hasil cek-manuskrip sugiyarto

by Sugiyarto Cek-sugiyarto

Submission date: 08-Sep-2021 08:36AM (UTC+0700)
Submission ID: 1643410824
File name: Manuscript_Sugiyarto1_FINAL_1.docx (96.15K)
Word count: 4578
Character count: 23858

Dimension Reduction of Multivariate Fuzzy Time Series using Core and Reduct Method

Sugiyarto Surono[1*] , Rani Nuraeni[2], Joko Eliyanto[3], Salafudin[4]
sugiyarto@math.uad.ac.id, rani1700015051@webmail.uad.ac.id, joko1907050003@webmail.uad.ac.id, salafudin@iainpekalongan.ac.id
[1,2,3] Department Mathematics, FAST UAD Yogyakarta
[4] Department Of Math. Education, IAIN Pekalongan Jawa Tengah

Abstract

Time series is widely known as one of the easily available and accessible data. However, its forecast usually requires more than one factor and its movement also tends to fluctuate. These characteristics are normally used in assuming its fuzziness level. This data can also be defined as a multivariate time series such that the high number of variables used in the forecasting process leads to a high level of computational weight. Therefore, this research was used to develop a forecasting method for multivariate time series data with significantly lower computational weight. The multivariate fuzzy time series was combined with the dimension reduction method to reduce the dimension of the data while maintaining the important information. Moreover, the reduction dimension process was first applied to obtain data with a lower number of dimensions before the forecasting process was initiated. The core and reduct method which is one of the concepts in Rough Theory was applied in the dimension reduction process. The results showed that the core and reduct method has the ability to decrease the computational time up to 80.9% without significantly reducing the accuracy level which was maintained at 99%. It was concluded that the proposed combination of methods was able to predict with a high level of accuracy and lower computational weight than the standard method.

Keywords: Data time series, multivariate fuzzy time series, rough set theory, core and reduct.

1. Introduction

Data time series is a series of observations conducted sequentially over a certain time interval. It is usually applied as a supporting factor during the decision-making process for forecasts [1]. This is due to its ability to provide detailed values in the form of numbers from the previous time interval. Moreover, it is also believed to have a repeating pattern and this means it has the ability to repeat a past occurrence in the present or future. Data time series is also used for predictions based on the assumption that the future function is dependent on the past function. This means it is possible to observe and predict the possible occurrences at a certain timeline through the application of past data [2] [3] [4].

Forecasting is a method of predicting certain future observation values by considering and observing the data acquired in the past or present. It is usually used to determine the working pattern of selected data. The method emerged as one of the important tools to produce an effective and efficient plan [5]. Meanwhile, time series forecasting is defined as a process of predicting certain future occurrences by considering the historical pattern of the data analyzed in the form of data series. One of the methods normally used in predicting this time series data is Multivariate Fuzzy Time Series (MFTS) [6]. It is a data forecasting method which involves the use of fuzzy theorem as the basis [7] [8] [9]. This system works by capturing the pattern of a particular data to project its future pattern using more than one variable [10] [11] [12].

The problem usually observed in the application of MFTS is the high level of computational weight and an increase in complexity during the forecasting process. This is associated with the relatively high number of variables applied [13][14]. Moreover, the unavailability of the exact pattern rule also affects the termination of certain variables with a possible significant impact on the accuracy of the forecasting process. There is, therefore, the need to find a method which has the ability to reduce the computational weight without significantly altering the level of accuracy. Furthermore, the objective of the reduction dimension method is to transform the high dimension data into the lower dimension data but while maintaining the main characteristics of the data [15]

* Corresponding author. Sugiyarto, Departement of Mathematics, Ahmad Dahlan University, Yogyakarta, Indonesia. E-mail: sugiyarto@math.uad.ac.id.

[16]. The method is expected to produce new information in the form of a pattern of rules to be applied as the reference in the forecasting process in order to improve the performance of the computational weight. It is also important to note that core and reduct is one of the techniques in rough set theory directed towards reducing the dimension of a certain dataset to ensure only informational core remain [17] [18] [19] [20].

Few studies were discovered to be related to forecasting and dimension reduction. For example, Hendrik Fery Herdiyatmoko [21] applied RST to reduce fire accident parameters or attributes in order to minimize complexity in data analysis. The results were later applied as a reference to determine the evacuation route for fire accidents.

Tahseen A. Jilani et al. [10] also studied the use of MFTS in predicting the number of cars accidents in Belgium using one main variable which is the death toll due to the accident and other four supporting variables. The pattern for the cause of the accident was analyzed to determine the fuzzy premises and assurances to measure the car and life insurance.

I Made Candra Satria [12] also compared the forecasting results from the Fuzzy Time Series (FTS) and MFTS using three variables which include the number of Australian tourists, Inflation in Indonesia, and AUD/IDR exchange rate. It was discovered that the vacation number of the Australian tourists in Bali producing using MFTS was more accurate than FTS.

This current research was designed to focus on the application of MFTS in the forecasting process using the basic concept of FTS developed to solve linguistic variables problems. It is important to note that MFTS is one of the improvements made on FTS [22] [23] and involves the consideration of more than one variable or factor. One of these is used as the main while others are supporting variables. This approach also combines linguistic variables with the fuzzy logic analytic process in the time series data to avoid using unwanted data. This method is, however, unable to determine the appropriate variables to be applied in the forecasting process despite its ability to process more than one variable [24] [9]. Moreover, the use of unwanted variables in forecasting also increases the time required for the computational system to operate, and Rough Set Theory has been reported to be one of the methods of dimension reduction proven to have the ability to reduce this computational weight. The main problem solved in this research is how to reduce the computational weight in the forecasting process. This involved evaluating the effectiveness of combining the MFTS and Rough Set Theory method for the forecasting process with a high number of variables. Therefore, the main objective was to produce a forecasting process with a high level of accuracy and lower computational weight than the standard method. The proposed method was applied to the sample data observed to be suitable with the desired condition. It is also important to note that only numerical data was used for the development just as in the MFTS method.

2. Method

2.1 Multivariate Fuzzy Time Series (MFTS)

Multivariate Fuzzy Time Series (MFTS) was developed based on Fuzzy Time Series (FTS). The concept was invented by Song and Chissom to determine the forecasting problem associated with historical data in the form of linguistic value. The basic principle to measure the MFTS method is if U is the universe with $U = \{u_1, u_2, \dots, u_n\}$, then the fuzzy set of $A_i = (i = 1,2, \dots, n)$ can be defined as:

$$A_i = \frac{\mu_{A_i}(u_1)}{u_1} + \frac{\mu_{A_i}(u_2)}{u_2} + \cdots + \frac{\mu_{A_i}(u_n)}{u_n} \qquad (1)$$

Where μ_{A_i} is the membership function of the fuzzy set A_i and $\mu_{A_i}(u_1)$ is the membership degree of μ_k from u_k in the fuzzy set A_i, $k = 1,2, \dots, n$.

Definition 2.1 [25] A universe of $Y(t)$ with $(t = \cdots, 0,1,2, \dots, n, \dots)$ is given as the *subset* from the real number (R) used to define the fuzzy set of $A_i(t)$. If $F(t)$ is the set from $A_1(t), A_2(t), \dots, A_n(t)$, then $F(t)$ is called *Fuzzy Time Series* (FTS) on $Y(t)$ with $(t = \cdots, 0,1,2, \dots, n, \dots)$.

Song and Chissom used the fuzzy relation equation to develop the forecasting model based on the assumption that the observation on time t is only dependent on the observation result of the previous time which is defined as follows.

Definition 2.2 [25] Suppose $F(t)$ is only caused by $F(t-1)$ and symbolized as $F(t-1) \rightarrow F(t)$, then the fuzzy relation between $F(t)$and $F(t-1)$ can be illustrated as the fuzzy relation equation $F(t) = F(t-1)^{\circ}R(t-1)$. The "°" in this equation is the operator of the max-min composition while the R relation is described as the first-order model from $F(t)$.

Definition 2.3 [26] Suppose $F(t-1) = A_i$ and $F(t) = A_j$, then it is possible to define the *Fuzzy Logical Relationship* (FLR) as $A_i \rightarrow A_j$, where A_i is described as the Current State and A_j is the Next. Moreover, if there is an FLR acquired from the *state* A_2, then there is a transition to the other state of A_j, $j = 1,2, \dots, n$, in the form of:

$$A_2 \to A_1, A_2 \to A_4, A_2 \to A_2$$

Then, FLR can be classified as the Fuzzy Logical Relationship Group (FLRG) as follows:

$$A_2 \to A_1,\ A_4, A_2$$

Definition 2.4 [26] If $F(t)$ is caused by more of the fuzzy set $F(t-n), F(t-n+1), \dots, F(t-1)$, then the fuzzy relationship will be represented by $A_{i1}, A_{i2}, \dots, A_{in} \to A_j$, where $F(t-n) = A_{i1}, (t-n+1) = A_{i2}, \dots, F(t-1) = A_{in}$. This relationship is, therefore, known as the n-order Fuzzy Time Series model or Multivariate Fuzzy Time Series.

The basic concept of Fuzzy was developed by L.Zadeh, improved by Song and Chissom, and further developed by Cheng et al. with slight changes applied in the interval and weight determination using Fuzzy Logical Relationship (FLR). Moreover, the weight was measured by inserting all the relationships and measurements based on the same FLR repetition. This means the forecasting process does not necessarily need a certain complex learning system, thereby, leading to the relatively easy application and development.

2.2 Rough Set Theory (RST)

Rough set theory was popularized by Zdiszlaw Pawlak in 1982 [3]. It is a methodology design to classify and analyze imprecise, uncertain, or uncompleted information and knowledge. It is, however, possible to apply the Core and Reduct method for the reduction process in this theory.

Definition 2.5 [27] A certain set of data is presented as a table in the rough set with the rows use for the cases, occurrence, patients, or certain objects while the columns indicate the variable attributes, observation, properties, etc. This table is generally described as the Information Systems, notated by IS, and defined as follows:

$$IS = (U, P) \tag{2}$$

Where U is the unlimited set not emptied from the objects and defined as Universe while A is the limited set not emptied from the attributes, where: $a: U \to V_a$ for each $a \in A$ set of V_a is described as the value set from a.

Definition 2.6 [27] Suppose $IS = (U, P)$ is the *information system* and $B \subseteq P$, then, the *indiscernibility* of the objects based on the attribute of B which is symbolized as $IND_{IS}(B)$ can be defined as:

$$IND_{IS}(B) = \{(x, x') \in U^2 | \forall_a \in B\ a(x) = a(x')\} \tag{3}$$

$IND_{IS}(B)$ is the *B-indiscernibility relation* and $IND_{IS}(B)$ is the equivalence relation. Therefore, if $(x, x') \in IND_{IS}(B)$, then, the objects of x and x' are indiscernible towards each other by attributes of B. The class considered to be equivalent with *B-indiscernibility relation* is denoted by $[x]_B$ and tagged the "equivalent class".

Definition 2.7 [27] Suppose $IS = (U, P)$ is the *information system* $B \subseteq P$ and $a \in B$, then, a is described as the dispensable variable in the attributes of B if:

$$IND_{IS}(B) = IND_{IS}(B) - \{a\} \tag{4}$$

Meanwhile, if a is considered indispensable it means it is very important in the attributes of B.

The set of B is considered independent when all its attributes are required. Meanwhile, each subset B' from B described as the reduct from B is B' is independent and this means $IND_{IS}(B') = IND_{IS}(B)$. Therefore, it is possible to define a reduct as the set of attributes which has the ability to show that the forecasting process is as accurate as the result produced when all the attributes are applied. The attributes not considered as reduct are those to be erased and without any effect on the forecasting process.

Definition 2.8 [27] Suppose $B \subseteq P$ and *Core* from B are the set of all the indispensable attributes of B, then, it is possible to define the core as:

$$Core(B) = \cap\, Red(B) \tag{5}$$

Where $Red(B)$ is the set of all reduct results on B. The core is, therefore, included in each of the reductions due to the fact that it is the intersection from all of the reduct results. This means each of the core attributes is included in the few reduction processes. Therefore, it is possible to consider core as the most important part of the attributes due to the absence of indispensable attributes in the analyzed set.

2.3 Multivariate Fuzzy Time Series Algorithm

Fuzzy Logical Relation (FLR) is one of the important factors in the Multivariate Fuzzy Time Series theory affecting the accuracy of the forecasting process [9]. It was provided as a slightly different method by Cheng to determine the interval and measure weight. The process involves inserting all the relationships and measurements based on the same FLR repetition. Moreover, the steps of the time series data forecasting process using the Fuzzy time series weighted based on the Cheng method are as follows:

First Step: Determine the universe

The universe for the main variable U can be defined as follows:

$$U = [D_{min}, D_{max}] \tag{6}$$

Where Dmin and Dmax is the minimum and maximum value from the historical data respectively. They are used to divide the set of the universe into few intervals of $U = \{u_1, u_2, \dots, u_n\}$ at an equal distance.

Second Step: Determine the range of the interval

The interval range was measured by applying the frequency distribution as indicated in the following stages:

a) Determine the range through the following equation:

$$R = D_{max} - D_{min} \tag{7}$$

Where R is the range while Dmin and Dmax is the minimum and maximum value as previously indicated.

b) Determine the number of interval classes using the following Sturges equation:

$$K = 1 + 3.322 \times \log(n) \tag{8}$$

Where, n is the number of data.

c) Measure the interval range using the following equation:

$$l = \frac{range\ data\ (R)}{Number\ of\ class\ interval(K)} \tag{9}$$

The interval was divided using the suitable number and length to ensure each can be measured using the following relationships:

$$\begin{aligned} u_1 &= [D_{min}, D_{min} + l] \\ u_2 &= [D_{min} + l, D_{min} + 2l] \\ &\vdots \\ u_n &= [D_{min} + (n-1)l, D_{min} + nl] \end{aligned} \tag{10}$$

e. Determine the median value with m_i by applying the following equation.

f.

$$m_i = \frac{lower\ limit + upper\ limit}{2} \tag{11}$$

Where, i is the number of fuzzy sets.

Third step: Determine the fuzzy set.

The fuzzy set was formed by observing the number of different frequencies. Those with higher values were divided into equal values of h interval while the second most high number were divided into equal value of h-1 interval. This measurement was repeated up to the period it is impossible to divide the interval further.

Fourth Step: Define the fuzzy sell for all the universes in each variable.

Each of the fuzzy sets in each variable was defined based on the already determined number of intervals as follows:

The fuzzy set for the main variable of A_i was acquired through:

$$A_i = \sum_{j=1}^{n} \frac{\mu_{ij}}{u_i} \tag{12}$$

Where, μ_{ij} is the membership degree measured using the following equation:

$$\mu_{ij} = \begin{cases} 1 & , i = j \\ 0.5 & , j = i - 1\ dan\ j = i + 1 \\ 0 & , for\ the\ others \end{cases}$$

It is possible to explain the Equation of A_i through the following set of rules:

a. Rule 1: If the historical data of Y_j is u_i, then, the membership degree of u_i is 1, u_{i+1} is 0.5, and the other is 0.

b. Rule 2: If the historical data of Y_j is u_i with $1 < i < n$, then, the membership degree of u_i is 1, u_{i+1} and u_{i-1} are 0.5, and the other is 0.

c. Rule 3: If the historical data of Y_j is u_n, then, the membership degree of u_n is 1, u_{i-1} is 0.5 and the other is 0.

Therefore, the fuzzy set from the main variable A_i can be defined as:

$$\begin{aligned} A_1 &= \frac{1}{u_1} + \frac{0.5}{u_2} + \cdots + \frac{0}{u_n} \\ A_2 &= \frac{0.5}{u_1} + \frac{1}{u_2} + \cdots + \frac{0}{u_n} \\ &\vdots \\ A_n &= \frac{0}{u_1} + \cdots + \frac{0.5}{u_{n-1}} + \frac{1}{u_n} \end{aligned} \quad (13)$$

Where, u_i $(i = 1,2, \ldots, n)$ is the element of the universe (U) while the symbol "/" represents the membership degree of $\mu_{ij}(u_i)$ towards A_i and the values are 0, 0.5, and 1.

Fifth Step: Fuzziness process of the historical data.

The fuzziness process involves transforming the actual data into the fuzzy linguistic value which is usually conducted to determine the suitable fuzzy set for each data.

Sixth Step: Determine the Fuzzy Logical Relationship (FLR).

The fuzzy logical relationship (FLR) is defined as $A_i \to A_j$. A_i is described as the current state $Y(t-1)$ and A_j is the next state in time t. The acquired FLR is usually illustrated as follows:

A 1 → A 2, A 2 → A 5, A 3 → A 2, A 1 → A 4, A 1 → A 5.

Seventh Step: Transform the relationship weight of FLR into Fuzzy Logical Relationship Group (FLGR).

This involved inserting all the relationship weight and determining the weight based on the same order and iteration. The FLR with the same current state (A_i) were classified into one group through the weighting process. Therefore, the equal order of FLR was illustrated as:

$(t = 1)$	$A_1 \to A_1$	Given weight 1
$(t = 2)$	$A_1 \to A_2$	Given weight 2
$(t = 3)$	$A_1 \to A_1$	Given weight 3
$(t = 4)$	$A_1 \to A_1$	Given weight 4

Where, t is the time. The weight acquired from the FLR relationship inserted into the weighting matrix (W) is defined as follows:

$$\boldsymbol{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix} \quad (14)$$

Where, **W** is the weighting matrix and w_{ij} is the weight in row i and column j with $i = 1,2,\ldots,n$ and $j = 1,2,\ldots,n$.

Eighth Step: Transform the weighting matrix ($\boldsymbol{W}$) into the standardized weighting matrix ($\boldsymbol{W}^*$).

The equation of the standardized weighing matrix ($\boldsymbol{W}^*$) is stated as follows:

$$\boldsymbol{W}^* = \begin{bmatrix} w_{11}^* & w_{12}^* & \cdots & w_{1n}^* \\ w_{21}^* & w_{22}^* & \cdots & w_{2n}^* \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1}^* & w_{n2}^* & \cdots & w_{nn}^* \end{bmatrix} \quad (15)$$

Where $\boldsymbol{W}^*$ is the standardized weighing matrix with $w_{ij}^* = \frac{w_{ij}}{\sum_{j=1}^{n} w_{ij}}$.

Ninth Step: Determine the Defuzzification of the predicted value.

The standardized weighing matrix ($\boldsymbol{W}^*$) was multiplied with the median value (m_i) from the main factor to produce the forecasting value. Therefore, the forecasting measurement was defined as:

$$F(t) = w_{i1}^*(m_1) + w_{i2}^*(m_2) + \cdots + w_{1p}^*(m_p) \quad (16)$$

Where $F(t)$ is the forecasting result with $w_{ij}^* = \frac{w_{ij}}{\sum_{j=1}^{n} w_{ij}}$.

This means if FLRG from A_i is an empty set ($A_i \to \emptyset$), then, the forecast from $F(t)$ is m_i which is the median value of the interval u_i.

3 Results and Discussion

3.1 Data Analysis

The climate data retrieved from the official site of BMKG (Meteorology, Climatology, and Geophysics Institute) on daily weather recorded by the Climatology station in Sleman, Yogyakarta were used as the variables in this research. The data were collected between 24 April 2018 and 26 April 2021 with the focus on 10 variables which include the

Minimum temperature (^{0}C), Maximum temperature (^{0}C), Average humidity (%), Rainfall (mm), Duration of sunshine (hour), Wind maximum velocity (m/s), Wind direction during its maximum velocity (^{0}), Average wind velocity (m/s), and the wind direction (^{0}). Moreover, Exploratory Data Analysis (EDA) was applied to determine the condition of the dataset used in order to improve the accuracy of the following forecasting process by avoiding the loss of important information from the dataset. It also has the ability to save the time required for the forecasting process. Summarily, EDA was used to handle the possibility of missing information, identify the most important variables, and also test the hypothesis and outliers.

Table. 1 Initial Climate Dataset in D.I Yogyakarta

No	Tn	Tx	RH_avg	···	ddd_x	ff_avg	Suhu
1	23	32.6	81	···	270	2	27.5
2	21	32.6	74	···	170	2	26.5
3	22	31.5	77	···	270	2	26.3
⋮	⋮	⋮	⋮	···	⋮	⋮	⋮
1094	22.7	33.1	78	···	240	2	27.3
1095	23.2	31.2	74	···	220	2	27
1096	22	30.8	76	···	230	2	26.8

3.1.1 Rough Set Theory

A dimension reduction process was required through the Rough Set Theory (RST) before the initiation of the forecasting process. This involved using the dataset used as the input and output in dimension reduction as the system information. Moreover, the core was extracted through the reduction process of this system information and the results are presented as follows:

Table. 2 Dimension Reduction Result

No	Tn	Tx	RH_avg	RR	ss
1	23	32.6	81	0	7.5
2	21	32.6	74	0	10.4
3	22	31.5	77	0	10.5
⋮	⋮	⋮	⋮	⋮	⋮
435	22.7	33.1	78	15	6.6
436	23.2	31.21864734	74	0	9.1
437	22	30.8	76		9.8

3.1.2 Multivariate Fuzzy Time Series

- Forecasting result

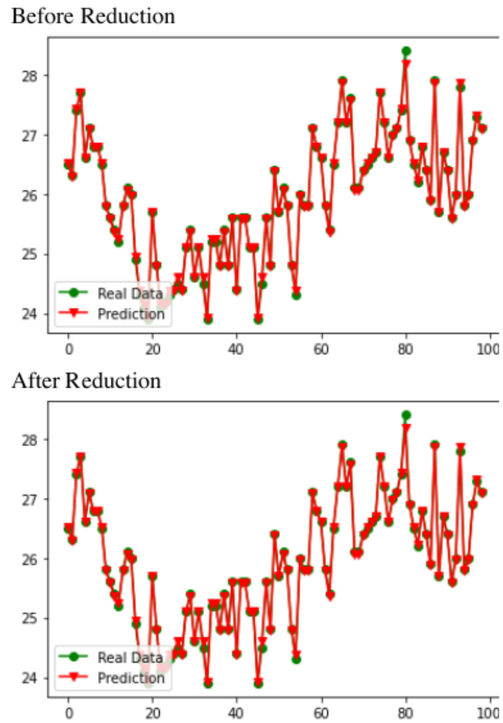

Figure 1. Forecasting result

Figure 1 shows that the Multivariate fuzzy time series provides similar results with and without reduction. This means the dimension reduction applied was able to maintain the accuracy level of the forecasting process. This, consequently, indicates the effectiveness of the MFTD.

- Accuracy

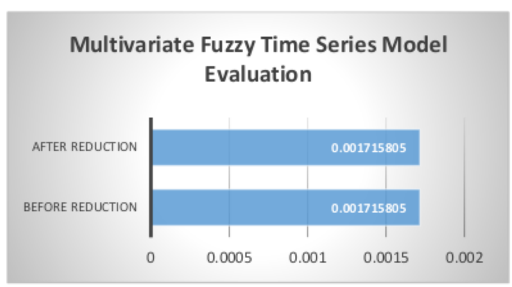

Figure 2. Model Evaluation

The improved method shows the core and reduct concept in the RST was able to maintain the accuracy level of the forecasting process. Figure 2 shows the dataset process produced a similar level of accuracy of 99% with or without reduction. This means the RST has the ability to maintain the accuracy level required in the forecasting process.

- Computational time

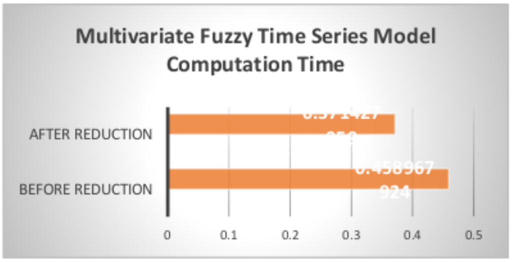

Figure 3. Computational time

The most important finding from the application of this newly developed method was the significant improvement in the computational time after a dimension reduction process of 80.9% was applied as indicated in Figure 2. Moreover, Figure 3 also shows that the dataset simplified through the reduction process was able to significantly improve the computational time and this also indicates the effectiveness of the core and reduct method.

The findings of this research are observed to be in line with the results of previous studies [28] [29] [30] [31] [10]. Moreover, Cheng et al. [30] studied the same sector by applying high-order MFTS based on RST to forecast TAIEX (Taiwan Stock Exchange Capitalization Weighted Stock Index) for a 10-year period. However, this present study improved on the research of Chen (1996) and Yu (2004) by adding 4 more different period parameters and orders.

4. Conclusion

MFTS with dimension reduction was found to be more effective in forecasting than those without dimension reduction. This means the concept of core and reduct in RST has the ability to improve the efficiency of computational time and weight in the MFTS method. The improvement also aided the forecasting process without reducing the level of accuracy. Meanwhile, it is important to note that the core of the analyzed dataset was the output from the dimension reduction process.

References

[1] M. Geurts, G. E. P. Box, and G. M. Jenkins, "Time Series Analysis: Forecasting and Control," *J. Mark. Res.*, 1977, doi: 10.2307/3150485.

[2] S. J. Taylor and B. Letham, "Forecasting at Scale," *Am. Stat.*, 2018, doi: 10.1080/00031305.2017.1380080.

[3] Z.-Y. Wang, Y.-C. Lin, S.-J. Lee, and C.-C. Lai, "A Time Series Forecasting Method," *ITM Web Conf.*, 2017, doi: 10.1051/itmconf/20171203008.

[4] C. Chatfield, *Time-series forecasting*. 2000.

[5] C. Beaumont, S. Makridakis, S. C. Wheelwright, and V. E. McGee, "Forecasting: Methods and Applications," *J. Oper. Res. Soc.*, 1984, doi: 10.2307/2581936.

[6] D. Kartini, H. Rusdiani, and A. Farmadi, "Analisis Pengaruh Banyak Orde pada Metode Multivariate High-Order Fuzzy Time Series untuk Prediksi Duga Muka Air Waduk," *J. Edukasi dan Penelit. Inform.*, 2019, doi: 10.26418/jp.v5i1.27265.

[7] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series - Part I," *Fuzzy Sets Syst.*, 1993, doi: 10.1016/0165-0114(93)90355-L.

[8] H. T. Nguyen and B. Wu, "Fuzzy time series analysis and forecasting," *Stud. Fuzziness Soft Comput.*, 2006, doi: 10.1007/11353492_9.

[9] N. Güler Dincer and Ö. Akkuş, "A new fuzzy time series model based on robust clustering for forecasting of air pollution," *Ecol. Inform.*, vol. 43, no. December, pp. 157–164, 2018, doi: 10.1016/j.ecoinf.2017.12.001.

[10] T. A. Jilani, S. M. A. Burney, and C. Ardil, "Multivariate High Order Fuzzy Time Series Forecasting for Car Road Accidents," *World Acad. Sci. Eng. Technol.*, vol. 2, no. 1, pp. 288–293, 2007.

[11] A. Fakhru Rosyad, Farikhin, and J. E. Suseno, "Information Systems of Forecasting Incidence Rates of Dengue Fever Disease Using Multivariate Fuzzy Time Series," *E3S Web Conf.*, vol. 202, 2020, doi: 10.1051/e3sconf/202020214005.

[12] I. M. C. SATRIA, I. K. G. SUKARSA, and K. JAYANEGARA, “PERAMALAN JUMLAH WISATAWAN AUSTRALIA YANG BERKUNJUNG KE BALI MENGGUNAKAN MULTIVARIAT FUZZY TIME SERIES,” *E-Jurnal Mat.*, 2015, doi: 10.24843/mtk.2015.v04.i03.p094.

[13] K. TAN, “Data Mining Cluster Analysis : Basic Concepts and Algorithms Lecture Notes for Chapter 8 Introduction to Data Mining by What is Cluster Analysis ?,” pp. 1–54, 2018.

[14] P. Esling and C. Agon, “Time-series data mining,” *ACM Comput. Surv.*, 2012, doi: 10.1145/2379776.2379788.

[15] C. Wu, Y. Yue, M. Li, and O. Adjei, “The rough set theory and applications,” *Engineering Computations (Swansea, Wales)*. 2004, doi: 10.1108/02644400410545092.

[16] M. S. Raza and U. Qamar, “An incremental dependency calculation technique for feature selection using rough sets,” *Inf. Sci. (Ny).*, 2016, doi: 10.1016/j.ins.2016.01.044.

[17] J. Komorowski, L. Polkowski, and A. Skowron, “Rough sets: A tutorial,” in *Rough fuzzy hybridization: A new trend in decision-making*, S. K. Pal and A. Skowron, Eds. Springer, 1999, pp. 3–98.

[18] M. Sudha and A. Kumaravel, “Comparative Analysis between Rough set theory and Data mining algorithms on their Prediction,” *Glob. J. Pure Appl. …*, vol. 13, no. 7, pp. 3249–3260, 2017, [Online]. Available: https://www.ripublication.com/gjpam17/gjpamv13n7_35.pdf.

[19] M. Abdolrazzagh-Nezhad, “Enhanced cultural algorithm to solve multi-objective attribute reduction based on rough set theory,” *Math. Comput. Simul.*, vol. 170, pp. 332–350, 2020, doi: 10.1016/j.matcom.2019.11.005.

[20] M. Sudha, “Intelligent decision support system based on rough set and fuzzy logic approach for efficacious precipitation forecast,” *Decis. Sci. Lett.*, vol. 6, no. 1, pp. 95–106, 2017, doi: 10.5267/j.dsl.2016.6.002.

[21] H. F. Herdiyatmoko, “REDUKSI ATRIBUT BERDASAR MATRIK DISCERNIBILITY TEORI ROUGH SET DALAM SIMULASI BENCANA KEBAKARAN,” *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, 2017, doi: 10.24176/simet.v8i1.846.

[22] F. Aditya, D. Devianto, and M. Maiyastri, “PERAMALAN HARGA EMAS INDONESIA MENGGUNAKAN METODE FUZZY TIME SERIES KLASIK,” *J. Mat. UNAND*, 2019, doi: 10.25077/jmu.8.2.45-52.2019.

[23] S. Panigrahi and H. S. Behera, “A study on leading machine learning techniques for high order fuzzy time series forecasting,” *Eng. Appl. Artif. Intell.*, vol. 87, no. August 2019, p. 103245, 2020, doi: 10.1016/j.engappai.2019.103245.

[24] N. Tran, T. Nguyen, B. M. Nguyen, and G. Nguyen, “A Multivariate Fuzzy Time Series Resource Forecast Model for Clouds using LSTM and Data Correlation Analysis,” 2018, doi: 10.1016/j.procS.2018.07.298.

[25] F. Alpaslan, O. Cagcag, C. H. Aladag, U. Yolcu, and E. Egrioglu, “A novel seasonal fuzzy time series method,” *Hacettepe J. Math. Stat.*, 2012.

[26] R. C. Tsaur, “A fuzzy time series-Markov chain model with an application to forecast the exchange rate between the Taiwan and us Dollar,” *Int. J. Innov. Comput. Inf. Control*, vol. 8, no. 7 B, pp. 4931–4942, 2012.

[27] S. Rissino and G. Lambert-Torres, “Rough Set Theory — Fundamental Concepts, Principals, Data Extraction, and Applications,” in *Data Mining and Knowledge Discovery in Real Life Applications*, 2009.

[28] L. Lei, W. Chen, B. Wu, C. Chen, and W. Liu, “A building energy consumption prediction model based on rough set theory and deep learning algorithms,” *Energy Build.*, vol. 240, p. 110886, 2021, doi: 10.1016/j.enbuild.2021.110886.

[29] C. H. Cheng, H. J. Teoh, and T. L. Chen, “High-order fuzzy time series based on rough set for forecasting TAIEX,” *Proc. Sixth Int. Conf. Mach. Learn. Cybern. ICMLC 2007*, vol. 3, no. August, pp. 1354–1358, 2007, doi: 10.1109/ICMLC.2007.4370355.

[30] N. Tran, T. Nguyen, B. M. Nguyen, and G. Nguyen, “A Multivariate Fuzzy Time Series Resource Forecast Model for Clouds using

LSTM and Data Correlation Analysis," *Procedia Comput. Sci.*, vol. 126, no. November, pp. 636–645, 2018, doi: 10.1016/j.procS.2018.07.298.

[31] N. Fauziah, S. Wahyuningsih, and Y. N. Nasution, "Peramalan Mengunakan Fuzzy Time Series Chen (Studi Kasus : Curah Hujan Kota Samarinda)," *Statistika*, vol. 4, no. 2, pp. 52–61, 2016.

[32] C. H. Cheng, T. L. Chen, H. J. Teoh, and C. H. Chiang, "Fuzzy time-series based on adaptive expectation model for TAIEX forecasting," *Expert Syst. Appl.*, 2008, doi: 10.1016/j.eswa.2006.12.021.

hasil cek-manuskrip sugiyarto

ORIGINALITY REPORT

16%	11%	12%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Sheng-Tun Li, Yi-Chung Cheng. "A Stochastic HMM-Based Forecasting Model for Fuzzy Time Series", IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2010 Publication	1%
2	docshare.tips Internet Source	1%
3	Submitted to Institute of Graduate Studies, UiTM Student Paper	1%
4	tjzhifei.github.io Internet Source	1%
5	www.studymode.com Internet Source	1%
6	journal.uad.ac.id Internet Source	1%
7	www.e3s-conferences.org Internet Source	1%

8	Riki Ariyanto, R Heru Tjahjana, Titi Udjiani. "Analysis of The Efficiency SA (Simulated Annealing) Model And Fuzzy Time Series (Case Study: Inventory System PT Y)", Journal of Physics: Conference Series, 2021 Publication	1%
9	Submitted to Syiah Kuala University Student Paper	1%
10	Submitted to King's College Student Paper	<1%
11	www.ripublication.com Internet Source	<1%
12	zh.scientific.net Internet Source	<1%
13	Abhishekh, Sanjay Kumar. "Handling higher order time series forecasting approach in intuitionistic fuzzy environment", Journal of Control and Decision, 2019 Publication	<1%
14	www.eusflat.org Internet Source	<1%
15	Cheng, Yi-Chung, and Sheng-Tun Li. "Fuzzy Time Series Forecasting With a Probabilistic Smoothing Hidden Markov Model", IEEE Transactions on Fuzzy Systems, 2012. Publication	<1%

16	Yu, H.K.. "Weighted fuzzy time series models for TAIEX forecasting", Physica A: Statistical Mechanics and its Applications, 20050415 Publication	<1 %
17	Submitted to Universiti Tenaga Nasional Student Paper	<1 %
18	Annisa Eka Haryati, Sugiyarto. "Clustering with Principal Component Analysis and Fuzzy Subtractive Clustering Using Membership Function Exponential and Hamming Distance", IOP Conference Series: Materials Science and Engineering, 2021 Publication	<1 %
19	Submitted to Stefan cel Mare University of Suceava Student Paper	<1 %
20	Hexiang Bai, Yong Ge, Jin-Feng Wang, Yi Lan Liao. "Using rough set theory to identify villages affected by birth defects: the example of Heshun, Shanxi, China", International Journal of Geographical Information Science, 2010 Publication	<1 %
21	Wangren Qiu, Ping Zhang, Yanhong Wang. "Fuzzy Time Series Forecasting Model Based on Automatic Clustering Techniques and	<1 %

Generalized Fuzzy Logical Relationship", 'Hindawi Limited', 2015
Internet Source

22	Yousif Alyousifi, Mahmod Othman, Ibrahima Faye, Rajalingam Sokkalingam, Petronio C. L. Silva. "Markov Weighted Fuzzy Time-Series Model Based on an Optimum Partition Method for Forecasting Air Pollution", International Journal of Fuzzy Systems, 2020 Publication	<1 %
23	res.mdpi.com Internet Source	<1 %
24	Nevin Güler Dincer, Özge Akkuş. "A new fuzzy time series model based on robust clustering for forecasting of air pollution", Ecological Informatics, 2018 Publication	<1 %
25	Yousif Alyousifi, Mahmod Othman, Akram A. Almohammedi. "A Novel Stochastic Fuzzy Time Series Forecasting Model Based on a New Partition Method", IEEE Access, 2021 Publication	<1 %
26	cronfa.swan.ac.uk Internet Source	<1 %
27	grammars.grlmc.com Internet Source	<1 %

journalofscience.org

28	Internet Source	<1%
29	ocs.unud.ac.id Internet Source	<1%
30	sersc.org Internet Source	<1%
31	www.hindawi.com Internet Source	<1%
32	"Computational Intelligence in Data Mining", Springer Science and Business Media LLC, 2020 Publication	<1%
33	Hsien-Lun Wong. "An Evaluation of Comparison between Multivariate Fuzzy Time Series with Traditional Time Series Model for Forecasting Taiwan Export", 2009 WRI World Congress on Computer Science and Information Engineering, 03/2009 Publication	<1%
34	Naresh Kumar, Seba Susan. "Particle swarm optimization of partitions and fuzzy order for fuzzy time series forecasting of COVID-19", Applied Soft Computing, 2021 Publication	<1%
35	Zuhaimy Ismail, Riswan Efendi, Mustafa Mat Deris. "Application of Fuzzy Time Series	<1%

	Approach in Electric Load Forecasting", New Mathematics and Natural Computation, 2015 Publication	
36	content.iospress.com Internet Source	<1 %
37	core.ac.uk Internet Source	<1 %
38	dspace.uii.ac.id Internet Source	<1 %
39	jwcn-eurasipjournals.springeropen.com Internet Source	<1 %
40	www.scribd.com Internet Source	<1 %
41	"Intelligent Systems", Springer Science and Business Media LLC, 2020 Publication	<1 %
42	Chao-Dian Chen. "A new method for forecasting the TAIEX based on high-order fuzzy logical relationships", 2009 IEEE International Conference on Systems Man and Cybernetics, 10/2009 Publication	<1 %
43	Chi-Chen Wang, Yi-Hsien Tu, Hsien-Lun Wong. "The Comparison between Multivariate Fuzzy Time Series and Traditional Time Series Modeling to Forecasting China Exports", 2009	<1 %

Fourth International Conference on Innovative Computing, Information and Control (ICICIC), 2009
Publication

44	Cagdas Hakan Aladag, Ufuk Yolcu, Erol Egrioglu, Ali Z. Dalar. "A new time invariant fuzzy time series forecasting method based on particle swarm optimization", Applied Soft Computing, 2012 Publication	<1 %
45	Hsuan-Shih Lee, Ming-Tao Chou †. "Fuzzy forecasting based on fuzzy time series", International Journal of Computer Mathematics, 2004 Publication	<1 %
46	Liu, Nian Ping, Hong Tu Wang, and Zhi Gang Yuan. "Evaluation Indexes of Sand Liquefaction Analysed by Rough Set", Advanced Materials Research, 2011. Publication	<1 %

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		