	15.2355	null	1

TABLE 13. Results of Pearson and Spearman between 22 inputs and output DBO-S.

Inputs	r-value(Pearson)	p-value(Pearson)	r-value(Spearman)	p-value(Spearman)
1.Q-E	-0.1342	0.0089	-0.1220	0.0175
2.ZN-E	0.0160	0.7562	0.0304	0.5551
3.PH-E	0.0723	0.1599	0.0971	0.0591
4.DBO-E	0.3636	0.0000	0.4060	0.0000
5.DQO-EE	0.2814	0.0000	0.3552	0.0000
6.SS-E	0.0076	0.8828	0.1265	0.0137
7.SSV-E	0.1920	0.0002	0.1805	0.0004
8.SED-E	0.1135	0.0271	0.2324	0.0000
9.COND-E	0.1081	0.0354	0.1434	0.0051
10.PH-P	0.0564	0.2731	0.0764	0.1377
11.DBO-P	0.3238	0.0000	0.3464	0.0000
12.SS-P	0.0394	0.4446	0.1553	0.0024
13.SSV-P	0.1890	0.0002	0.1985	0.0001
14.SED-P	0.1645	0.0013	0.2989	0.0000
15.COND-P	0.1280	0.0126	0.1698	0.0009
16.PH-D	-0.0212	0.6805	0.0169	0.7424
17.DBO-D	0.4918	0.0000	0.5045	0.0000
18.DQO-D	0.4343	0.0000	0.4601	0.0000
19.SS-D	0.2250	0.0000	0.2768	0.0000
20.SSV-D	0.1626	0.0015	0.1639	0.0014
21.SED-D	0.1850	0.0003	0.2363	0.0000
22.COND-D	0.1428	0.0053	0.2064	0.0001

The results of the attribute reduction experiments for DBO-S are reported in TABLEs 9 to 12.

From TABLEs 9 to 12, we conclude that the FIMDMF is better than the VPDRS and FNRS in the attribute reduction.

The experiment results of Pearson and Spearman methods are shown in TABLE 13. From TABLE 13, we found that the Pearson and Spearman methods could not find out the correlated input attributes of the output DBO-S attribute, because almost all absolute r-valuse of Pearson and Spearmn methods are smaller than 0.5. The Pearson and Spearman methods are not effective in this example. This is because almost all the input attributes are nonlinear with the output attribute DBO-S.

However, we can easily plot the variation tendency of input attributes' FIMDMF values according to FIMDMF experiment results. We plot the FIMDMF values of DBO-P, DBO-D, DQO-E and DQO-D as examples in Fig.4 and Fig.5, and so do other input attributes. From the Fig.4 and Fig.5, DBO-P, DBO-D, DQO-E and DQO-D are strongly positively correlated with the DBO-S. DBO-P is positively correlated with DBO-S stronger than DBO-D, while DQO-E is stronger than DQO-D. These results are consistent with the real situation in the water treatment, because the biological demand of oxygen and chemical demand of oxygen to the secondary settler are more important than to the primary settler for DBO-S in practice. Compared with other methods, the FIMDMF can easily identify the correlation variation tendency in spite of the nonlinear relationship among different attributes. The results are consistent with the parameter discussion of the fuzzy monotone inclusive model in Section III-E.

C. COMPARISONS AGAINST UCI CONCRETE COMPRESSIVE STRENGTH DATA

In this experiment, UCI Concrete Compressive Strength data set [44] was used and the data information was described in TABLE 14. According to [45], all input attributes affect the concrete compressive strength apparently except for the fly ash. So the core input attributes to the concrete compressive

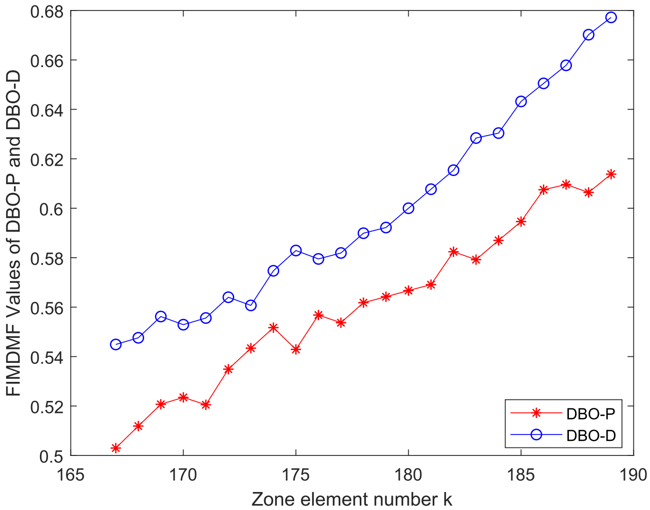

FIGURE 4. FIMDMF values of DBO-P and DBO-D correlation variation tendency for DBO-S.

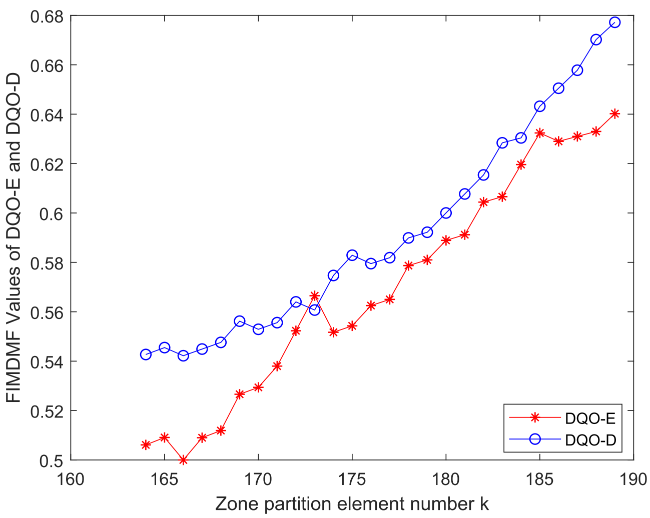

FIGURE 5. FIMDMF values of DQO-E and DQO-D correlation variation tendency for DBO-S.

strength is $core = \{1, 2, 4, 5, 6, 7, 8\}$. The FNRS is ineffective for this dataset. We report the attribute reduction results in TABLE 15 to TABLE 17 only for FIMDMF, FMF and VPDRS.

From the results in the tables, the FIMDMF performs best in terms of the error rate and better in the runtime for the attribute reduction.

For correlation analysis, the experiment results of Pearson and Spearman methods are shown in TABLE 18. From TABLE 18, we found that Pearson and Spearman methods could not find out the correlative relationship between eight inputs and the output(Concrete compressive strength), because most absolute r-values are smaller than 0.5. The Pearson and Spearman methods are not effective in this experiment because the concrete compressive strength is a highly nonlinear function of age and ingredients according to the data description [44].

TABLE 14. UCI concrete compressive strength data information.

Number	Attribute Name	Description	Data type
1	Cement	Input variable	Quantitive
2	Blast Furnace Slag	Input variable	Quantitive
3	Fly Ash	Input variable	Quantitive
4	Water	Input variable	Quantitive
5	Superplasticizer	Input variable	Quantitive
6	Coarse Aggregate	Input variable	Quantitive
7	Fine Aggregate	Input variable	Quantitive
8	Age	Input variable	Quantitive
9	Concrete compressive strength	Output variable	Quantitive

TABLE 15. FIMDMF result.

Runtime	Fuzzy monotone increasing	Fuzzy monotone decreasing	er
2.5932	1,2,5,8	4,6,7	0

TABLE 16. FMF result.

Runtime	Fuzzy monotone increasing	Fuzzy monotone decreasing	er
1.1311	1,2,5,8	3,6,7	0.286

TABLE 17. VPDRS results.

Parameter β	Runtime	Reduction set	er
0.5	19533.6474	All attributes(no reduction)	0.181
0.6	17433.5018	All attributes(no reduction)	0.181
0.7	16990.4256	All attributes(no reduction)	0.181
0.8	16682.2497	All attributes(no reduction)	0.181
0.9	16395.4207	All attributes(no reduction)	0.181
1.0	15420.7622	All attributes(no reduction)	0.181

TABLE 18. Results of Pearson and Spearman between the 8 inputs and the output.

Inputs	r-value(Pearson)	p-value(Pearson)	r-value(Spearman)	p-value(Spearman)
1.Cement	0.4978	0.0000	0.4776	0.0000
2.Blast Furnace Slag	0.1348	0.0141	0.1625	0.0000
3.Fly Ash	-0.1058	0.6753	-0.0780	0.0123
4.Water	-0.2896	0.0000	-0.3084	0.0000
5.Superplasticizer	0.3661	0.0000	0.3476	0.0000
6.Coarse Aggregate	-0.1649	0.0001	-0.1835	0.0000
7.Fine Aggregate	-0.1672	0.0001	-0.1800	0.0000
8.Age	0.3289	0.0000	0.5960	0.0000

However, we can easily plot the variation tendency of input attributes' FIMDMF values according to FIMDMF experimental results. According to TABLE 15, we plot the positive correlation or fuzzy monotone increasing input attributes with the output attribute(Concrete compressive strength) in Fig.6, and the negative correlation or fuzzy monotone decreasing input attributes in Fig.7. As x-axis values, the element number k of partition intervals is between $\lfloor n/2 \rfloor - \lfloor 0.02 \times n \rfloor$ and $\lfloor n/2 \rfloor$(495 to 515). From the Fig.6 and Fig.7, we found that the FIMDMF method can easily reveal the positive or negative correlation variation tendency between the nonlinear inputs and output, and easily find which inputs is stronger in positive or negative correlation with the output. From the two figures, the cement is the strongest positive correlation with the concrete compressive strength, while the water is the strongest negative correlation with it within the certain range, which is consistent with the actual situation. However, Pearson, Spearman and other compared methods have no such feature. The results are consistent with the parameter discussion of fuzzy monotone inclusive model in Section III-E.

D. DISCUSSIONS

According to our experimental results on the continuous input and output data, we can conclude that FIMDMF is able to achieve the small error rate and use less runtime in feature

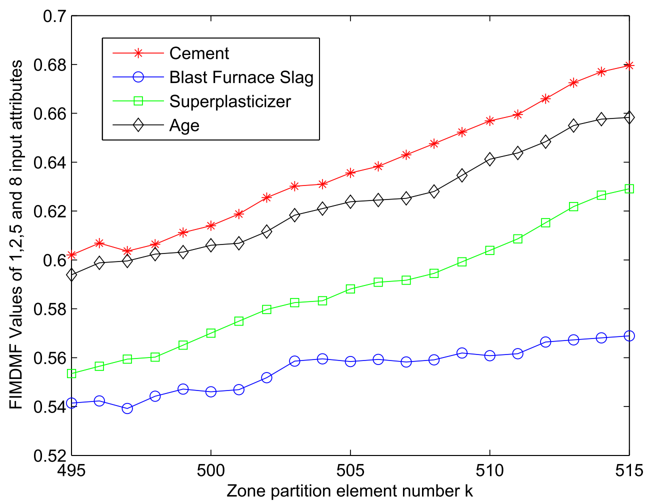

FIGURE 6. FIMDMF values of 1,2,5, and 8 inputs positive correlation variation tendency for the output.

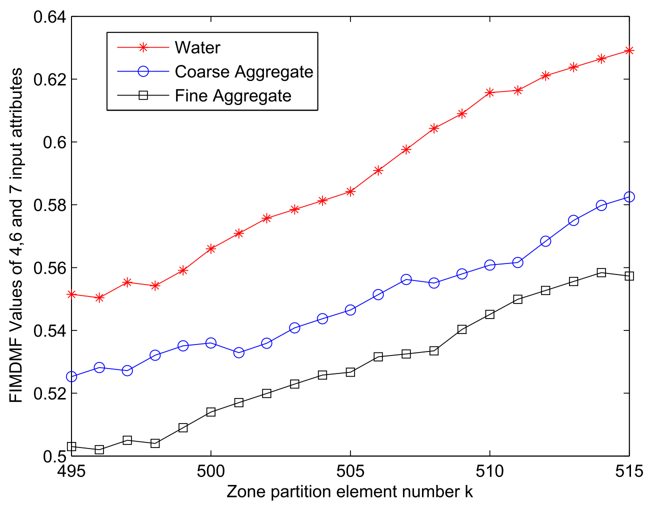

FIGURE 7. FIMDMF values of 4,6, and 7 inputs negative correlation variation tendency for the output.

selection, compared with other three methods for the continuous data [24]–[32]. In addition, the output of selected attributes by FIMDMF is easily explained in terms of which reduction attributes are more important in the set of core reduction attributes.

Based on the element partial order relationship, a fuzzy monotone relationship has a little similarity to the dominance relationship of rough sets. In fact, their differences are as follows: (1) The fuzzy monotone relationship is based on the fuzzy and monotone, rather than the rough sets theory. In contrast, the dominance relationship is based on the rough sets theory; (2) The fuzzy monotone relationship is used to search for a multielement relationship among different intervals for one attribute, while the dominance relationship is for a binary relationship among some attributes; (3) For feature selection or attribute reduction, the fuzzy monotone relationship methods can be used to determine whether the quantity variation of an input condition attribute is similar to that of the certain output decision attribute or not, by considering input condition attributes one by one. However, the dominance relationship methods determine whether some attributes' classification approximates the certain decision attributes' classification for objects or not; and (4) The fuzzy monotone method achieves feature selection or attribute reduction without using heuristic methods. But the dominance relationship methods always use heuristic methods, because it is a NP-hard problem to find optimized feature selection or attribute reduction set in the rough sets theory [23], [33].

Compared with both other monotone test methods, one is presented in [32] named FMF above, the other is presented in [21]. This paper method FIMDMF is a little similar as the method FMF, and FMF method is to select a minimum value in one interval to compare with the values of neighbor partition interval to gain the value of fuzzy monotone test, however, FIMDMF method is to measure the elements' partial order to gain the value of fuzzy monotone test between the two neighbor intervals at the same time, and the partition interval methods of FMF and FIMDMF are different too. The FIMDMF method is very different from the method presented in [21]. Some different features are discussed as follows: (1) FIMDMF method measures the elements' partial order relationship between the two partition neighbor interval, then gain values of fuzzy monotone test through all two partition neighbor intervals for many times interval partition, but the method in [21] is to measure total neighbor elements' relationship to gain a value of monotone test; (2) FIMDMF method is presented according to the third property of Definition 1; (3) FIMDMF method can be used to analyze the nonlinear relationship through the variation tendency of different partition values of fuzzy membership function, however, the method in [21] can attain only one value for monotone test and cannot be used to analyze the nonlinear relationship. And the method in [21] cannot be used to describe the fuzzy monotone relationship.

Compared with feature selection or attribute reduction methods for classification, the FIMDMF method is based on the fuzzy monotone relationship. The purpose of the other methods for classification is to find out a set of the reduced condition attributes with similar classification discrimination capabilities with respect to the decision attribute. But they cannot be used for finding out the quantity variation of the reduced input condition attributes correlated with the output decision attribute. In contrast, our method is able to achieve this. So some characteristics of FIMDMF method are summarized as follows: (1) Based on the novel fuzzy monotone relationship, it is not for the classification; (2) It can directly be applied to numeric or continuous attributes without any discretization; and (3) It does not aim to find the reduced number of condition attributes with the similar capacity of classification discrimination with respect to a decision attribute. Rather, it finds the reduced number of condition attributes that mostly affect the continuous quantity variation of a decision attribute in a set of continuous data.

Compared with the correlation analysis methods Pearson and Spearman, the fuzzy monotone relationship method can be directly applied to deal with nonlinear relationships, and the FIMDMF method can easily reveal the correlation variation tendency among the different attributes and easily do the analysis in spite of the linear or nonlinear relationship.

VI. CONCLUSION

In this paper, we have introduced new concepts of the fuzzy monotone and the fuzzy inclusive monotone, and defined a fuzzy monotone relationship based on an inclusion degree. For deeply examining the fuzzy monotone relationship between input and output attributes, we have presented and proved several propositions. A decision membership function is then deduced from these propositions. According to the proposed decision membership function, a new algorithm has been presented. Against several sets of continuous or numeric data in both input and output ends, the experiment results and comparisons indicated that the proposed method is effective for feature selection or attribute reduction. The fuzzy monotone relationship can be directly applied to nonlinear relationships for correlation analysis. It is suitable for revealing the correlation variation tendency in the linear or nonlinear relationship among different attributes.Then this paper presents a new effective way to correlation analysis, especially for nonlinear correlation analysis. It also presents a new effective way to feature selection or attribute reduction for continuous data set in both input and output ends. Our future work will further research and apply the fuzzy monotone relationship together with its methods to more and more fields.

REFERENCES

[1] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–353, Jun. 1965.

[2] A. A. Ewees, M. A. Elaziz, R. M. Arafa, and R. M. Ghoniem, "Improved approach based on fuzzy rough set and sine-cosine algorithm: A case study on prediction of osteoporosis," *IEEE Access*, vol. 8, pp. 203190–203202, 2020.

[3] C. Wang, Y. Huang, M. Shao, and X. Fan, "Fuzzy rough set-based attribute reduction using distance measures," *Knowl.-Based Syst.*, vol. 164, pp. 205–212, Jan. 2019.

[4] P. Chen, M. Lin, and J. Liu, "Multi-label attribute reduction based on variable precision fuzzy neighborhood rough set," *IEEE Access*, vol. 8, pp. 133565–133576, 2020.

[5] L. Wang, Q. Ma, and J. Meng, "Incremental fuzzy association rule mining for classification and regression," *IEEE Access*, vol. 7, pp. 12110–12195, 2019.

[6] C. Wang, Q. He, D. Chen, and Q. Hu, "A novel method for attribute reduction of covering decision systems," *Inf. Sci.*, vol. 254, pp. 181–196, Jan. 2014.

[7] E. T. Lee and T. Chou, "Fuzzy monotone functions and applications," *Kybernetes*, vol. 30, no. 1, pp. 84–97, Feb. 2001.

[8] L. M. Pang, K. M. Tay, C. P. Lim, and H. Ishibuchi, "A new monotone fuzzy rule relabeling framework with application to failure mode and effect analysis methodology," *IEEE Access*, vol. 8, pp. 144908–144930, 2020.

[9] Y. W. Kerk, K. M. Tay, and C. P. Lim, "Monotone interval fuzzy inference systems," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 11, pp. 2255–2264, Nov. 2019.

[10] Y. W. Kerk, K. M. Tay, and C. P. Lim, "Monotone fuzzy rule interpolation for practical modelling of the zero-order TSK fuzzy inference system," *IEEE Trans. Fuzzy Syst.*, early access, Feb. 4, 2021, doi: 10.1109/TFUZZ.2021.3057239.

[11] E. Van Broekhoven and B. De Baets, "Only smooth rule bases can generate monotone Mamdani–Assilian models under center-of-gravity defuzzification," *IEEE Trans. Fuzzy Syst.*, vol. 17, no. 5, pp. 1157–1174, Oct. 2009.

[12] P. Husek, "Monotonic smooth Takagi–Sugeno fuzzy systems with fuzzy sets with compact support," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 3, pp. 605–611, Mar. 2019.

[13] L. M. Pang, K. M. Tay, and C. P. Lim, "Monotone fuzzy rule relabeling for the zero-order TSK fuzzy inference system," *IEEE Trans. Fuzzy Syst.*, vol. 24, no. 6, pp. 1455–1463, Dec. 2016.

[14] H. Seki, H. Ishii, and M. Mizumoto, "On the monotonicity of fuzzy-inference methods related to T–S inference method," *IEEE Trans. Fuzzy Syst.*, vol. 18, no. 3, pp. 629–634, Jun. 2010.

[15] P. Husek, "On monotonicity of Takagi–Sugeno fuzzy systems with ellipsoidal regions," *IEEE Trans. Fuzzy Syst.*, vol. 24, no. 6, pp. 1673–1678, Dec. 2016.

[16] J.-M. Won and F. Karray, "Toward necessity of parametric conditions for monotonic fuzzy systems," *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 2, pp. 465–468, Apr. 2014.

[17] B. De Baets, "Implication-based models of monotone fuzzy rule bases," *Fuzzy Sets Syst.*, vol. 232, pp. 134–155, Dec. 2013.

[18] S. Mandal and B. Jayaram, "Monotonicity of SISO fuzzy relational inference with an implicative rule base," *IEEE Trans. Fuzzy Syst.*, vol. 24, no. 6, pp. 1475–1487, Dec. 2016.

[19] C. Li, J. Yi, and G. Zhang, "On the monotonicity of interval type-2 fuzzy logic systems," *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 5, pp. 1197–1212, Oct. 2014.

[20] Y. W. Kerk, C. Y. Teh, K. M. Tay, and C. P. Lim, "Parametric conditions for a monotone TSK fuzzy inference system to be an N-ary aggregation function," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 7, pp. 1864–1873, Jul. 2021.

[21] C. Y. Teh, Y. W. Kerk, K. M. Tay, and C. P. Lim, "On modeling of data-driven monotone zero-order TSK fuzzy inference systems using a system identification framework," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 6, pp. 3860–3874, Dec. 2018.

[22] Z. Wang, X. Zheng, and H. G. Pan, "Information entropy multi-decision attribute reduction fuzzy rough set for dust particulate imagery characteristic extraction," *IEEE Access*, vol. 8, pp. 77865–77874, 2020.

[23] C. Wang, M. Shao, B. Sun, and Q. Hu, "An improved attribute reduction scheme with covering based rough sets," *Appl. Soft Comput.*, vol. 26, pp. 235–243, Jan. 2015.

[24] C. Wang, M. Shao, Q. He, Y. Qian, and Y. Qi, "Feature subset selection based on fuzzy neighborhood rough sets," *Knowl.-Based Syst.*, vol. 111, no. 1, pp. 173–179, 2016.

[25] S. Greco, B. Matarazzo, and R. Slowinski, "Rough approximation by dominance relations," *Int. J. Intell. Syst.*, vol. 17, no. 2, pp. 153–171, Feb. 2002.

[26] M. Inuiguchi, Y. Yoshioka, and Y. Kusunoki, "Variable-precision dominance-based rough set approach and attribute reduction," *Int. J. Approx. Reasoning*, vol. 50, no. 8, pp. 1199–1214, Sep. 2009.

[27] H. Y. Zhang and S. Y. Yang, "Feature selection and approximate reasoning of large-scale set-valued decision tables based on α-dominance-based quantitative rough sets," *Inf. Sci.*, vol. 378, pp. 328–347, Jan. 2017.

[28] W. H. Xu, *Ordered Information System Rough Sets*. Beijing, China: Science Press, 2013.

[29] L.-H. Chen and C.-C. Tu, "Dominance-based ranking functions for interval-valued intuitionistic fuzzy sets," *IEEE Trans. Cybern.*, vol. 44, no. 8, pp. 1269–1282, Aug. 2014.

[30] B. Huang, H.-X. Li, and D.-K. Wei, "Dominance-based rough set model in intuitionistic fuzzy information systems," *Knowl.-Based Syst.*, vol. 28, pp. 115–123, Apr. 2012.

[31] X. Yang, Y. Qi, D.-J. Yu, H. Yu, and J. Yang, "α-dominance relation and rough sets in interval-valued information systems," *Inf. Sci.*, vol. 294, pp. 334–347, Feb. 2015.

[32] J. Liang, F. Luo, and Y. G. Xu, "Fuzzy rough monotone dependence algorithm based on decision table and its application," *J. South China Univ. Technol.*, vol. 39, no. 7, pp. 7–12, 2011.

[33] D. Ye, Z. Chen, and S. Ma, "A novel and better fitness evaluation for rough set based minimum attribute reduction problem," *Inf. Sci.*, vol. 222, pp. 413–423, Feb. 2013.

[34] Y. Qian, J. Liang, and C. Dang, "Consistency measure, inclusion degree and fuzzy measure in decision tables," *Fuzzy Sets Syst.*, vol. 159, no. 18, pp. 2353–2377, Sep. 2008.

[35] Z. B. Xu, J. Y. Liang, C. Y. Dang, and K. S. Chin, "Inclusion degree: A perspective on measures for rough set data analysis," *Inf. Sci.*, vol. 141, nos. 3–4, pp. 227–236, Apr. 2002.

[36] G. J. Klir, *An Overview of Fundamentals, Applications and Personal Views*. Beijin, China: Beijing Normal Univ. Press, 2000.

[37] J. Liang, "Fuzzy rough monotone data mining algorithm with it's application research in wastewater treatment," Ph.D. dissertation, School Autom. Sci. Eng., South China Univ. Technol., Guangdong, China, 2011.

[38] J. D. Gibbons and S. Chakraborti, *Nonparametric Statistical Inference*, vol. 149, no. 3. Boca Raton, FL, USA: CRC Press, 2014.

[39] J. Hauke and T. Kossowski, "Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data," *Quaestiones Geographicae*, vol. 30, no. 2, pp. 87–93, Jun. 2011.

[40] T. D. Gauthier, "Detecting trends using Spearman's rank correlation coefficient," *Environ. Forensics*, vol. 2, no. 4, pp. 359–362, Dec. 2001.

[41] D. Dua and C. Graff, *UCI Machine Learning Repository*. Irvine, CA, USA: Univ. California, 2017. [Online]. Available: http://archive.ics.uci.edu/ml/datasets/Water+Treatment+Plant

[42] I. Santín, C. Pedret, and R. Vilanova, "Applying variable dissolved oxygen set point in a two level hierarchical control structure to a wastewater treatment process," *J. Process. Control*, vol. 28, pp. 40–55, Apr. 2015.

[43] T. Yang, W. Qiu, Y. Ma, M. Chadli, and L. Zhang, "Fuzzy model-based predictive control of dissolved oxygen in activated sludge processes," *Neurocomputing*, vol. 136, pp. 88–95, Jul. 2014.

[44] Y. I-Cheng, *UCI Machine Learning Repository*. Irvine, CA, USA: Univ. California, 2007. [Online]. Available: http://archive.ics.uci.edu/ml/datasets/Concrete+Compressive+Strength

[45] I.-C. Yeh, "Design of high-performance concrete mixture using neural networks and nonlinear programming," *J. Comput. Civil Eng.*, vol. 13, no. 1, pp. 36–42, Jan. 1999.

JIN LIANG received the M.S. degree in computer software and theory from South China Normal University, Guangzhou, China, in 2004, and the Ph.D. degree in control theory and control engineering from the South China University of Technology, Guangzhou, in 2011.

From 2004 to 2009, he was a Lecturer and a Researcher in education with the School of Information Technology, South China Normal University. From 2010 to 2021, he was an Assistant Professor of education with the School of Information Technology, South China Normal University. He is currently an Assistant Professor with the School of Software, South China Normal University. His research interests include fuzzy sets, fuzzy logic, fuzzy systems, data mining, and machine learning.

ZUO PENG ZHANG received the bachelor's degree in management from Shenzhen University, and the master's degree in software engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2011. From 2013 to 2021, he has been working as an Experimenter with the Experimental Center, School of Software, South China Normal University, Foshan, China. His research interests include natural language processing and intelligent computing, basic research of natural language processing in education, and AI technology + education.