HASIL CEK_MBAK GRETA JURNAL 3

by Cek_mbak Greta Jurnal 3 Mbak Greta Jurnal 3

Dimension Reduction of Multivariate Fuzzy Time Series by Core and Reduct

Sugiyarto Surono[*], Rani Nuraeni[2], Joko Eliyanto[3]

Abstract

In this era, time series is widely known as one of the data where we can always find in a daily basis. Forecasting on this type of data usually needs an involvement from more than one factor. Generally, the movement of time series data tends to fluctuate. This characteristics is assumed as the fuzziness level on time series data. This data also can be defined as multivariate time series. Forecasting process applied in the data with high number of variables will impacted on the high level of computational weight. This paper tries to develop the method for forecasting process on multivariate time series data with computational weight which significantly lower. In this paper, multivariate fuzzy time series method will be combined with dimension reduction method. The application of dimension reduction is to reduce the dimension of the data but still maintaining the important information inside the data. Before the forecasting process is initiated, reduction dimension process is applied first to acquire the data with lower number of dimension. This paper used Core and reduct method which known as one of the concept in Rough set theory to do the dimension reduction process. The result of the application of core and reduct method is proven to be able to decrease the computational time up to 80.9% without significantly decrease the level of accuracy which still can be maintained in 99%. This paper concludes that the proposed combination of methods is able to produce the prediction with 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 observation conducted sequentially in certain time interval. Data time series is able to be applied as the supporting factor on the decision-making in forecasting process [1]. This application can be done because data time series is able to provide values in detail in the form of number on the previous time interval. Other than that, data time series is believed to have the repeating pattern, which means that the period occurred in the past will be repeated in the moment or in the future. Data time series also can be applied to predict with the assumption that the function in the future is dependent on the function in the past. Hence, we can observe and predict what will be happen in certain timeline through the application of the data in the past [2] [3] [4].

Forecasting is a method to predict certain observation value in the future by considering and observing the data acquired in the past or in the moment. Forecasting is applied to determine how the pattern of certain data works. This method emerged as one of the important tools on producing effective and efficient plan [5]. Time series forecasting process is a process of predicting certain occurrence in the future by considering the historical pattern of the analyzed data in the form of data series. One of the method which applied to predict the time series data is Multivariate Fuzzy Time Series (MFTS) [6]. MFTS is a data forecasting method with the application of fuzzy theorem as its basis [7] [8] [9]. The system of MFTS forecasting process worked by capturing the pattern of certain data to be applied to project the pattern of the data in the future with the usage of only more than one variables [10] [11] [12].

The problem which usually emerge in the application of MFTS is the high level of computational weight and the increase of complexity during the forecasting process. This problem occurs because of the number of applied variables is relatively high [13 [14]. In addition to that, the unavailability of the exact pattern rule also affecting the termination of certain variables which may significantly impacting the accuracy of forecasting process. These problems leads in the need of the method which able to reduce the computational weight without significantly altering

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

the level of accuracy. The objective of the reduction dimension method is to transform the high dimension data into the lower dimension data but still maintain the main characteristics of the data [15] [16]. The output of this method is new information in the form of pattern of rules which able to be applied as the reference in the forecasting process, so that the performance of the computational weight can be improved. Core and reduct is one of the technique in rough set theory which aims to decrease the dimension of certain dataset until its informational core are left [17] [18] [19] [20].

There are few previous research which related to the forecasting and dimension reduction. The research worked by Hendrik Fery Herdiyatmoko [21] applying RST with the objective to decrease the parameter or attributes of the fire accident which resulting in the decrease of the complexity for the fire accident data analysis. The output of this research is applied as the reference for determining the evacuation route for fire accidents.

Another research worked by Tahseen A Jilani and co. [10] studying MFTS to predict the number of car accident in Belgium with the application of one main variable which is the death toll of the accident and four supporting variables. Furthermore, this research also using MFTS to analyze the pattern of the cause of the accident in determining the fuzzy premises and assurances for the car and life insurance measurement.

In addition to that, the other research worked by I Made Candra Satria [12] analyzing the comparison of the forecasting result from the FTS and MFTS. There are three variables applied in this research, which are the number of Australian tourists, Inflation in Indonesia, and the currency rate of AUD to IDR. The output of this research shows that the vacation number of the Australian tourists in Bali with the application of MFTS produce better accuracy than the application of FTS.

This paper focuses on the application of MFTS in forecasting process. This forecasting process applying the basic concept of FTS method by the approach which developed to solve the linguistic variables problem. MFTS is one of the improvement from the Fuzzy Time Series (FTS) method [22] [23]. MFTS considers more than one variables or factors where one of the factor is functioned as the main variable and the other as the supporting variables. This approach combines the linguistic variables with the fuzzy logic analytic process in the time series data to overcome the usage of the unwanted data. It generally known that although Multivariate Fuzzy Time Series (MFTS) has the ability to processing more than one variables, this method is not able to determine which variables is proper to applied in the forecasting process [24] [9]. The application of these unwanted variables affecting the forecasting method applied by the computational system takes longer time to produce. Furthermore, it also generally known that Rough Set Theory method is one of the method of dimension reduction which proven to be able to decrease the computational weight. The main problem which this paper wanted to solve is how to decrease the computational weight in the forecasting process. This paper tries to identify how effective is the combination of MFTS and Rough Set Theory method to forecasting process with high number of variables. The main objective of this paper is to produce the forecasting process with high level of accuracy and lower computational weight than the standard method. As the process of the method improvement, we applied this method in the sample data which suitable with the desired condition. This method only developed in the numerical data just as how the MFTS method is developed.

2. Method

2.1 Multivariate Fuzzy Time Series (MFTS)

Multivariate Fuzzy Time Series (MFTS) is developed based on Fuzzy Time Series (FTS). This concept of this method is invented by Song and Chissom which applied to determine the forecasting problem with the historical data in the form of linguistic value. The basic principle of the MFTS method measurement 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] Suppose there is given the universe of $Y(t)$ with $(t = \cdots, 0,1,2, \dots, n, \dots)$ as the *subset* from the real number (R) which defined as 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 using the fuzzy relation equation to develop the forecasting model with the assumption that the observation on time of t only dependent in the observation result of the previous time, which defined as follows.

Definition 2.2 [25] Suppose $F(t)$ 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 of $F(t) = F(t-1)°R(t,t-1)$. In this equation, ' ° ' is the operator of the max-min composition. The R relation is described as the first order model from $F(t)$.

Definition 2.3 [26] Suppose $F(t-1) = A_i$ and $F(t) = A_j$, *Fuzzy Logical Relationship* (FLR) can be defined as $A_i \rightarrow A_j$ where A_i is described as Current State and A_j is described as Next State. If there is FLR which acquired from the *state* A_2, then the transition is made to the other state of A_j , $j = 1,2,\ldots,n$, in the form of:

$$A_2 \rightarrow A_1, A_2 \rightarrow A_4, A_2 \rightarrow A_2$$

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

$$A_2 \rightarrow A_1, A_4, A_2$$

Definisi 2.4 [26] Jika $F(t)$ disebabkan oleh lebih banyak himpunan *fuzzy* $F(t-n), F(t-n+1), \ldots, F(t-1)$, relasi *fuzzy* direpresentasikan oleh $A_{i1}, A_{i2}, \ldots, A_{in} \rightarrow A_j$, dimana $F(t-n) = A_{i1}, (t-n+1) = A_{i2}, \ldots, F(t-1) = A_{in}$. Relasi ini disebut model *Fuzzy Time Series* orde ke-n atau *Multivariate Fuzzy Time Series.*

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

The basic concept of Fuzzy is developed by L.Zadeh. This concept then improved by Song and Chissom and further developed by Cheng and co., where the slight change in the interval and weight determination are applied by the application of Fuzzy Logical Relationship (FLR). This weight measurement is worked by inserting all of the relationship and measured based on the same FLR repetition. The forecasting process by using this method does not necessarily need certain learning system from the complex one which affecting in the relatively easy application and development.

[Type here]

2.2 Rough Set Theory (RST)

Rough set theory is popularized by Zdiszlaw Pawlak in 1982]. This theory is a methodology focusing in the classification and analysis of imprecise, uncertainty, or uncompleted information and knowledge.

The reduction process in Rough Set Theory can be applied through Core and Reduct method

Definition 2.5 [27] In the rough set, certain set of data is represented as a table, where the rows represents cases, occurrence, patients, or certain objects and the columns represents variable attributes, observation, properties, etc. This table is generally described as Information systems which notated by IS and defined as follows:

$$IS = (U, P) \quad (2)$$

Where U is the unlimited set which are not empty from the objects and defined as Universe and A is the limited set which are not empty from the attributes where: $a: U \rightarrow V_a$ 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 symbolized as $IND_{IS}(B)$,can be defined as:

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

$IND_{IS}(B)$ is defined as *B-indiscrenibility relation*. $IND_{IS}(B)$ is the equivalence relation. If $(x, x') \in IND_{IS}(B)$, then, the objects of x and x' are the objects which are indiscernible towards each other by attributes of B. The class which equivalent with *B-indiscrenibility relation* is notated by $[x]_B$ and described as equivalent class.

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

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

On the contrary, if a is considered as indispensable then a is very important in attributes of B.

The set of B is independent if all of its attributes are required. Each of subset B^i from B is described as the reduct from B is B^i is independent and $IND_{IS}(B') = IND_{IS}(B)$. Thus, reduct can be defined as the set of the attributes which are able to provide the accuracy of the forecasting process equal to the result if all of the attributes are applied. The attributes which not considered as reduct is the attributes which will be erased and wouldn't affecting the result of the forecasting process.

Definition 2.8 [27] Suppose $B \subseteq P$ and $Core$ from B is the set of all the indispensable attributes of B, then Core can be defined as:

$$Core(B) = \cap \, Red(B) \qquad (5)$$

Where $Red(B)$ is the set of all the reduct result on B. Because core is the intersection from all of the reduct result, then core is included in each of the reduction, which means that each of the core attributes is included on few reduction process. In other words, core can be considered as the most important part of the attributes because there are no indispensable attributes in the analyzed set.

2.3 Multivariate Fuzzy Time Series Algorithm

In the Multivariate Fuzzy Time Series theory, Fuzzy Logical Relation (FLR) is one of the important factor which affecting the accuracy of the MFTS forecasting process [9]. Cheng method provides a slight different method in the interval determination and the weight measurement through the application of FLR. This method is worked by inserting all of the relationship and measured based on the same FLR repetition. The first step of the time series data forecasting process with the application of Fuzzy time series weighted based on the Cheng method is as follows:

First Step: Determining the universe.

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

$$U = [D_{min}, D_{max}] \qquad (6)$$

Where Dmin and Dmax is the minimum and maximum value from the historical data which divide the set of the universe into few interval of $U = \{u_1, u_2, \ldots, u_n\}$ with the equal distance.

Second Step: Determining the wide of the interval

The measurement of the wide of the interval is done by applying the frequency distribution through these following steps:

a) Determining the range through this following equation:

$$R = D_{max} - D_{min} \qquad (7)$$

Where R is range and Dmin and Dmax is the minimum and maximum value.

b) Determining the number of the interval class through the Sturges equation which defined as follows:

$$K = 1 + 3.322 \times \log(n) \qquad (8)$$

Where n = number of data.

c) Measuring the wide of the interval through this following equation:

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

Then, divide the interval with the suitable number and the length of the interval so that each of the interval is measured with:

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

e. Determining the median value with m_i by applying this following equation.

$$m_i = \frac{lower\ limit + upper\ limit}{2} \qquad (11)$$

Where i is the number of fuzzy set

Third step: Determining the fuzzy set.

Set of fuzzy is formed by observing the number of frequency which are different. For the most number of frequency, divide its value into the equal value of h interval. Then, for the second most number of frequency, divide its value into the equal value of h-1 interval. This measurement then repeated until the interval which cannot be divided again.

Fourth Step: Define the set of Fuzzy for all the universe in each variables.

Each of the fuzzy set in each of the variables is defined in the number of interval which already determined as follows:

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

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

With μ_{ij} is the membership degree which measured by this following equation:

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

The equation of A_i can be explained through these 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} is 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 set of fuzzy from the main variable of 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) and symbol "/" represents the membership degree of $\mu_{ij}(u_i)$ towards A_i where its values are 0,0.5, and 1.

Fifth Step: Fuzziness process of the data towards historical data.

Fuzziness process is processed by transforming the actual data into the fuzzy linguistic value. This step is done to find the suitable set of fuzzy for each of the data.

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

The fuzzy logical relationship is defined as $A_i \rightarrow A_j$. A_i is described as the current state $Y(t-1)$ and A_j is the next state in time of-t. Suppose there is acquired FLR which 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: Transforming the relation weight FLR into Fuzzy Logical Relationship Group (FLGR).

This step is initiated by inserting all of the relationship weight and determining the weight based on the same order and iteration. FLR which have the same current state (A_i) then classified into one group in the form of weighting process. Suppose there are certain equal order of FLR illustrated as:

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

Where t described as time. Then, the acquired weight from FLR relation inserted into the form of weighting matrix (W) which its equation 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 w ij is the weight in row of-i and column of-j with $i = 1,2, \ldots, n; and \; j = 1,2, \ldots, n$.

Eighth Step: Transforming the weighing matrix ($\boldsymbol{W}$) into the form of standardized weighing matrix ($\boldsymbol{W^*}$).

The equation of the standardized weighing matrix ($\boldsymbol{W^*}$) is illustrated 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: Determining the Defuzzification of the predicted value.

To produce the forecasting value, standardized weighing matrix ($\boldsymbol{W^*}$) is multiplied with the median value (m_i). The median value which will be applied in the forecasting process is the median value from the main factor. Thus, the forecasting measurement will be 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 $w_{ij}^{*} = \frac{w_{ij}}{\sum_{j=1}^{n} w_{ij}}$.

If FLRG from A_i is an empty set ($A_i \rightarrow \emptyset$) then the forecasting from $F(t)$ is m_i which is the median value from interval u_i.

3 Result and Discussion

3.1 Data Analysis

The variable applied in this research is the climate data acquired from the official site of BMKG (Meteorology, Climatology, and Geophysics Institute) in the form of daily weather data from the Climatology station in Sleman, Yogyakarta. Data is collected since 24 April 2018 until 26 April 2021 with 10 variables, such as 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 most of wind direction (^{0}). To determine the condition of the applied dataset, the application of Exploratory Data Analysis (EDA) needs to be initiated. This evaluation is important to improve the accuracy of the following forecasting process. Without the application of EDA, there will be a possibility of the loss important information from the dataset. EDA also able to save the time of forecasting process. For the conclusion, EDA is functioned for handling the possibility of loss information, identifying the most important variables, and testing 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

Before forecasting process is initiated, dimension reduction process through rough set theory (RST) needs to be run first. In RST, dataset which act as input and output in dimension reduction can be perceived as system information. Based on rough set theory, core of the dataset can be extracted through the reduction process of those system information. The result of dimension reduction process worked in this paper is 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

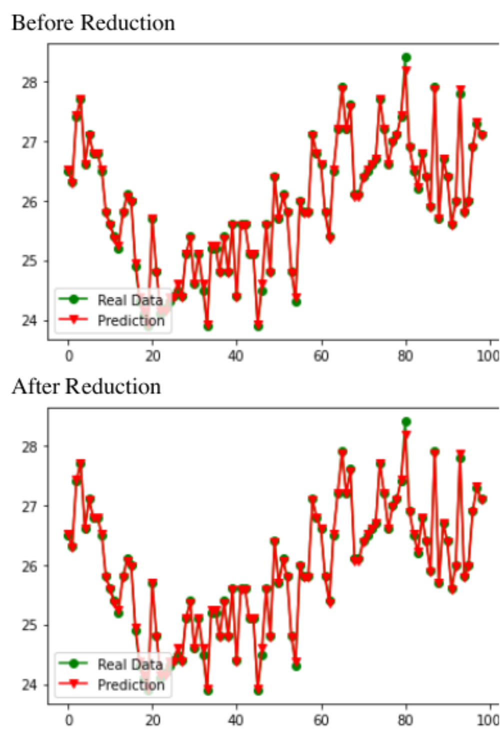

Figure 1. Forecasting result

The forecasting result in figure 1 illustrate that the Multivariate fuzzy time series with and without reduction provides the similar result. This shows that the application of dimension reduction in this research is able to maintain the accuracy level of the forecasting process which also can be implied that MFTD forecasting method is working very well.

- Accuracy

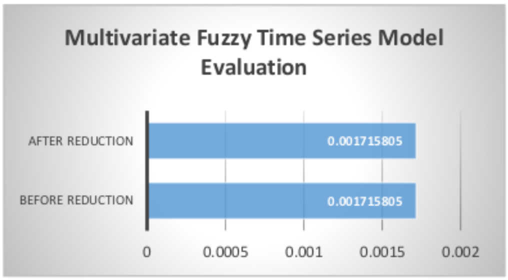

Figure 2. Model Evaluation

The result shown by this new improved method illustrate that the core and reduct concept in RST is able to well-maintained the accuracy level of forecasting process. In figure 2, we can observe that the dataset with or without reduction process are producing the similar level of accuracy, which is 99%. This concludes that RST is proven to be able to maintain the accuracy level in forecasting process.

- Computational time

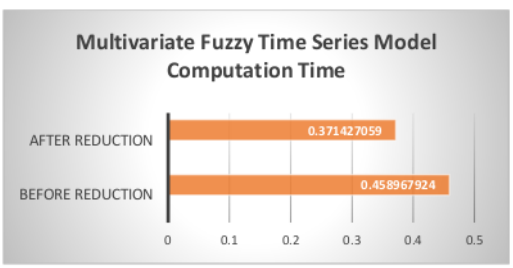

Figure 3. Computational time

The most important result from the application of this new developed method is that whether core and reduct concept in RST is able to improve the computational time. Figure 2 shows that the computational time is significantly improved after 80.9% dimension reduction process is applied.

In figure 3, we also can observe that the dataset which have been simplified through reduction process is able to significantly improve the computational time. The improvement on computational time on forecasting process through MFTS method indicate that the dimension reduction core and reduct works very well.

The output resulted in this research also linear towards the results from the previous research [28] [29] [30] [31] [10]. Cheng and co. [30] also did the research on the same sector with the application of high-order MFTS based on RST to forecast TAIEX (Taiwan Stock Exchange Capitalization Weighted Stock Index) for 10 years period. This research then further developed by Chen (1996) and Yu (2004) by adding 4 more different period parameter and different order.

4. Conclusions

In forecasting process, MFTS with the application of dimension reduction works better than without the dimension reduction application. This means that the concept of core and reduct in RST is proven to be able to improve the efficiency of computational time and computational weight on MFTS method.

The output resulted from this dimension reduction process is the core from the analyzed dataset. Core and reduct application focuses on the efficiency improvement of the computational time and computational weight in the forecasting method. This improvement also able to produce without decreasing the accuracy level of forecasting process through MFTS combined with RST.

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_MBAK GRETA JURNAL 3

ORIGINALITY REPORT

16%	10%	12%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Singh, S.R.. "A simple method of forecasting based on fuzzy time series", Applied Mathematics and Computation, 20070301 Publication	1%
2	journal.uad.ac.id Internet Source	1%
3	Submitted to Institute of Graduate Studies, UiTM Student Paper	1%
4	Sheng-Tun Li, , and 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%
5	docshare.tips Internet Source	1%
6	tjzhifei.github.io Internet Source	1%
7	www.e3s-conferences.org Internet Source	

1 %

8	www.hindawi.com Internet Source	1 %
9	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 %
10	Submitted to Syiah Kuala University Student Paper	<1 %
11	Arnita, N Afnisah, F Marpaung. "A Comparison of The Fuzzy Time Series Methods of Chen, Cheng and Markov Chain in Predicting Rainfall in Medan", Journal of Physics: Conference Series, 2020 Publication	<1 %
12	Zuhaimy Ismail, Riswan Efendi, Mustafa Mat Deris. "Application of Fuzzy Time Series Approach in Electric Load Forecasting", New Mathematics and Natural Computation, 2015 Publication	<1 %
13	www.ripublication.com Internet Source	<1 %
14	Yu, H.K.. "Weighted fuzzy time series models for TAIEX forecasting", Physica A: Statistical	<1 %

Mechanics and its Applications, 20050415
Publication

15 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 <1 %
Publication

16 Ching-Hsue Cheng, Hia-Jong Teoh, Tai-Liang Chen. "High-Order Fuzzy Time Series Based on Rough Set for Forecasting Taiex", 2007 International Conference on Machine Learning and Cybernetics, 2007 <1 %
Publication

17 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 <1 %
Publication

18 Mochammad Taufik, Ashri Shabrina Afrah, Endah Septa Sintiya, Didik Hariyanto. "A comparative study of time-series models for forecasting the indonesian gold price", Proceedings of the 5th International Conference on Sustainable Information Engineering and Technology, 2020 <1 %
Publication

19	grammars.grlmc.com Internet Source	<1 %
20	www24.us.archive.org Internet Source	<1 %
21	Tahseen Ahmed Jilani. "M-Factor High Order Fuzzy Time Series Forecasting for Road Accident Data", Advances in Soft Computing, 2007 Publication	<1 %
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	linknovate.com Internet Source	<1 %
24	www.coursehero.com Internet Source	<1 %
25	www.eusflat.org Internet Source	<1 %
26	Emrah Gulay, Okan Duru. "Hybrid modeling in the predictive analytics of energy systems and prices", Applied Energy, 2020 Publication	<1 %

27	Silvia Rissino, Germano Lambert-Torres. "Chapter 3 Rough Set Theory — Fundamental Concepts, Principals, Data Extraction, and Applications", IntechOpen, 2009 Publication	<1 %
28	docplayer.net Internet Source	<1 %
29	jurnal.umk.ac.id Internet Source	<1 %
30	www.jise.ir Internet Source	<1 %
31	"Information Computing and Applications", Springer Science and Business Media LLC, 2013 Publication	<1 %
32	"Intelligent Systems", Springer Science and Business Media LLC, 2020 Publication	<1 %
33	Abdul Aziz, Khoirun Nissa Isti Khomah, Samgadi Palgunadi Yohanes. "Prediction The Price of National Groceries Using Average Based Fuzzy Time Series With Song - Chissom and Markov Chain Approach", 2018 International Conference on Applied Information Technology and Innovation (ICAITI), 2018 Publication	<1 %

34	Agustiyara, Eko Priyo Purnomo, Rijal Ramdani. "Using Artificial Intelligence Technique in Estimating Fire Hotspots of Forest Fires", IOP Conference Series: Earth and Environmental Science, 2021 Publication	<1 %
35	Alireza Maheri, Shahin Jalili, Yousef Hosseinzadeh, Reza Khani, Mirreza Miryahyavi. "A comprehensive survey on cultural algorithms", Swarm and Evolutionary Computation, 2021 Publication	<1 %
36	M. H. Fazel Zarandi, A. Molladavoudi, A. Hemmati. "Fuzzy time series based on defining interval length with Imperialist Competitive Algorithm", 2010 Annual Meeting of the North American Fuzzy Information Processing Society, 2010 Publication	<1 %
37	export.arxiv.org Internet Source	<1 %
38	www.exeter.ac.uk Internet Source	<1 %
39	www.fuzzy.org.tw Internet Source	<1 %
40	www.mdpi.com Internet Source	<1 %

41 www.studymode.com
Internet Source
<1 %

42 www.tandfonline.com
Internet Source
<1 %

43 Hye-young Jung, Jin-hee Yoon, Seung-hoe Choi. "Fuzzy time series reflecting the fluctuation of historical data", 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery, 2010
Publication
<1 %

44 Yueh-Ju Lin, Chi-Chen Wang. "The dynamic analysis of Baltic exchange dry index", International Mathematical Forum, 2014
Publication
<1 %

45 Abhishekh, Sanjay Kumar. "Handling higher order time series forecasting approach in intuitionistic fuzzy environment", Journal of Control and Decision, 2019
Publication
<1 %

46 Erol Egrioglu, Eren Bas, Ufuk Yolcu, Mu Yen Chen. "Picture fuzzy time series: Defining, modeling and creating a new forecasting method", Engineering Applications of Artificial Intelligence, 2020
Publication
<1 %

47 Muhammad Summair Raza, Usman Qamar. "Understanding and Using Rough Set Based
<1 %

Feature Selection: Concepts, Techniques and Applications", Springer Science and Business Media LLC, 2019
Publication

48	downloads.hindawi.com Internet Source	<1 %
49	dspace.uii.ac.id Internet Source	<1 %
50	www.ukessays.com Internet Source	<1 %
51	ijettjournal.org Internet Source	<1 %
52	M. Sudha, A. Kumaravel. "Quality of classification with LERS system in the data size context", Applied Computing and Informatics, 2018 Publication	<1 %

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		