Journal of Computer Science and Cybernetics, V.37, N.1 (2021), 23–42
DOI 10.15625/1813-9663/37/1/15852

A TIME SERIES FORECASTING MODEL BASED ON LINGUISTIC FORECASTING RULES

PHAM DINH PHONG*

Faculty of Information Technology, University of Transport and Communications, 3 Cau Giay Street, Lang Thuong Ward, Dong Da District, Ha Noi, Viet Nam

Abstract. The fuzzy time series (FTS) forecasting models have been studied intensively over the past few years. The existing FTS forecasting models partition the historical data into subintervals and assign the fuzzy sets to them by the human expert's experience. Hieu et al. proposed a linguistic time series by utilizing the hedge algebras quantification to converse the numerical time series data to the linguistic time series. Similar to the FTS forecasting models, the obtained linguistic time series can define the linguistic, logical relationships which are used to establish the linguistic, logical relationship groups and form a linguistic forecasting model. In this paper, we propose a linguistic time series forecasting model based on the linguistic forecasting rules induced from the linguistic, logical relationships instead of the linguistic, logical relationship groups proposed by Hieu. The experimental studies using the historical data of the enrollments of University of Alabama and the daily average temperature data in Taipei show the outperformance of the proposed forecasting models over the counterpart ones. Then, to realize the proposed models in Viet Nam, they are also applied to the forecasting problem of the historical data of the average rice production of Viet Nam from 1990 to 2010.

Keywords. Hedge algebras, defuzzification, linguistic time series, linguistic logical relationship group.

1. INTRODUCTION

In recent decades, there have been many researches related to the forecasting problems published with the aim of improving the accuracy of forecasting results and reducing computational time. The fuzzy time series model firstly introduced by Song and Chissom in 1993 is based on the fuzzy set theory, in which the fuzzy set is considered as the computational semantics of linguistic words. This forecasting model is applied to forecast the enrollments of the University of Alabama [1, 2, 3]. Their researches are originated from observations of weather in a particular place in North America, where the weather data is described in terms of linguistic words such as *good, very good, quite good, very very good, cool, very cool, quite cool, hot, very hot, cold, very cold, very very cold.* Their studies have opened up a new field of research that has attracted a considerable amount of researches both in terms of methodology and application.

The methodological oriented research includes the model improvement by simplifying the calculation method proposed by Chen in [4], optimizing historical data partition intervals [5, 6, 7, 8, 9, 10, 11], applying the high-order fuzzy time series models [12, 13, 11], applying multi-factor fuzzy time series model [14, 15], improving the fuzzy defuzzification techniques

*Corresponding author.
E-mail address: phongpd@utc.edu.vn

[8, 9, 16, 17], ... The application-oriented research includes the problems of the enrollment forecasting [4, 5, 13, 8, 1, 3], temperature forecasting [10, 11, 15], stock forecasting [9, 10, 11, 17, 15], tourism demand forecasting [18], car road accident forecasting [16, 19], etc.

In spite of the remarkable achievements, in the research on the fuzzy time series forecasting model, there are still some problems that have not been optimally resolved. *The first* is how to partition the universe of discourse of the historical data into appropriate interval lengths and how many intervals are reasonable? If the number of intervals is too small, the forecasting result may give low forecasting accuracy due to lack of information, while choosing too many intervals may reduce the meaning of the fuzziness of the linguistic values. To solve this problem, the authors have applied individually or in combination with different techniques such as genetic algorithms [5, 13, 10], simulation annealing techniques, ant colony optimization, particle swarm optimization [6, 8, 9, 16, 20], granule computing, data clustering [20], hedge algebras [21, 22, 23], etc. to determine the best partition intervals. At present, how many partition intervals for each forecasting problem is reasonable still remains an open question. Is there any other effective solution to fuzzify the historical data and handle computing directly with linguistic words?

The second is the effective construction of fuzzy logical relationships, fuzzy logical relationship groups. The authors have studied to solve this problem by building high-order fuzzy time series models [13, 11], multi-factor fuzzy time series models [18, 15]. In [24], Dieu proposes the time-variant fuzzy logical group to replace the time-invariant one in Yu's model [17] in order to form a time-variant fuzzy time series model.

The third is the application of suitable fuzzy defuzzification techniques to improve the forecasting results. Various fuzzy defuzzification techniques have been proposed with their own advantages and disadvantages, a fuzzy defuzzification technique can be good for the first-order fuzzy time series model, but poor for the high-order fuzzy time series model and vice versa; a fuzzy defuzzification technique can be good for a designated forecasting problem, but poor for other forecasting problems and vice versa.

The fourth is the weakness of the methodology. Firstly, it is the matter of word semantics: humans have their habit of using linguistic words in natural language and therefore time series data can also be in the form of linguistic words. However, in the existing forecasting models, linguistic words are just the linguistic labels of the fuzzy sets designed for fuzzy time series. Besides, there is no formalism that connects the fuzzy sets with the linguistic labels. Secondly, it is the determination of the fuzzy sets used for a particular time series. For any variable $\mathcal{X}$ whose numerical domain is the universe of discourse $U_{\mathcal{X}}$, the use of the linguistic words associated with $\mathcal{X}$ is objective because the semantics of the words of $\mathcal{X}$ are often understood the same among the users. Meanwhile, the fuzzy sets designed on $U_{\mathcal{X}}$ depend a lot on the designer. Therefore, the question naturally arises is that is there a formalism for directly handling linguistic words in solving time series forecasting problem?

Hedge algebras (HA) [25, 26, 27] has effective applications in data mining [28, 29, 30, 31], fuzzy control [32], image processing [33], fuzzy time series [23, 34], etc. The HA exploits the semantic order of words in the linguistic value domain of the linguistic variable to form a mathematical formal basis for the linkage of fuzzy set based computational semantics with the inherent semantics of linguistic words.

Solving most of the problems mentioned above by a methodology seems to be impossible. To provide a formalism for directly handling linguistic words, Hieu et al. proposed a linguistic

time series forecasting model [21, 22] based on HA quantification [25, 26, 27] to convert numerical time series into linguistic time series. Similar to the fuzzy time series forecasting model, the obtained linguistic time series determines the linguistic, logical relationships (LLRs) which are used to establish the linguistic, logical relationship groups (LLRGs) and form a linguistic time series forecasting (LTS) model. Therefore, instead of partitioning a given numerical time series data into the intervals and defining the fuzzy sets on them, LTS model uses the linguistic words directly and it is a crucial condition to form a formalism for directly handling linguistic words. The crisp forecasted results are computed simply based on the real semantics of the used linguistic words which are generated automatically from the inherent semantics of the words.

In [24], Dieu proposed the time-variant fuzzy logical group (FLRG) concept in such a way that only put into the right-hand side of a FLRG the fuzzy sets occurred at or before the current forecasting time t (the time of the left-hand side of the FLRG under consideration). The aim of creating of the time-variant FLRG is to prevent the presence of a fuzzy set (an element) at a time after t in right-hand side of it.

In this paper, inspired by the time-variant FLRG, the *linguistic forecasting rule* (LFR) for a certain time t induced from the LLRs instead of the LLRGs [21, 22] is proposed in such a way that only the right-hand sides of the LLRs with the same left-hand sides occurred before or at the time t are put in chronological order into the right-hand side of the same LFR under consideration. Therefore, a new linguistic time series forecasting model based on the LFRs is built to improve the forecasted results. In addition, a new formula for calculating the crisp forecasted values more efficiently is used instead of the one used in [21, 22]. The experimental results over the time series datasets of the historical enrollment of University of Alabama observed from 1971 to 1992 and the daily average temperature observed from June 1996 to September 1996 in Taipei show that the proposed linguistic time series forecasting models are more efficient than the ones proposed in [21, 22]. Besides, to realize the proposed models, they are also applied to the forecasting problem of the historical data of the average rice production of Viet Nam from 1990 to 2010.

The rest of the paper is organized as follows: Section 2 is some basic concepts of hedge algebras. Section 3 describes some concepts of fuzzy time series and fuzzy time series forecasting models. Section 4 presents the linguistic time series forecasting model and the one based on linguistic forecasting rules. Section 5 shows the applications of the proposed forecasting models. Conclusion and remarks are included in Section 6.

2. SOME CONCEPTS OF HEDGE ALGEBRAS

For each linguistic variable $\mathcal{X}$ whose word-domain, denoted by $\mathrm{Dom}(\mathcal{X})$, is the set of the words in the natural language. We can easily observe that the words in the $\mathrm{Dom}(\mathcal{X})$ can be induced from two primary words, e.g., "*small*" and "*large*" (so-called the generator words) by applying linguistic hedges such as "*very*", "*little*". The generated words such as "*very small*", "*small*", "*little small*", etc. are linearly ordered based on their inherent qualitative semantics and they are comparable. It inspired Ho et al. to introduce a mathematic structure, so-called hedge algebras (HA) [26, 27]. A hedge algebra $\mathcal{AX}$ of variable $\mathcal{X}$ is an order-based structure $\mathcal{AX} = (X, G, C, H, \leq)$, where

- $X \subseteq \mathrm{Dom}(\mathcal{X})$ is a linguistic word set of variable $\mathcal{X}$.

- $G = \{c^-, c^+\}$ is a set of the generator words, where $c^- \leq c^+$, c^- and c^+ are the negative and the positive words, respectively.
- $C = \{\boldsymbol{0}, W, \boldsymbol{1}\}$ satisfies the order relation $\boldsymbol{0} \leq c^- \leq W \leq c^+ \leq \boldsymbol{1}$, where $\boldsymbol{0}$ and $\boldsymbol{1}$ are the least and the greatest constants, respectively; W is the neutral constant.
- $H = H^- \cup H^+$ is a set of linguistic hedges of variable $\mathcal{X}$, where H^- and H^+ are two sets of negative and positive linguistic hedges, respectively.
- $\leq$ is an order relation induced by the inherent qualitative semantics of the words of variable $\mathcal{X}$.

Each word x in X is represented as a string, i.e., either $x = c$ or $x = \sigma c$, where $c \in \{c^-, c^+\}$ and $\sigma = h_m \dots h_1$, $h_j \in H$, $j = 1, \dots, m$. Put $H(x) = \{\sigma x, \sigma \in H\}$, so $X = H(c^-) \cup H(c^+) \cup C$. $\mathcal{AX}$ is the linear hedge algebras if all hedges in H and all linguistic words in X are linearly ordered, respectively. Some primary properties of the linear hedge algebras are as follows.

The negative generator word c^- has negative sign, denoted by sign(c^-) = -1. Similarly, the positive generator word c^+ has positive sign, denoted by sign(c^+) = +1.

The negative hedges make the semantics of the generator words decreased, whereas, the positive hedges make the semantics of the generator words increased. For example, "*small*" $\leq$ "*less small*" and "*less large*" $\leq$ "*large*", the hedge "*less*" makes the semantics of "*small*" and "*large*" decreased, whereas, "*very small*" $\leq$ "*small*" and "*large*" $\leq$ "*very large*", the hedge "*very*" makes the semantics of "*small*" and "*large*" increased. The negative hedges are denoted by $H^- = \{h_{-q}, \dots, h_{-1}\}$, where $h_{-q} \leq \dots \leq h_{-2} \leq h_{-1}$. The positive hedges are denoted by $H^+ = \{h_1, \dots, h_p\}$, where $h_1 \leq h_2 \leq \dots \leq h_p$. Therefore, $H = H \cup H^+$. if $h \in H^+$ then sign(h) = +1, whereas if $h \in H$ then sign(h) = -1.

If the hedge h makes the semantics of the hedge k increased, h is positive with respect to k, whereas, if the hedge h makes the semantics of the hedge k decreased, h is negative with respect to k. The negativity and positivity of the hedges do not depend on the linguistic words on which they act. For example, "*very*" is positive with respect to "*less*", we have "*small*" $\leq$ "*less small*" then "*less small*" $\leq$ "*very less small*", or "*less old*" $\leq$ "*old*" then "*very less old*" $\leq$ "*less old*". If the hedge h strengthens the effect trend of the hedge k, sign(h, k) = +1, whereas, if the hedge h weakens the effect trend of the hedge k, sign(h, k) = -1. Therefore, the sign of a word $x = h_m h_{m-1} \dots h_2 h_1 c$ can be defined by

$$\text{sign}(x) = \text{sign}(h_m, h_{m-1}) \times \dots \times \text{sign}(h_2, h_1) \times \text{sign}(h_1) \times \text{sign}(c).$$

The word sign meaning is that sign(kx) = +1 $\rightarrow x \leq kx$ and sign(kx) = -1 $\rightarrow kx \leq x$.

On the semantic aspect, the semantics of the set of linguistic words $H(x)$, $x \in X$, which is generated from the linguistic word x are changed by using the linguistic hedges in H but they still convey the original semantics of the word x. Therefore, $H(x)$ reflects the fuzziness of x and the length of $H(x)$ can be used to express the *fuzziness measure* of x, denoted by $fm(x)$. When $fm(x)$ is mapped to a sub-interval in the normalized space [0, 1] following the order structure of X by a mapping v, it is called the *fuzziness interval* of x, denoted by $\Im(x)$.

Let $\mathcal{AX}$ be a linear hedge algebras. A function fm: $X \rightarrow [0, 1]$ is said to be a *fuzziness measure* of words in X provided that it satisfies the following properties:

(F1): $fm(c^-) + fm(c^+) = 1$ and $\sum_{h\in H} fm\,(hu) = fm\,(u)$, for $\forall u \in X$;

(F2): $fm(x) = 0$ for all $H(x) = x$, especially, $fm(\boldsymbol{0}) = fm(W) = fm(\boldsymbol{1}) = 0$;

(F3): $\forall x, y \in X, \forall h \in H$, the proportion $\dfrac{fm\,(hx)}{fm\,(x)} = \dfrac{fm\,(hy)}{fm\,(y)}$ which does not depend on any particular linguistic word on X is called the fuzziness measure of the hedge h, denoted by $\mu(h)$.

From (F1) and (F3), $fm(x)$, where $x = h_m \dots h_1 c$ and $c \in \{c^-, c^+\}$, can be recursively computed that $fm(x) = \mu(h_m) \dots \ \mu(h_1) fm(c)$, where $\sum_{h\in H} \mu\,(h) = 1$. The fuzziness measure of a word in X can be easily computed when the values of $fm(c)$ and $\mu(h_j) \in H$ are given.

Semantically quantifying mappings (SQMs): The semantically quantifying mapping of $\mathcal{AX}$ is an order-preserved mapping $v : X \to [0, 1]$ provided that it satisfies the following conditions:

(SQM1): it preserves the order based structure of X, i.e. $x \le y \to v(x) \le v(y), \forall x \in X$;

(SQM2): It is one-to-one mapping and $v(x)$ is dense in [0, 1].

Let fm be a fuzziness measure on X, $\sum_{i=-q}^{-1} \mu\,(h_i) = \alpha$, $\sum_{i=1}^{p} \mu\,(h_i) = \beta$, $\alpha, \beta > 0$ and $\alpha + \beta = 1$. $v\,(x)$ is computed recursively based on fm as follows

1. $v(W) = \theta = fm(c^-)$, $v(c^-) = \theta - \alpha fm(c^-) = \beta fm(c^-)$, $v(c^+) = \theta + \alpha fm(c^+)$;
2. $v\,(h_j x) = v\,(x) + \text{sign}\,(h_j x) \left(\sum_{i=\text{sign}(j)}^{j} fm(h_i x) - \omega\,(h_j x)\, fm\,(h_j x) \right)$,
   where $j \in [-q, p] = \{j\text{: } -q \le j \le p \ \&\ j \ne 0\}$ and
   $\omega(h_j x) = \frac{1}{2}[1 + \text{sign}(h_j x)\text{sign}(h_p h_j x)(\beta - \alpha)] \in \{\alpha, \beta\}$.

3. FUZZY TIME SERIES FORECASTING MODELS

3.1. Some basic concepts of fuzzy time series

The fuzzy time series forecasting model was introduced by Song and Chissom in 1993 [1, 2, 3] and enhanced by Chen [4] with a simple defuzzification technique but more accurate. Some basic concepts of fuzzy time series are as follows.

Definition 1. (Fuzzy time series) [1, 2, 3]. Let $Y(t)$ (t= ...0, 1, 2,...) be a subset of R^1, where t is the temporal variable. $Y(t)$ is the universe of discourse U on which the fuzzy sets $f_i(t)$, $i = 1, 2, \dots$ are defined. If $F(t)$ is a series of fuzzy sets $f_i(t)$ ($i = 1, 2, \dots$) then $F(t)$ is called a fuzzy time series on $Y(t)$.

Definition 2. (Fuzzy logical relationship) [1]. At the times t and $t-1$, if there exists a fuzzy relationship $R(t-1, t)$ between $F(t-1)$ and $F(t)$ such that $F(t) = F(t-1) * R(t-1, t)$, where * is an operator then $F(t)$ is said to be inferred from $F(t-1)$. The relationship between $F(t-1)$ and $F(t)$ is defined by the notation $F\,(t-1) \to F\,(t)$. If $F(t-1) = A_i$ and $F(t) = A_j$, the logical relationship between $F(t-1)$ and $F(t)$ is denoted by $A_i \to A_j$, where A_i is the left-hand side (current state) and A_j is the right-hand side (next state) of the fuzzy relation.

Definition 3. The fuzzy logical relationships (FLRs) which have the same left-hand side can be grouped together and they are called fuzzy logical relationship groups (FLRGs). Assume that there are fuzzy logical relationships $A_i \to A_{j1},\ A_i \to A_{j2}, ...,\ A_i \to A_{jn}$. They can be put into a group denoted as $A_i \to A_{j1},\ A_{j2}, ...,\ A_{jn}$.

In the Chen's model [4], the fuzzy sets in the right-hand side of every FLRG are unique. Whereas, a fuzzy set in the right-hand side of a FLRG can be repeated in the Yu's model [17]. For example, if there are the FLRs $A_i \to A_k$, $A_i \to A_j$, $A_i \to A_k$, the FLRG will be $A_i \to A_k$, A_j in the Chen's model and $A_i \to A_k$, A_j, A_k in the Yu's model.

Definition 4. (Time-variant fuzzy logical relationship groups) [6, 9]. The FLR is specified by the relation $F(t-1) \to F(t)$. Let $F(t-1) = A_i(t-1)$ and $F(t) = A_j(t)$, we have the relationship $A_i(t-1) \to A_j(t)$. If at the time t we have the FLRs $A_i(t-1) \to A_j(t),\ A_i(t1-1) \to A_{j1}(t1), ...,\ A_i(tk-1) \to A_{jk}(tk)$, the FLRG $A_i(t-1) \to A_j(t),\ A_{j1}(t1),\ A_{j2}(t2), ...,\ A_{jk}(tk)$ with $t1, t2, ..., tk \le t$ is called the time-variant FLRG.

3.2. The fuzzy time series forecasting model of Chen

Chen enhanced the Song and Chissom model [1, 2, 3] by using simplified arithmetic operations on fuzzy logical relationship groups instead of min-max composition operations in fuzzy logical relationships. The brief of Chen's forecasting model [4] is as follows:

Step 1. Partition the universe of discourse of the time series U into the equal length intervals $u_1,\ u_2,\ \ldots,\ u_p$.

Step 2. Define the fuzzy sets on the universe of discourse U.

Step 3. Fuzzify the universe of discourse of U.

Step 4. Establish the FLRs and the FLRGs.

Step 5. Forecast and defuzzify the fuzzy output data to get the crisp forecasted values. In this step, the forecasting and defuzzification principles are defined. The principles are as follows:

- Principle 1. If $A_i \to A_j$ and the maximum value of the membership function of A_j occurs at u_j and the midpoint of u_j is m_j, the forecasted value at the time j is m_j.
- Principle 2. If there is the fuzzy logical relationship group $A_i \to A_{j1}, A_{j2}, ..., A_{jk}$, where A_i is the fuzzy set of a year, assume k, then the fuzzy forecasted value is $A_{j1}, A_{j2}, ..., A_{jk}$. If $m_{j1}, m_{j2}, ..., m_{jk}$ are the midpoints of the intervals $u_{j1}, u_{j2}, ..., u_{jk}$ respectively, the crisp forecasted value of year $k+1$ is computed as the following formula
$$\mathrm{CFV}_{k+1} = \frac{m_{j1} + m_{j2} + ... + m_{jk}}{k}. \tag{1}$$
- Principle 3. If $A_i \to \emptyset$, the fuzzy forecasted value is A_i and the crisp forecasted value is m_i which is the midpoint of interval u_i.

3.3. The fuzzy time series forecasting model of Yu

In the Yu's model [17], a fuzzy set can be repeated in the right-hand side of a FLRG. Therefore, to resolve recurrent fuzzy logical relationships and reflect the different importance

among them, the fuzzy sets in the right-hand side of the FLRGs are assigned different weights in chronological order. In the forecasting and defuzzification step (step 5), the Principle 2 of the Chen's model is changed. Specifically, if there is a FLRG $A_i \to A_{j1}, A_{j2}, ..., A_{jk}$, where A_i is the fuzzy set of a year, assume k, then the fuzzy forecasted value is $A_{j1}, A_{j2}, ..., A_{jk}$. If $m_{j1}, m_{j2}, ..., m_{jk}$ are the midpoints of the intervals $u_{j1}, u_{j2}, ..., u_{jk}$, respectively, the crisp forecasted value of year $k + 1$ is computed as the following formula

$$\mathrm{CFV}_{k+1} = \frac{1 \times m_{j1} + 2 \times m_{j2} + ... + k \times m_{jk}}{1 + 2 + ... + k}. \tag{2}$$

3.4. The time-variant fuzzy time series forecasting model of Dieu

In [24], Dieu has proposed a time-variant fuzzy logical relationship groups which formed a new time-variant fuzzy time series forecasting model in such a way that the time-variant fuzzy logical relationship groups presented in Definition 4 were used in the forecasting model of Yu instead of the time-invariant fuzzy logical relationship groups presented in Definition 3. Then, he et al. applied various optimization algorithms and defuzzication techniques to improve the forecasting accuracy of their new models [20].

4. LINGUISTIC TIME SERIES FORECASTING MODEL

4.1. The linguistic time series forecasting model

The fuzzy time series concept is really attractive because the linguistic words with fuzzy set based semantics are used to solve the time series forecasting problems. Although the fuzzy time series forecasting models are originated from time series of linguistic data, there are no studies in this field that can directly handle linguistic data with their inherent semantics in the natural language. This prompted Hieu et al. to study and introduce the linguistic time series forecasting model [22]. The linguistic time series concept introduced in [22] is inspired by the fuzzy time series concept introduced by Song and Chissom in [2].

Definition 5. (Linguistic time series) [22]. Let $\mathcal{X}$ be a set of linguistic words in the natural language of a linguistic variable $\mathcal{X}$ defined on the universe of discourse $U_{\mathcal{X}}$ to describe its numerical quantities. Any series $L(t)$, $t = 0, 1, 2, ...$, where $L(t)$ is a collection of words of $\mathcal{X}$, is called a linguistic time series (LTS).

$L(t)$ is a finite subset because in practical applications only a few words of $\mathcal{X}$ are used. The concept of linguistic, logical relationship (LLR) defined from linguistic time series is similar to the fuzzy logical relationship concept defined in Definition 2. The LLR has the form $X_i \to X_j$, where X_i and X_j are the linguistic words of the linguistic variable $\mathcal{X}$ at the time t and $t+1$, respectively. The LLRs which have the same left-hand side are grouped into linguistic, logical relationship group (LLRG) of the form $X_i \to X_{j_1}, X_{j_2}, \ldots, X_{j_n}$.

In the formalism of HA [26, 27], the inherent semantics of a linguistic word x is quantified by three quantitative semantic aspects: the fuzziness measure $fm(x)$, the fuzziness interval $\Im(x)$ and the semantically quantifying mapping value (SQM-value) $f_{\mathcal{X}}(x)$. It is crucial that the qualitative semantics of the linguistic variable $\mathcal{X}$ must formally determine three quantitative semantic aspects of $\mathcal{X}$ described above. This means that HA-based formalism can establish a formal linkage between the meaning of linguistic data and their respective

quantitative semantics, allowing to provide a formal basis for handling directly linguistic data in LTS to solve the time series forecasting problems.

A LTS forecasting model developed based on the formalism of HA to solve the time series forecasting problems is described as follows (see [22]):

Step 1. Determine the universe of discourse of the linguistic variable $\mathcal{X}$, establish HA structure by selecting two generator words, the relative sign table of hedges, two fuzzy parameters $\theta = fm(c^-)$ and $\alpha = \mu(L)$ and an integer number to limit the maximum length of the declared linguistic words.

Step 2. Calculate the SQM-value $v(x)$ of the declared linguistic words.

Step 3. Transform the SQM-values of the linguistic words from the normalized universe $[0,\ 1]$ into the real numerical semantic value domain of the universe of discourse of the variable $\mathcal{X}$.

Step 4. Semantize the historical data. The semantics of each data point is determined based on the real numerical semantic value which is closest to the data point.

Step 5. Establish the LLRs and group them into the LLRGs.

Step 6. Calculate the forecasted values based on the established LLRGs and the crisp value calculation principles.

4.2. The proposed linguistic time series forecasting model based on linguistic forecasting rules

In the LTS forecasting model [22] presented above, the time-invariant LLRGs are used to induce the forecasting rules for calculating the crisp forecasted values. That is, the LLRs which have the same left-hand sides occurred after the time t are still grouped into the same group with the LLRs (also having the same left-hand sides) occurred before or at the time t.

In this paper, inspired by the time-variant FLRG concept proposed by Dieu [24], the *linguistic forecasting rule* (LFR) for a certain time t is induced from the LLRs in such a way that only the right-hand sides of the LLRs which have the same left-hand sides occurred before or at the current forecasting time t are put in chronological order into the right-hand side of the same LFR under consideration. For example, assume that we have the LLRs: *Little small* $\rightarrow$ *Little small* and *Little small* $\rightarrow$ *medium* occurred at the years 1977 and 1978, respectively. The LFRs for those years are *Little small* $\rightarrow$ *Little small* and *Little small* $\rightarrow$ *Little small*, *medium,* respectively. Note that, although those two LLRs have the same left-hand sides, the word *medium* in the right-hand side of the LLR occurred at the year 1978 is not put into the right-hand side of the LFR for the year 1977 because that LLR occurred after the year 1977. The procedure for generating a LFR for the current forecasting time t is as follows:

Step 1. Create a new LFR $\wp$ for the current forecasting time t whose both left-hand side and right-hand sides are empty.

Step 2. Add to the left-hand side and the right-hand side of $\wp$ the words in the left-hand side and the right-hand side of the LLR occurred at the time t.

Step 3. Find all LLRs which have the same left-hand side with $\wp$ occurred before the time t and then put the right-hand sides of them into the righ-hand side of $\wp$ in chronological order.

Besides, in the step of calculating the crisp forecasted value, we apply a more efficient weighted calculation formula to replace the one applied in [26, 27]. With the above ideas, we refine the forecasting procedure of the LTS in [22] into a new one (Step 2 and Step 3 of the model in [22] are included in Step 2) as hereafter:

Step 1. Determine the syntactic and qualitative semantics of the linguistic variable $\mathcal{X}$ by defining two generator words, the set of hedges H, the relative sign table of hedges and a positive integer λ determining the maximum length of the declared linguistic words. Determine the universe of discourse of $\mathcal{X}$ based on the historical data and determine the selected linguistic word set.

Step 2. *Quantify the semantics of the selected words.* In the formalism of HA, the quantitative semantics of the linguistic variables is determined by the values of two fuzziness parameters $\theta = fm(c^-)$ and $\alpha = \mu(L)$ (can be determined by the human experts or the trial-error method). When the values of θ and α are specified, the SQM-values of the declared word set are calculated. Then, the SQM-values of the word set selected to use are linearly transformed from the normalized universe [0, 1] into the real numerical semantic value domain of the universe of discourse of $\mathcal{X}$.

Step 3. *Semantize the historical data.* Transforms the given historical data into a linguistic time series in such a way that for each given timestamp of the historical data, choose a linguistic word from the selected word set so that its real numerical semantics is closest to the value of the historical data at that timestamp.

Step 4. *Establish the LLRs and the LFRs.* The LLRs are generated by scanning the obtained linguistic time series and then they are used to generate the LFRs. As described above, a LFR is of the form: $l_i \to l_{j_1}, l_{j_2}, \ldots, l_{j_n}$, where l_i is the word of time $t-1, l_{j_k} (1 \le k \le n)$ is the word in the right-hand side of the LLR which occurred before or at the time t and its left-hand side is l_i.

Step 5. *Calculate the crisp forecasted results.* Based on the crisp value calculation principles and the semantics of the LFRs, calculate the crisp forecasted values by using the real semantic value of the linguistic words and applying the weights in chronological order to the right-hand side of the LFRs.

5. APPLICATIONS OF THE PROPOSED TIME SERIES FORECASTING MODEL BASED ON LINGUISTIC FORECASTING RULES

In order to show the performance of the proposed approach, some experimental studies are executed to compare the performance of the proposed LTS forecasting model based on linguistic forecasting rules with the forecasting models examined by Chen [4], Yu [17] and Hieu [22] using the historical data of the enrollments of University of Alabama observed from 1971 to 1992, and Chen and Hwang [35] and Hieu [21] using the daily average temperature data observed from June 1996 to September 1996 in Taipei. The comparative studies aim at showing that the proposed forecasting models outperform their counterparts. Then, to show their realization in Viet Nam, they are applied to the forecasting problem of the historical data of the average rice production (thousand ton per year) of Viet Nam from 1990 to 2010.

5.1. Forecast the enrolments of the University of Alabama

Based on the proposed LTS forecasting model in the previous section, the steps of the forecasting procedure of the enrollments of University of Alabama are as follows:

Step 1. Choose two generator words $c^- = small(s)$ and $c^+ = large(l)$ and two hedges $Little(L) \in H^-$, $Very(V) \in H^+$. The declared linguistic words have their maximum length of 2 ($\lambda = 2$), so we have: $X_{(2)} = \{\boldsymbol{0}$, *Very small*, *small*, *Little small*, *medium*, *Little large*, *large*, *Very large*, $\boldsymbol{1}\}$, where $\boldsymbol{0}$ and $\boldsymbol{1}$ are the two constants with the smallest semantics (*Extremely small*) and the largest semantics (*Extremely large*), respectively. However, to ensure comparative meaning with the existing models, only 7 linguistic words are used to describe the universe of discourse, so the two constants $\boldsymbol{0}$ and $\boldsymbol{1}$ are not used and we have the set of selected linguistic words: $U_{\mathcal{X},\mathcal{L}} = \{$*Very small*, *small*, *Little small*, *medium*, *Little large*, *large*, *Very large*$\}$. The universe of discourse of linguistic variable $\mathcal{X}$ is $U_{\mathcal{X}} =$ [13000, 20000]. Put $U_{\mathcal{X}min} = 13000$ and $U_{\mathcal{X}max} = 20000$.

Step 2. *Quantify the semantics of the selected words.* Compute the SQM-values based on the fuzziness parameter values of the linguistic variable $\mathcal{X}$. The fuzziness parameter values include: $fm\,(c^-) = fm\,(small) = \theta$ and $(Little) = \alpha$ which are the fuzziness measure values of the generator word c^- and the negative hedge *Little*, respectively. The values of these two parameters can be determined by human's experts, or by trial and error method. In the experiments for this forecasting problem, the values of $fm(c^-) = 0.46$ and $(Little) = 0.52$ are chosen by human experts. The numerical semantics of the declared word set is determined by their SQM-values and linearly transformed to the real numerical semantic domain of the universe of discourse $U_{\mathcal{X}} =$ [13000, 20000] by the formula

$$v_R(l_i) = U_{\mathcal{X}min} + (U_{\mathcal{X}max} - U_{\mathcal{X}min}) \times v(l_i), \tag{3}$$

where $l_i \in X_{(2)}$, $v\,(l_i)$ is the SQM value of the word l_i, $U_{\mathcal{X}\min}$ and $U_{\mathcal{X}\max}$ are the lower bound and the upper bound of $U_{\mathcal{X}}$, respectively. Specifically, with the set $U_{\mathcal{X},L}$ and the values of $\theta = fm(c^-)$ and $\alpha = \mu(L)$ specified above we have the real numerical semantics the declared word set: $U_{\mathcal{X},R} = \{13742, 14546, 15416, 16220, 17163, 18186, 19129\}$, where each value of it is computed by the equation (3).

Step 3. Transform the historical data of the enrollments of the University of Alabama observed from 1971 to 1992 (the column "**Enrollment**" in the Table 5.1. into the linguistic words from the selected word set. For example, the enrollment data of the year of 1973 is 13867 which is closest to the real numerical semantics of 13742 of the word *Very small* in $U_{\mathcal{X},L}$. Hence, it is assigned the linguistic word *Very small*. In the same manner of the enrollment data for other years, all historical data are transformed to the linguistic time series and shown in the column "**Linguistic time series**" in Table 5.1..

Step 4. Scan the linguistic time series to generate the LLRs and the results are shown in the column "**Linguistic logical relationship**" in Table 5.1.. Generate LFRs from the LLRs and the results are shown in Table 5.1..

Step 5. Calculate the forecasted values based on the LFRs. The following two crisp value calculation principles are applied:

- Principle 1. If a LFR takes the form: $l_i \to l_{j1}, l_{j2}, \ldots, l_{jp}, p \geq 1$, where l_i is the linguistic word of a certain year, say year k, and $v_R\,(l_{j1})\,, v_R\,(l_{j2})\,, \ldots, v_R\,(l_{jp})$ are

Table 1: The historical data of the enrollments of the University of Alabama observed from 1971 to 1992 and the linguistic logical relationships.

Year	Enrollment	Linguistic time series	Linguistic logical relationship
1971	13055	*Very small*	
1972	13563	*Very small*	*Very small* $\rightarrow$ *Very small*
1973	13867	*Very small*	*Very small* $\rightarrow$ *Very small*
1974	14696	*small*	*Very small* $\rightarrow$ *small*
1975	15460	*Little small*	*small* $\rightarrow$ *Little small*
1976	15311	*Little small*	*Little small* $\rightarrow$ *Little small*
1977	15603	*Little small*	*Little small* $\rightarrow$ *Little small*
1978	15861	*medium*	*Little small* $\rightarrow$ *medium*
1979	16807	*Very large*	*medium* $\rightarrow$ *Very large*
1980	16919	*Very large*	*Very large* $\rightarrow$ *Very large*
1981	16388	*medium*	*Very large* $\rightarrow$ *medium*
1982	15433	*Little small*	*medium* $\rightarrow$ *Little small*
1983	15497	*Little small*	*Little small* $\rightarrow$ *Little small*
1984	15145	*Little small*	*Little small* $\rightarrow$ *Little small*
1985	15163	*Little small*	*Little small* $\rightarrow$ *Little small*
1986	15984	*medium*	*Little small* $\rightarrow$ *medium*
1987	16859	*Little large*	*medium* $\rightarrow$ *Little large*
1988	18150	*Large*	*Little large* $\rightarrow$ *large*
1989	18970	*Very large*	*Large* $\rightarrow$ *Very large*
1990	19328	*Very large*	*Very large* $\rightarrow$ *Very large*
1991	19337	*Very large*	*Very large* $\rightarrow$ *Very large*
1992	18876	*Very large*	*Very large* $\rightarrow$ *Very large*

the real numerical semantics of the words $l_{j1}, l_{j2}, \ldots, l_{jp}$, respectively, then the crisp forecasted value of the year $k+1$ is calculated by the formula

$$\mathrm{CFV}_{k+1} = \frac{1 \times v_R(l_{j1}) + 2 \times v_R(l_{j2}) + \ldots + p \times v_R(l_{jp})}{1 + 2 + \ldots + p}. \tag{4}$$

- Principle 2. If the linguistic word of year k is l_i and the right-hand side of the LFR is empty (no linguistic word exists) then the crisp forecasted value of the year $k+1$ is $v_R(l_i)$.

  The proposed linguistic time series forecasting model based on LFRs with the application of the formula (4) is denoted by **IV_LTS4**.

In order to show the proposed linguistic time series forecasting model based on LFRs more efficiently than the LTS model proposed in [22] (denoted by Hieu 2020), the crisp forecasted value calculation formulas should be the same. Therefore, the formula (4) of the proposed forecasting model is replaced by the following formula (5) to calculate the crisp forecasted values as in [21, 22] (denoted by **IV_LTS5**).

$$\mathrm{CFV}_{k+1} = \frac{v_R(l_{j1}) + v_R(l_{j2}) + \ldots + v_R(l_{jp})}{p}. \tag{5}$$

Table 2: The list of linguistic forecasting rules

Year	Rule	Linguistic Forecasting Rules
1972	Rule 1	*Very small* → *Very small*
1973	Rule 2	*Very small* → *Very small, Very small*
1974	Rule 3	*Very small* → *Very small, Very small, small*
1975	Rule 4	*Small* → *Little small*
1976	Rule 5	*Little small* → *Little small*
1977	Rule 6	*Little small* → *Little small, Little small*
1978	Rule 7	*Little small* → *Little small, Little small, medium*
1979	Rule 8	*medium* → *Little large*
1980	Rule 9	*Little large* → *Little large*
1981	Rule 10	*Little large* → *Little large, medium*
1982	Rule 11	*medium* → *Little large, Little small*
1983	Rule 12	*Little small* → *Little small, Little small, medium, Little small*
1984	Rule 13	*Little small* → *Little small, Little small, medium, Little small, Little small*
1985	Rule 14	*Little small* → *Little small, Little small, medium, Little small, Little small, Little small*
1986	Rule 15	*Little small* → *Little small, Little small, medium, Little small, Little small, Little small, medium*
1987	Rule 16	*medium* → *Little large, Little small, Little large*
1988	Rule 17	*Little large* → *Little large, medium, large*
1989	Rule 18	*large* → *Very large*
1990	Rule 19	*Very large* → *Very large*
1991	Rule 20	*Very large* → *Very large, Very large*
1992	Rule 21	*Very large* → *Very large, Very large, Very large*

Furthermore, to confirm that the proposed forecasting models are more efficient than the existing forecasting models, the forecasted results of the two proposed models **IV_LTS4** and **IV_LTS5** are compared with the ones of the forecasting models of Song [2] (**Song 1993a**), Chen [4] (**Chen 1994**) and Hieu [21, 22] (**Hieu 2020**) and shown in Table 5.1. and also visualized graphically in Figure 5.1.. The criterion used to evaluate the compared models is the mean square error (MSE), where F_i is the forecasted value and A_i is the actual value

$$MSE = \left(\frac{1}{N}\right)\sum_{i=1}^{N}(F_i - A_i)^2. \tag{6}$$

The analysis of the data in Table 5.1. shows that the forecasting model **IV_LTS4** applying the formula (4) has the MSE value of 106216, much better than the one of the forecasting model **IV_LTS5** applying the formula (5) has the MSE value of 154606. In comparison with some existing forecasting models, both the proposed models **IV_LTS4** and **IV_LTS5** have better MSE values than the ones of the models of **Hieu 2020**, **Song 1993a** and **Chen 1996** (106216 and 1546066 in comparison with 262326, 423027 and 407507, respectively). Therefore, we can state that handling the forecasting problem of the enrollments of University of

Alabama with the application of the linguistic time series forecasting model based on LFRs receives better forecasted results than the compared forecasting models, the one with the application of the formula (4) is better than the one with the application of the formula (5).

Table 3: The forecasted results of the forecasting models for the enrollments of University of Alabama observed from 1971 to 1992.

Year	Enrollment	Song 1993a	Chen 1996	Hieu 2020	IV_LTS5	IV_LTS4
1971	13055	-	-	-	-	-
1972	13563	14000	14000	14537	13742	13742
1973	13867	14000	14000	14537	13742	13742
1974	14696	14000	14000	14537	14010	14144
1975	15460	15500	15500	15534	15416	15416
1976	15311	16000	16000	15534	15416	15416
1977	15603	16000	16000	15534	15416	15416
1978	15861	16000	16000	16019	15684	15818
1979	16807	16000	16000	16019	17163	17163
1980	16919	16813	16833	17162	17163	17163
1981	16388	16813	16833	17162	16692	16535
1982	15433	16789	16833	16019	16290	15999
1983	15497	16000	16000	15534	15617	15657
1984	15145	16000	16000	15534	15577	15577
1985	15163	16000	16000	15534	15550	15531
1986	15984	16000	16000	15514	15646	15703
1987	16859	16000	16000	16019	16581	16581
1988	18150	16813	16833	17162	17190	17360
1989	18970	19000	19000	19217	19129	19129
1990	19328	19000	19000	19217	19129	19129
1991	19337	19000	19000	19217	19129	19129
1992	18876	-	19000	19217	19129	19129
MSE		**423027**	**407507**	**262326**	**154606**	**106216**
RMSE		**650.4**	**638.36**	**512.18**	**393.2**	**325.9**

5.2. Forecast the daily average temperature in Taipei

The above proposed models are applied to forecast the daily average temperature in Taipei from June to September 1996 (column **AV** in Table 5.2.. The temperature forecasted results of the models of **IV_LTS5** and **IV_LTS4** are compared together and compared with the one of Hieu in [21] and the one of Chen and Hwang in [35].

With the observed minimum and maximum temperature values of 23.3 and 31.6, respectively, we can set the value interval of the universe of discourse of the variable $\mathcal{X}$ to be $U_{\mathcal{X}}$ = [23, 32]. Two generator words are $c^- = cool$ (c) and $c^+ = hot$ (h) and two hedges are $Little$ $(L) \in H^-$ and $Very$ $(V) \in H^+$. The maximum length of the selected words are 2, so we have $X_{(2)}$= {*Very cool*, *cool*, *Little cool*, *normal*, *Little hot*, *hot*, *Very hot*}. Two values of the fuzziness parameters are chosen as $fm(c^-) = 0.52$ and $\mu(L) = 0.528$, so the real numerical semantics of the selected words are $U_{\mathcal{X},R}$ = {24.04, 25.2, 26.51, 27.68, 28.76,

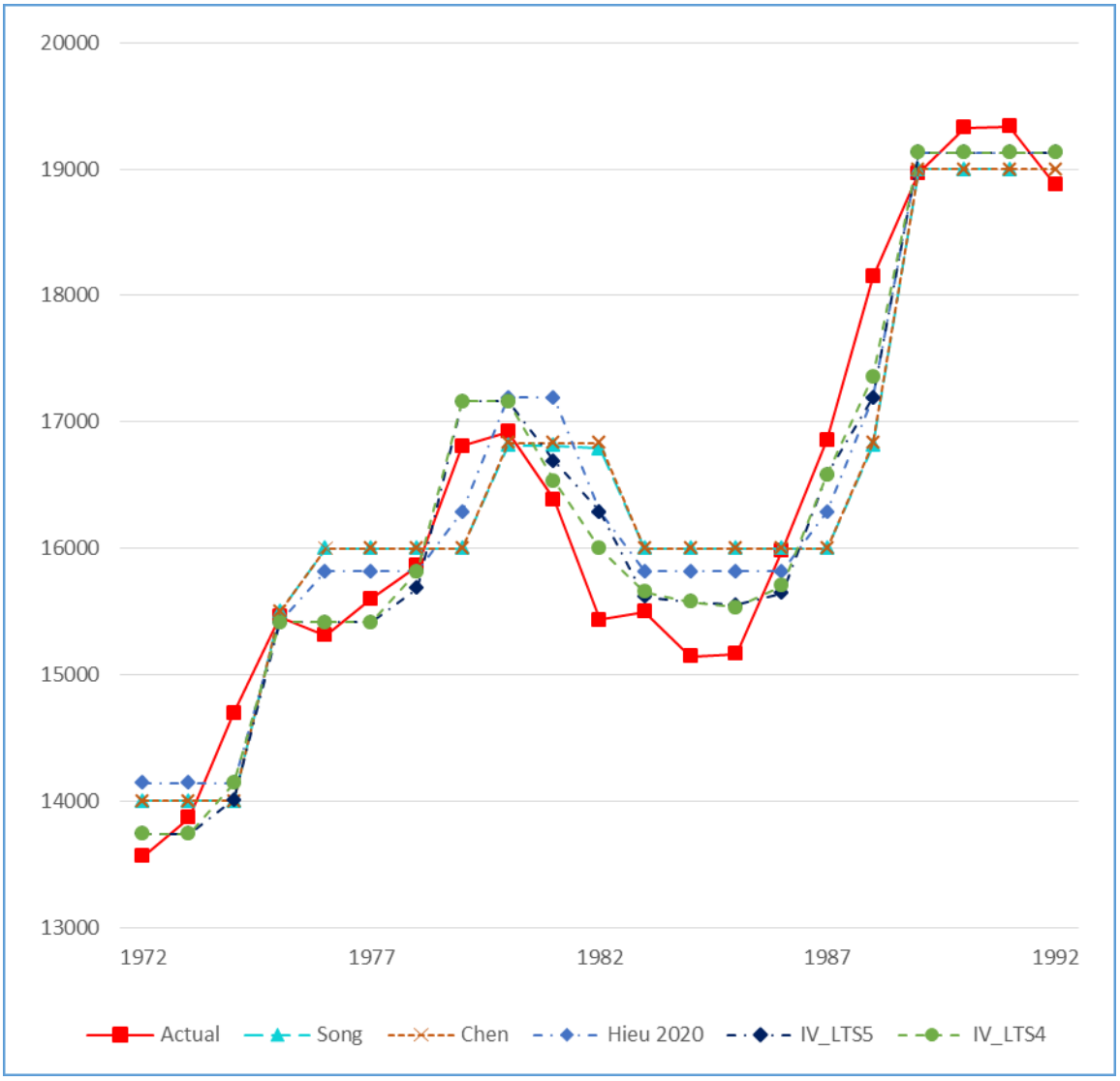

Figure 1: The comparison of the forecasted results of the enrollments of University of Alabama.

29.96, 31.04}}, where each value of it is computed by the equation (3). With those specified data, transform the observed historical data of the daily average temperature from June to September 1996 to the linguistic time series, establish the linguistic logical relationships and the linguistic forecasting rules. The temperature forecasted results of the models of **IV_LTS5** and **IV_LTS4** are shown in the columns **F5** and **F4** of Table 5.2., respectively.

The criterion used to evaluate the compared models are the mean absolute percentage error (MAPE), lower is better, as follows

$$\text{MAPE} = \frac{100\%}{N} \sum_i \left| \frac{F_i - A_i}{A_i} \right| \tag{7}$$

It is easy to calculate the MAPE values of the forecasting models based on the real

data and the forecasted data in Table 5.2.. The MAPE value of the model **IV_LTS4** is smaller than the one of the model **IV_LTS5** (2.36 in comparison with 2.57). Thus, for the daily average temperature forecasting problem, the forecasting model **IV_LTS4** is also more effective than the forecasting model **IV_LTS5**.

Table 4: The daily average temperature forecasted results of the models **IV_LTS5** and **IV_LTS4**.

Day	June			July			August			September		
	AV	F5	F4	AV	F5	F4	AV	F5	F4	AV	F5	F4
1	26.1	-	-	29.9	29.7	29.8	27.1	27.7	27.7	27.5	28.2	28.0
2	27.6	27.7	27.7	28.4	29.6	29.6	28.9	28.4	28.3	26.8	28.1	27.9
3	29.0	28.8	28.8	29.2	29.0	28.8	28.9	28.8	28.6	26.4	27.5	27.3
4	30.5	31.0	31.0	29.4	29.1	29.0	29.3	28.8	28.6	27.5	27.5	27.4
5	30.0	30.0	30.0	29.9	29.6	29.6	28.8	28.8	28.6	26.6	28.1	27.8
6	29.5	30.0	30.0	29.6	29.6	29.7	28.7	28.8	28.6	28.2	27.5	27.4
7	29.7	30.0	30.0	30.1	29.7	29.7	29.0	28.8	28.6	29.2	28.1	27.9
8	29.4	30.0	30.0	29.3	29.6	29.6	28.2	28.7	28.6	29.0	28.7	28.6
9	28.8	29.7	29.5	28.1	29.0	28.8	27.0	28.3	28.1	30.3	28.7	28.6
10	29.4	30.5	30.3	28.9	28.5	28.5	28.3	28.0	28.2	29.9	29.5	29.4
11	29.3	29.5	29.2	28.4	29.0	28.8	28.9	28.7	28.6	29.9	29.5	29.5
12	28.5	29.9	29.5	29.6	29.0	28.9	28.1	28.7	28.5	30.5	29.6	29.6
13	28.7	29.6	29.2	27.8	29.5	29.4	29.9	28.4	28.3	30.2	30.4	30.3
14	27.5	29.2	28.7	29.1	28.5	28.5	27.6	29.5	29.3	30.3	29.6	29.6
15	29.5	29.4	29.6	27.7	28.9	28.8	26.8	28.3	28.1	29.5	29.6	29.6
16	28.8	29.4	29.1	28.1	28.4	28.4	27.6	27.9	28.0	28.3	29.6	29.6
17	29.0	29.2	28.7	28.7	28.5	28.4	27.9	28.2	28.1	28.6	28.7	28.6
18	30.3	29.3	29.0	29.9	29.0	28.9	29.0	28.3	28.1	28.1	28.7	28.6
19	30.2	29.4	29.3	30.8	29.6	29.6	29.2	28.7	28.5	28.4	28.1	27.9
20	30.9	29.6	29.7	31.6	30.2	30.2	29.8	28.8	28.6	28.3	28.7	28.6
21	30.8	30.5	30.7	31.4	30.4	30.5	29.6	29.5	29.4	26.4	28.7	28.5
22	28.7	29.9	29.7	31.3	30.5	30.7	29.3	29.5	29.3	25.7	27.3	27.0
23	27.8	29.1	28.7	31.3	30.6	30.8	28.0	28.7	28.6	25.0	25.2	25.2
24	27.4	28.8	28.6	31.3	30.6	30.8	28.3	28.3	28.2	27.0	25.9	26.1
25	27.7	28.5	28.2	28.9	30.4	30.4	28.6	28.7	28.6	25.8	27.1	26.7
26	27.1	28.4	28.1	28.0	28.9	28.8	28.7	28.7	28.6	26.4	26.1	26.3
27	28.4	28.4	28.3	28.6	28.5	28.5	29.0	28.7	28.6	25.6	27.0	26.5
28	27.8	28.9	28.5	28.0	28.9	28.7	27.7	28.7	28.5	24.2	25.6	25.4
29	29.0	28.5	28.4	29.3	28.5	28.5	26.2	28.2	28.1	23.3	24.0	24.0
30	30.2	29.0	28.8	27.9	28.8	28.6	26.0	27.7	27.5	23.5	24.0	24.0
31				26.9	28.4	28.3	27.7	27.7	27.6		28.2	28.0

To compare with the daily average temperature forecasting models of Chen and Hwang [35] and Hieu [21], the historical data of the daily average temperature are divided by months for forecasting. In [35], Chen and Hwang applied many different algorithms with different window bases to evaluate the daily average temperature forecasting. We will compare the

results of our proposed models with the best one of them (denoted by Best of Chen's). When doing monthly forecasting, the selected word set $X_{(2)}$, the values of $fm(c^-)$ and $\mu(L)$ are kept unchanged. However, the minimum and maximum temperatures are different between months, so their universe of discourses are also different. The $U_{\mathcal{X}}$ of the months from June to September are [25.5, 31.5], [27.0, 32.0], [25.5, 31.0] and [23.0, 31.0], respectively, and the real numerical semantics of their word set are computed by formula (3). Hence, the LLRs are changed leading to the LLRGs and the LFRs are changed due to the change of the observed dataset. Intuitively seen in Figure 5.2., the MAPE value of **IV_LTS4** is the best, the second is of **IV_LTS5** and the third is of **Hieu 2020**. So, we can conclude that **IV_LTS4** is the best for forecasting the daily average temperature in comparison with the rest ones.

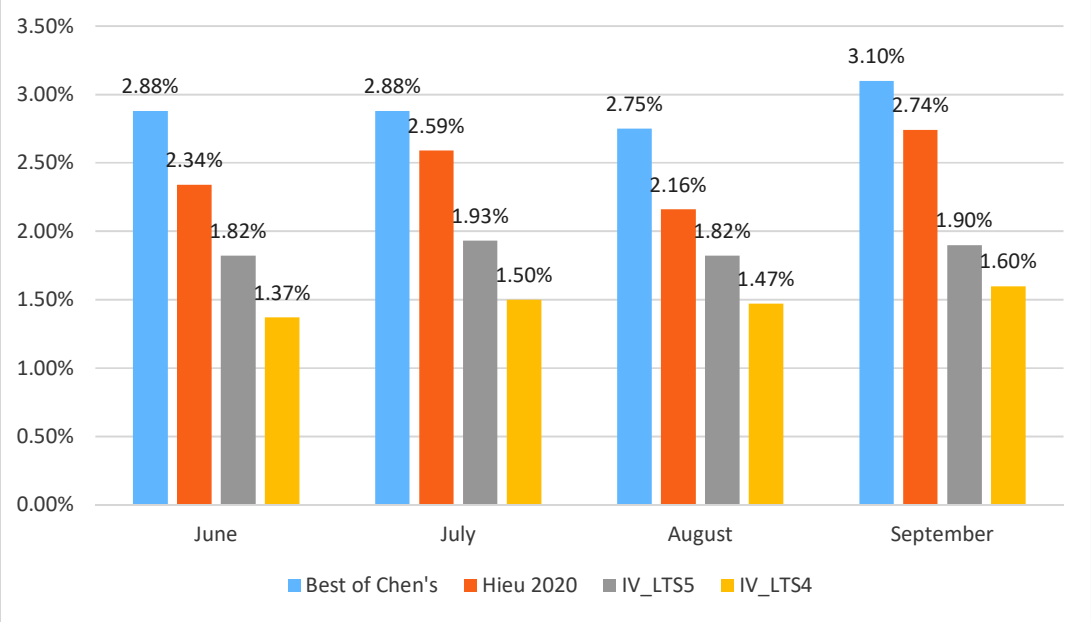

Figure 2: The comparison of the MAPE values of the daily average temperature forecasting by months among the forecasting models.

5.3. Forecast the average rice production of Viet Nam

As shown in the previous subsections, the proposed linguistic time series forecasting model is very efficient. In this subsection, we apply this model to a real dataset of the average rice production (thousand ton per year) of Viet Nam from 1990 to 2010 (shown in the column "Real values" in Table 5.3. and can be found on the Website of General Statistics Office of Vietnam (https://gso.gov.vn). The observed minimum and maximum production values are of 19225.1 and 39988.9, respectively, so the value interval of the universe of discourse of the variable $\mathcal{X}$ can be set to be $U_{\mathcal{X}} = [19000, 40000]$. Two generator words are $c^- = low(l)$ and $c^+ = high(h)$, two hedges are $Little(L) \in H^-$ and $Very(V) \in H^+$. We have $X_{(3)} = \{\boldsymbol{0}$, *Very Very low*, *Very low*, *Little Very low*, *low*, *Little Little low*, *Little low*, *Very Little low*, *medium*, *Very Little high*, *Little high*, *Little Little high*, *high*, *Little Very high*, *Very high*, *Very Very*

*high,**1***}. All words of $X_{(3)}$ are used, so $U_{\mathcal{X},L} = X_{(3)}$. Two fuzziness parameter values are chosen as $fm(c^-) = 0.52$ and $\mu(L) = 0.46$, so the real numerical semantics the declared word set $U_{\mathcal{X},R}$ = {19000, 20719.51, 22184.27, 23432.04, 24896.8, 26144.56, 27207.47, 28455.23, 29920.0, 31272.09, 32423.87, 33405.02, 34556.8, 35908.89, 37060.67, 38412.77, 40000.0}, where each value of it is computed by the equation (3). Based on those specified data, transform the observed historical data of the average rice production of Viet Nam from 1990 to 2010 to the LTS, establish the LLRs and the LFRs. The average rice production forecasted results of **IV_LTS4** model are shown in the column "**IV_LTS4**" of the Table 5.3.. The values of MSE, RMSE, ME and MAPE are also shown in the bottom of the Table 5.3., where ME is the mean error and its value is computed as

$$ME = (\frac{1}{N}) \sum_{i=1}^{N} |F_i - A_i|. \tag{8}$$

Table 5: The forecasted results of the proposed model for the average rice production of Viet Nam observed from 1990 to 2010.

Year	Real values	IV_LTS4	Year	Real values	IV_LTS4
1990	19225.1	-	2001	32108.4	32423.87
1991	19621.9	19000.0	2002	34447.2	33845.82
1992	21590.4	21122.85	2003	34568.8	34556.8
1993	22836.5	23432.04	2004	36148.9	35458.19
1994	23528.2	23432.04	2005	35832.9	35908.89
1995	24963.7	24408.55	2006	35849.5	35908.89
1996	26192.0	26144.56	2007	35942.7	35908.89
1997	27288.7	27207.47	2008	38729.8	36910.44
1998	28919.3	28455.23	2009	38950.2	38412.77
1999	31393.8	31272.09	2010	39988.9	39470.92
2000	32529.5	32423.87			
MSE				**317,184.4**	
RMSE				**563.19**	
ME				**391.0**	
MAPE				**1.294**	

It is easy to see that the value of ME of **IV_LTS4** is 391 and the value of MAPE is 1.294 which are good enough to realize the proposed forecasting model for this forecasting problem and it is a competitive forecasting model.

6. CONCLUSIONS

A new linguistic time series forecasting model based on linguistic forecasting rules which is enhanced from the linguistic time series forecasting model proposed by Hieu et al. by applying the linguistic forecasting rules instead of the linguistic, logical relationship groups is proposed in this paper. In addition, a new formula for calculating the crisp forecasted values is applied to improve the forecasted results. The enhancement of these linguistic time series forecasting models utilizing hedge algebras theory in comparison with the existing forecasting models is that the historical data is transformed into the linguistic time series based on

the real numerical semantics of the selected word set, defined by the SQM values, instead of partitioning the historical data into the intervals. This handling process is similar to the application users observing the given historical data in terms of their linguistic words. Therefore, the linguistic time series forecasting models are natural in general. The experimental studies made on two given historical data, the enrollment of University of the Alabama and the daily average temperature in Taipei have shown that the proposed forecasting models outperform their counterparts. Then, the realization of proposed forecasting model in Viet Nam is justified by applying to the forecasting problem of the average rice production of Viet Nam from 1990 to 2010. In this paper, the values of $fm(c^{-})$ and $\mu(Little)$ are chosen for the given forecasting problems by human experts. In next study, an optimization algorithm will be applied to get their optimal values to improve the forecasted results of the forecasting problems under consideration.

7. ACKNOWLEDGMENTS

This research is funded by University of Transport and Communications (UTC) under grant number T2021-CN-006.

The author thanks the reviewers for their valuable comments and suggestions to improve the paper quality.

REFERENCES

[1] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series - part i," *Fuzzy Set and Systems*, vol. 54, pp. 1–9, 1993.

[2] ——, "Fuzzy time series and its model," *Fuzzy Set and Systems*, vol. 54, pp. 269–277, 1993.

[3] ——, "Forecasting enrollments with fuzzy time series - part ii," *Fuzzy set and systems*, vol. 62, pp. 1–8, 1994.

[4] S.-M. Chen, "Forecasting enrollments based on fuzzy time series," *Fuzzy Set and Systems*, vol. 81, pp. 311–319, 1996.

[5] S.-M. Chen and C.-Y. Chung, "Forecasting enrollments of students by using fuzzy time series and genetic algorithms," *International Journal of Intelligent Systems*, vol. 17, no. 3, pp. 1–17, 2006a.

[6] Y.-L. Huang, Shi-JinnHorng, M. He, P. Fan, Tzong-WannKao, M. K. Khan, J.-L. Lai, and I.-H. Kuo, "A hybrid forecasting model for enrollments based on aggregated fuzzy time series and particle swarm optimization," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8014–8023, 2011.

[7] K. Huarng, "Effective lengths of intervals to improve forecasting in fuzzy time series," *Fuzzy Set and Systems*, vol. 123, no. 3, pp. 387–394, 2001b.

[8] I.-H. Kuo, S.-J. Horng, T.-W. Kao, T.-L. Lin, C.-L. Lee, and Y. Pan, "An improved method for forecasting enrolments based on fuzzy time series and particle swarm optimization," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6108–6117, 2009.

[9] I.-H. Kuo, Shi-JinnHorng, Y.-H. Chen, R.-S. Run, Tzong-WannKao, R.-J. Chen, J.-L. Lai, and T.-L. Lin, "Forecasting taifex based on fuzzy time series and particle swarm optimization," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1494–1502, 2010.

[10] L.-W. Lee, L.-H. Wang, and S.-M. Chen, "Temperature prediction and taifex forecasting based on fuzzy logical relationships and genetic algorithms," *Expert Systems with Applications*, vol. 33, no. 3, pp. 539–550, 2007.

[11] ——, "Temperature prediction and taifex forecasting based on high-order fuzzy logical relationships and genetic simulated annealing techniques," *Expert Systems with Applications*, vol. 34, no. 1, pp. 328–336, 2007.

[12] S.-M. Chen, "Forecasting enrollments based on high-order fuzzy time series," *Cybernetic and Systems*, vol. 33, pp. 1–16, 2002.

[13] S.-M. Chen and C.-Y. Chung, "Forecasting enrollments using high-order fuzzy time series and genetic algorithms," *International Journal of Intelligent Systems*, vol. 21, no. 5, pp. 485–501, 2006b.

[14] C.-H. Cheng, G.-W. Cheng, and J.-W. Wang, "Multi-attribute fuzzy time series method based on fuzzy clustering," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1235–1242, 2008.

[15] N.-Y. Wang and S.-M. Chen, "Temperature prediction and taifex forecasting based on automatic clustering techniques and two-factors high-order fuzzy time series," *Expert Systems with Applications*, vol. 36, no. 2, pp. 2143–2154, 2009.

[16] G. C. G. Shyi-Ming Chen, Xin-Yao Zou, "Fuzzy time series forecasting based on proportions of intervals and particle swarm optimization techniques," *Information Sciences*, vol. 500, pp. 127–139, 2019.

[17] H.-K. Yu, "Weighted fuzzy time series models for taiex forecasting," *Physica A: Statistical Mechanics and its Applications*, vol. 439, no. 3-4, pp. 609–624, 2005.

[18] C.-H. Wang and L.-C. Hsu, "Constructing and applying an improved fuzzy time series model: Taking the tourism industry for example," *Expert Systems with Applications*, vol. 34, no. 4, pp. 2732–2738, 2008.

[19] V. R. Uslu, E. Bas, U. Yolcu, and E. Egrioglu, "A fuzzy time series approach based on weights determined by the number of recurrences of fuzzy relations," *Swarm and Evolutionary Computation*, vol. 15, pp. 19–26, 2014.

[20] N. V. Tinh and N. C. Dieu, "A new hybrid fuzzy time series forecasting model based on combining fuzzy c-means clustering and particle swam optimization," *Journal of Computer Science and Cybernetics*, vol. 35, no. 3, pp. 267–292, 2019.

[21] N. D. Hieu, N. C. Ho, and V. N. Lan, "An efficient fuzzy time series forecasting model based on quantifying semantics of words," in *2020 RIVF International Conference on Computing and Communication Technologies (RIVF),*, Ho Chi Minh, Vietnam, 2020, pp. 1–6.

[22] ——, "Enrollment forecasting based on linguistic time series," *Journal of Computer Science and Cybernetics*, vol. 36, no. 2, pp. 119–137, 2020.

[23] N. C. Ho, N. C. Dieu, and V. N. Lan, "The application of hedge algebras in fuzzy time series forecasting," *Journal of Science and Technology*, vol. 54, no. 2, pp. 161–177, 2016.

[24] N. C. Dieu, "Fuzzy time-depending logical relationship groups in fuzzy time series models," *Journal of Science and Technology (In Vietnamese)*, vol. 52, no. 6, pp. 659–672, 2014.

[25] N. C. Ho and N. V. Long, "Fuzziness measure on complete hedges algebras and quantifying semantics of terms in linear hedge algebras," *Fuzzy Sets and Systems*, vol. 158, no. 4, pp. 452–471, 2007.

[26] N. C. Ho and W. Wechler, "Hedge algebras: an algebraic approach to structures of sets of linguistic domains of linguistic truth values," *Fuzzy Sets and Systems*, vol. 35, pp. 281–293, 1990.

[27] ——, "Extended algebra and their application to fuzzy logic," *Fuzzy Sets and Systems*, vol. 52, no. 3, pp. 259–281, 1992.

[28] N. C. Ho, W. Pedrycz, D. T. Long, and T. T. Son, "A genetic design of linguistic terms for fuzzy rule based classifiers," *International Journal of Approximate Reasoning*, vol. 54, no. 1, pp. 1–21, 2013.

[29] N. C. Ho, T. T. Son, and P. D. Phong, "Modeling of a semantics core of linguistic terms based on an extension of hedge algebra semantics and its application," *Knowledge-Based Systems*, vol. 67, pp. 244–262, 2014.

[30] N. C. Ho, H. V. Thong, and N. V. Long, "A discussion on interpretability of linguistic rule based systems and its application to solve regression problems," *Knowledge-Based Systems*, vol. 88, pp. 107–133, 2015.

[31] P. D. Phong, N. D. Du, N. T. Thuy, and H. V. Thong, "A hedge algebras based classification reasoning method with multi-granularity fuzzy partitioning," *Journal of Computer Science and Cybernetics*, vol. 35, no. 4, pp. 319–336, 2008.

[32] B. H. Le, L. T. Anh, and N. V. Binh, "Explicit formula of hedge-algebras-based fuzzy controller and applications in structural vibration control," *Applied Soft Computing*, vol. 60, pp. 150–166, 2017.

[33] N. H. Huy, N. C. Ho, and N. V. Quyen, "Multichannel image contrast enhancement based on linguistic rule-based intensificators," *Applied Soft Computing*, vol. 76, pp. 744–762, 2019.

[34] H. Tung, N. D. Thuan, and V. M. Loc, "The partitioning method based on hedge algebras for fuzzy time series forecasting," *Journal of Science and Technology*, vol. 54, no. 5, pp. 571–583, 2016.

[35] S.-M. Chen and J.-R. Hwang, "Temperature prediction using fuzzy time series," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 30, no. 2, pp. 263–275, 2000.

Received on January 22, 2021
Accepted on March 01, 2021