ORIGINAL PAPER

Long-term prediction of time series based on fuzzy time series and information granulation

Yunzhen Liu[1] · Lidong Wang[1]

Abstract

Time-series prediction involves forecasting future data by analyzing and modeling historical data. The prediction process involves analyzing and mining various features hidden in the data to predict future data. Compared with one-step forecasting, long-term forecasting is urgently needed, which contributes to capturing the overall picture of future trends and enables discovering prospective ranges and development patterns. This study presents a new long-term forecasting model named the TIG_FTS_SEL model, which is developed by integrating trend-based information granules (TIGs), fuzzy time series, and ensemble learning. First, a time series is converted into a series of equal-length trend-based information granules to capture the fluctuation range and trend information effectively. Then the trend-based information granules are fuzzified to form fuzzy time series, which contributes to realizing the long-term prediction at a high abstract level. Furthermore, different models are used to establish an ensemble long-term forecasting approach by introducing a selection strategy for individual models. The ensemble method performs the prediction tasks using part models with solid prediction performances while disregarding the remaining models. Finally, the developed model is verified by experiments on different time-series datasets. The results demonstrate the sound prediction performance and efficiency of the proposed model.

Keywords Fuzzy time series · Long-term prediction · Trend-based information granulation · Ensemble learning

1 Introduction

Time-series prediction involves forecasting future data by analyzing and modeling historical data, which has been extensively used in various fields, including economics (Kumar et al. 2022), medicine (Fang et al. 2022), industry (Liu et al. 2020), and other fields (Granata and Di Nunno 2021; Huang et al. 2022). Time-series prediction can be performed using different models, the most common models include classical models and machine learning methods.

The classical models, such as exponential smoothing (Brown 1959), AR (Box et al. 1976), MA (Box et al. 1976), ARIMA (Box et al. 1976), are based on statistical principles and are widely used for time-series analysis and forecasting. Implementing these models is relatively straightforward and can yield precise forecasts for particular types of time-series data. Machine learning methods, including the support vector machine (SVM) model (Vapnik 1995) and the artificial neural network (ANN) model (McCulloch and Pitts 1943), offer more flexible ways of handling complex patterns and relationships. However, both traditional statistical models and machine learning methods face challenges in handling uncertain historical data and can provide interpretable semantic results, particularly in long-term prediction tasks.

Fuzzy logic is a mathematical framework for addressing data uncertainty by allowing partial membership in multiple categories. Song and Chissom (1993b) introduce the fuzzy time-series model through the application of fuzzy logic, which aids in extracting rules from uncertain historical data. The classical fuzzy time-series model consists of several key steps: dividing the discourse into intervals,

✉ Lidong Wang
  ldwang@dlmu.edu.cn

  Yunzhen Liu
  1156269494@qq.com

[1] School of Science, Dalian Maritime University, Dalian 116026, China

defining the fuzzy sets, establishing fuzzy logic relationships, making inferences and defuzzifying the results.

Several advanced models for fuzzy time-series prediction incorporate the four fundamental steps of the fuzzy time-series framework. Chen (1996) introduces simple arithmetic operations to reduce computation overhead while implementing fuzzy time-series forecasting. Chen and Tanuwijaya (2011) propose a fuzzy time-series model that utilizes automatic clustering techniques to divide the universe of discourse into intervals of varying lengths. Iqbal et al. (2020) introduce an innovative method for fuzzy time-series prediction. This method combines fuzzy clustering and information granules and incorporates a weighted average approach to handle uncertainty in the data series. Intelligent optimization algorithms are used to obtain the optimal interval length, which improves the prediction accuracy of fuzzy time-series models (Chen and Chung 2006; Goyal and Bisht 2023; Chen et al. 2019; Chen and Phuong 2017; Zeng et al. 2019). Wang et al. (2013) present an advanced approach to enhance the prediction accuracy and interpretability of the model by utilizing fuzzy C-means clustering and information granules for determining unequal-length temporal intervals. Lu et al. (2015) construct information granules within the amplitude change space to divide intervals, considering the trend information and distribution of the time-series data.

Fuzzy logic relationships are extracted based on fuzzing discourse intervals to form the "If-Then" interpretable semantic rules. The fuzzy rule-based approach is a frequently used method for data modeling (Cheng et al. 2016; Askari and Montazerin 2015; Chen and Jian 2017; Gautam et al. 2018). The fuzzy rule model is usually combined with machine learning techniques to develop new uncertain data analysis methods. Huarng and Yu (2006) use back propagation neural networks to determine fuzzy relations, and then form a fuzzy time-series model. Subsequently, fuzzy relations are determined using a variety of artificial neural networks, which helps to improve model efficacy. These neural network models include feed-forward artificial neural network (FFANN) (Aladag et al. 2009), Pi-Sigma neural network (Bas et al. 2018), and generalized regression neural network (GRNN) (Panigrahi and Behera 2018). Panigrahi and Behera (2020) apply multiple methods, including long-short term memory (LSTM), support vector machines (SVM), and deep belief network (DBN) to determine fuzzy relations.

The majority of fuzzy time-series models primarily emphasize one-step forecasting, which strives to achieve higher accuracy at the numerical level. However, there is an increasing demand for long-term forecasts. The cumulative error may occur when the one-step prediction model directly attempts to make long-term predictions. Moreover, these models rely solely on a single model for determining fuzzy relations. It is widely acknowledged that prediction models invariably require parameters and data preprocessing, and different approaches can yield varying outcomes. Therefore, selecting a suitable model that can precisely forecast the results of most tasks is challenging. The combination of prediction models is rooted in the understanding that no single model can excel across all the data, but the fusion of multiple models has the potential power to yield an estimate closely aligned with the actual data. Thus, an ensemble approach that assigns weights to individual models is designed from different perspectives (Hao and Tian 2019; Song and Fu 2020; Kaushik et al. 2020), which can usually improve the prediction results compared to using a single model, particularly for weak learners. Several weighted determination methods are suggested to consider the different performance levels of various component models (Adhikari and Agrawal 2014; Maaliw et al. 2021). One of the weighting methods is based on an in-sample error weight scheme, where each weight is assigned inversely proportional to the corresponding in-sample error.

Zadeh (1979) introduces the concept of information granule, which is now considered a crucial foundation in granular computing. By granulating time-series data into information granules, the overall features of the time series within a specific period can be extracted to characterize the dynamic change process, rather than emphasizing precise values at a particular time point. The principle of justifiable granularity (Pedrycz and Vukovich 2001) is a guideline that should be adhered to when converting a numerical time series into a sequence of information granules, enabling the extraction of valuable information from the time series and facilitating the interpretability of the results. Most existing models focus mainly on amplitude information while neglecting the important trend information in time-series data, which is crucial for decision-making. For time series representation, the trend-based information granules (TIGs) developed by Guo et al. (2021) are more representative and informative, encompassing time-series amplitude and trend information. It offers a promising method for the long-term prediction of time series by treating abstract entities as a whole rather than numerical entities.

This study develops a new long-term forecasting model named TIG_FTS_SEL that combines TIGs, fuzzy time series, and ensemble learning. In the initial stage of the proposed approach, a given numerical time series is granulated into smaller units called granules. This allows for prediction at the granularity level, where each granule represents a fundamental unit that reflects time-series variation range and trend information. Then the fuzzy C-means clustering algorithm is applied to assign the semantic description for the time series features captured

by the information granules. Next, fuzzy relations are determined, and predictions are implemented using a variety of techniques, encompassing back propagation neural network (BPNN), SVM, LSTM, DBN, GRNN, and the fuzzy logic group (FLRG). Finally, an ensemble scheme based on weighted linear combination techniques is employed to integrate the forecasts derived from the individual models. This scheme selects models exhibiting superior performance by defining a predetermined number of models. The study's main contributions are outlined as follows:

- The proposed TIG_FTS_SEL model implements the prediction at the granularity level, which can reduce the cumulative error.
- The proposed TIG_FTS_SEL model adopts an ensemble approach to alleviate potential problems that may arise from relying solely on a single model, thus improving the prediction performance.

The organization of this study is as follows: an introduction to the theoretical foundations that underlie the construction of TIG and fuzzy time series is presented in Sect. 2. The entire process of the suggested long-term forecasting model is provided in Sect. 3. The experimental results of the proposed TIG_FTS_SEL model and other comparison models on seven datasets are exhibited in Sect. 4. The conclusions obtained are described in Sect. 5.

2 Trend-based information granulation and fuzzy time series

This section discusses the process of creating TIGs, as well as the concepts related to fuzzy time series.

2.1 Trend-based information granulation

Given a numerical time series $\{x_1, x_2, \ldots, x_n\}$, we granulate it by dividing it into q subsequences denoted by $S_1, S_2, \ldots, S_q$, and set the corresponding time-domain windows to $T_1, T_2, \ldots, T_q$. We illustrate the formation of TIGs with $S_i = \{x_{i1}, x_{i2}, \ldots, x_{in_i}\}$ as an example, where n_i denotes the size of S_i. Using Cramer's decomposition theorem as a guide, the sequence S_i can be represented as follows (Guo et al. 2021):

$$x_{it} = k_i t + c_i + u_{it}, \qquad t = 1, 2, \ldots, n_i \qquad (1)$$

where c_i is a constant that denotes the intercept of S_i, and k_i represents the slope of S_i, they can be estimated by applying the least squares estimation method. Then the interval information granule $\Omega_i^u = [a_i^u, b_i^u]$ is constructed on the residual error sequence $U_i = \{u_{i1}, u_{i2}, \ldots, u_{in_i}\}$ using the principle of justifiable granularity. The construction

process must fulfill two intuitive requirements (Pedrycz and Vukovich 2001), which are as follows:

1. Coverage: The interval information granule should contain as many data points as possible. The cardinality of Ω_i^u is considered a measure of its coverage, that is, $\mathrm{card}\{u_{it}|u_{it} \in \Omega_i^u\}$. In this study, an increasing function f_1 of this cardinality is used, and it can be expressed by $f_1(u) = \frac{u}{N}$, where N is the length of U_i.

2. Specificity: The length of an interval should be as specific as possible. The function of the size of the interval, i.e., $m(\Omega_i^u) = |b_i^u - a_i^u|$, is used as an indicator of specificity. More generally, this study considers a decreasing function f_2 of the length of the interval, which can be represented by $f_2(u) = 1 - \frac{u}{range}$, where $range = |\max(U_i) - \min(U_i)|$.

The principle of justifiable granularity emphasizes the need for a broad scope of coverage while maintaining a high level of specificity. However, specificity tends to decrease as coverage increases. To achieve a balance between these two conflicting aspects, we can rely on an indicator that considers the product of coverage and specificity, i.e., $f = f_1 \times f_2$. Using this indicator, the lower bound a_i^u and upper bound b_i^u of the information granule Ω_i^u can be determined independently as follows:

$$a_{i,\mathrm{opt}}^u = \max_{a_i^u \leq \mathrm{rep}(U_i)} V(a_i^u), \qquad (2)$$

$$b_{i,\mathrm{opt}}^u = \max_{b_i^u \geq \mathrm{rep}(U_i)} V(b_i^u), \qquad (3)$$

where:

$$V(a_i^u) = f_1\big(\mathrm{card}\{u_{it}|u_{it} \in [a_i^u, \mathrm{rep}(U_i)]\}\big) \times f_2\big(|\mathrm{rep}(U_i) - a_i^u|\big), \qquad (4)$$

$$V(b_i^u) = f_1\big(\mathrm{card}\{u_{it}|u_{it} \in [\mathrm{rep}(U_i), b_i^u]\}\big) \times f_2\big(|b_i^u - \mathrm{rep}(U_i)|\big), \qquad (5)$$

where $\mathrm{rep}(U_i)$ takes the mean value of the residual error sequence U_i. Following this process, an interval information granule $\Omega_i^u = [a_i^u, b_i^u]$ is constructed for U_i. Then the TIG can be represented as $G_i = \{\Omega_i, k_i\} = \{c_i + \Omega_i^u, k_i\} = \{[c_i + a_i^u, c_i + b_i^u], k_i\} = \{[a_i, b_i], k_i\}$.

2.2 Fuzzy time series

Let $U = \{u_1, u_2, \ldots, u_n\}$ be the universe of discourse. A fuzzy set A on U can be represented as $A = \{\mu_A(u_1), \mu_A(u_2), \cdots, \mu_A(u_n)\}$, where μ_A is the membership function of the fuzzy set A, $\mu_A : U \to [0, 1]$, $\mu_A(u_i)$ denotes the membership degree of u_i belonging to the fuzzy set A, and $1 \leq i \leq n$.

Definition 1 (Song and Chissom 1993b) Let $Y(t)$ be the universe of discourse and the fuzzy sets defined on it are

$f_i(t)$ $(i = 1, 2, \ldots)$. If $F(t)$ consists of $f_1(t), f_2(t), \ldots$, it is referred to as a fuzzy time series on $Y(t)$.

Definition 2 (Song and Chissom 1993b) Let $F(t) = A_i$ and $F(t - 1) = A_j$, where A_i and A_j are fuzzy sets. If $F(t)$ can be determined by $F(t - 1)$, the relationship is expressed by a fuzzy logical relationship $A_i \rightarrow A_j$.

Definition 3 (Song and Chissom 1993b) When $F(t)$ is determined by $F(t - 1), F(t - 2), \ldots, F(t - n)$, the relationship between them can be characterized as the n-th order fuzzy logical relationship $F(t - n), \ldots, F(t - 2), F(t - 1) \rightarrow F(t)$.

3 The fuzzy time-series model for long-term prediction

This section introduces a fuzzy time-series prediction model that utilizes TIGs, fuzzy time series, and ensemble learning for long-term forecasting. First, a numerical time series is transformed into a sequence of equal-length TIGs, and trend feature datasets, including the intercept dataset, the fluctuation range dataset, and the slope dataset, are constructed. Next, each trend feature dataset is fuzzified using the fuzzy C-means clustering algorithm. Then various machine learning methods are applied to the training dataset to determine fuzzy relations and calculate the prediction error of each model to select several methods with solid performance. The predicted values obtained by the selected models are integrated to produce the final predicted results of the test dataset. The framework of the proposed approach is presented in Fig. 1. For a given time series $x = \{x_1, x_2, \ldots, x_n\}$, the forecasting procedure includes the following steps:

Step 1: Granulate time series and construct trend feature datasets

1) Granulate a given numerical time series.

As discussed in Sect. 2.1, a specific time series x is converted into a collection of TIGs by employing a fixed time window size T, and a granular time series $G = \{G_1, G_2, \ldots, G_q\}$ is formed, where $G_i = \{[a_i, b_i], k_i\}$, and $q = n/T$.

2) Construct trend feature datasets.

Through the information granulation process, trend features are extracted from the original time series to construct trend feature datasets. Table 1 presents constructed trend feature datasets, including the intercept dataset, the fluctuation range dataset, and the slope dataset. The slope k_i, which is determined by Eq. (1), represents the changing trend of G_i. On the other hand, the intercept a_i characterizes the start level of the changing trend of G_i.

The fluctuation range can be calculated by $(b_i - a_i)$, indicating the fluctuation level of G_i.

Step 2: Fuzzify each trend feature dataset.

1) Clustering for each trend feature dataset.

Three trend feature datasets are clustered using the fuzzy C-means clustering algorithm, and clustering prototypes for each cluster are generated.

2) Divide the universe of discourse of each trend feature dataset into intervals of unequal length.

The clustering prototypes are sorted from smallest to largest for each trend feature dataset. Let V_i and V_{i+1} be two adjacent clustering prototypes, defining the lower bound interval_L_i and upper bound interval_U_i of the i-th interval as follows:

$$\text{interval\_U}_i = \frac{V_i + V_{i+1}}{2}, \tag{6}$$

$$\text{interval\_L}_{i+1} = \text{interval\_U}_i. \tag{7}$$

Equations (6) and (7) are not applicable for determining the lower bound interval_L_1 of the first interval and the upper bound interval_U_c of the last interval. Hence, these values are computed in the following form:

$$\text{interval\_L}_1 = V_1 - (\text{interval\_U}_1 - V_1), \tag{8}$$

$$\text{interval\_U}_c = V_c + (V_c - \text{interval\_L}_c), \tag{9}$$

where c is the number of clusters. Following the above-mentioned calculation procedure, the interval of unequal length is obtained. The midpoint interval_M_i of the i-th interval is determined based on its lower bound interval_L_i and upper bound interval_U_i:

$$\text{interval\_M}_i = \frac{\text{interval\_L}_i + \text{interval\_U}_i}{2}. \tag{10}$$

3) Define the fuzzy sets on the universe of discourse of each trend feature dataset.

Linguistic terms can be defined on a set of obtained intervals $u_1, u_2, \ldots, u_r$ and can be represented in the following form using fuzzy sets A_i:

$$A_i = f_{i1}/u_1 + f_{i2}/u_2 + \ldots + f_{ir}/u_r, \tag{11}$$

where $f_{ij} \in [0, 1]$ is the membership degree of the interval u_j belonging to the fuzzy set A_i, which is defined as follows:

$$f_{ij} = \begin{cases} 1 & j = i \\ 0.5 & j = i - 1 \ or; j = i + 1 \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Thus, each interval is associated with all fuzzy sets at different membership degrees. In this way, fuzzy sets on the intercept dataset, the fluctuation range dataset, and the slope dataset can be defined, respectively.
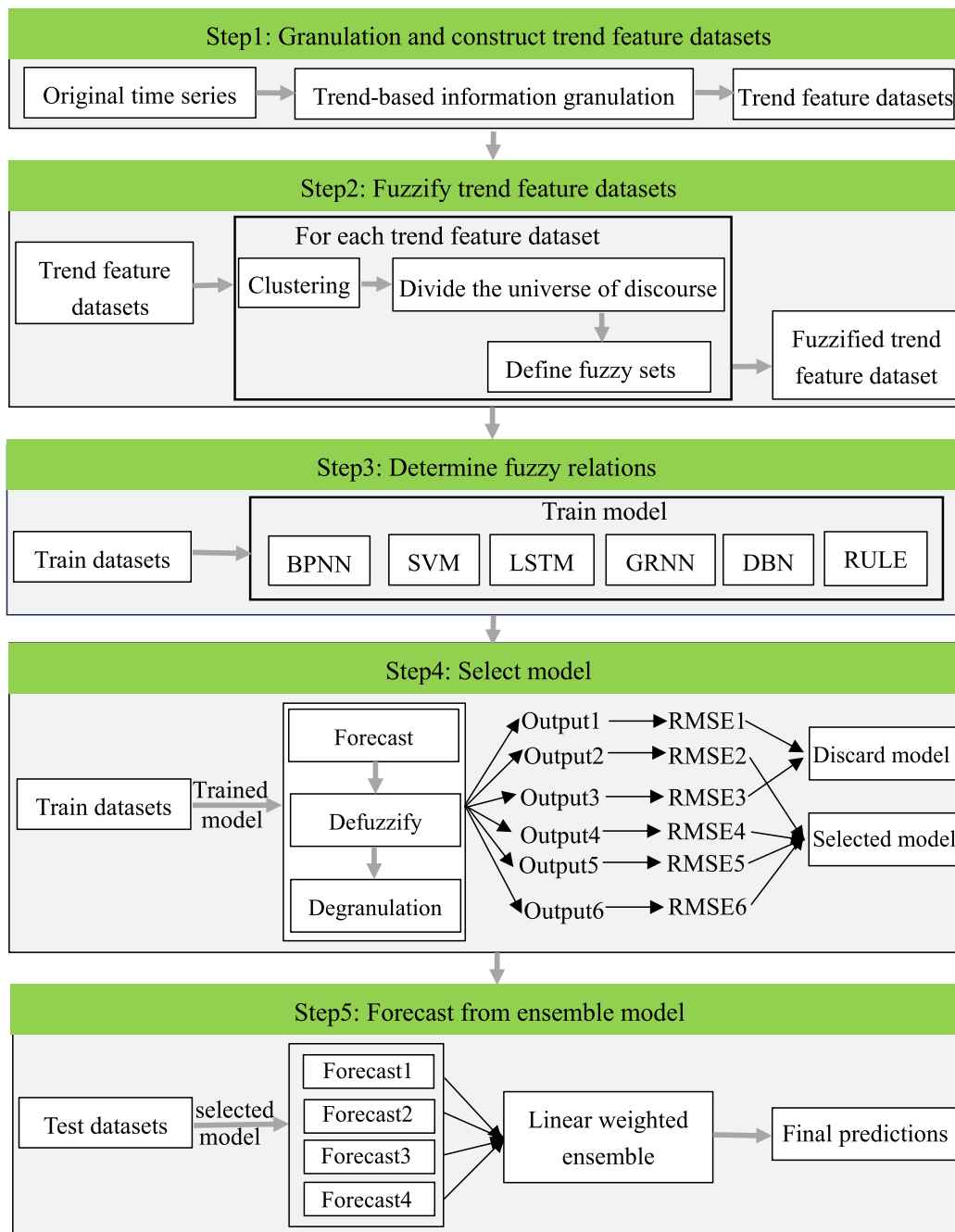
Fig. 1 The framework of the proposed TIG_FTS_SEL model

Table 1 Trend feature datasets

TIG	Intercept	Fluctuation range	Slope
G_1	a_1	$b_1 - a_1$	k_1
G_2	a_2	$b_1 - a_2$	k_2
G_3	a_3	$b_3 - a_3$	k_3
...	...	...	...
G_q	a_q	$b_q - a_q$	k_q

4) Fuzzify the element of each trend feature dataset.

If an element belongs to the interval u_i, it is fuzzified into the fuzzy set A_i. At this point, the fuzzification of parameters is achieved.

Step 3: Extract fuzzy relations.

After fuzzification of the three trend feature datasets, it is necessary to determine fuzzy relations. In this study, fuzzy relations are determined using BPNN (Rumelhart et al. 1986), SVM (Vapnik 1995), LSTM (Hochreiter and Schmidhuber 1997), DBN (Hinton et al. 2006), GRNN

(Specht et al. 1991), and FLRG (Song and Chissom 1993a), and the corresponding models are referred to as TIG_FTS_BPNN, TIG_FTS_SVM, TIG_FTS_LSTM, TIG_FTS_DBN, TIG_FTS_GRNN, and TIG_FTS_RULE, respectively. The index number i of the fuzzy set A_i is taken as the input and output of the models. For instance, consider three observations of $[A_5, A_4, A_1]$, where the input values of $[A_5, A_4]$ are 5 and 4, and the target value of A_1 is 1. In this step, min–max normalization is performed on the fuzzified data to constrain its values within the range between zero and one. The normalized value of an element in the fuzzified data can be calculated by dividing the difference between the element and the minimum value of the fuzzified data by the difference between the maximum and minimum values of the fuzzified data.

Step 4: Select individual models and defuzzfied.

The trained models are selected based on their performance on the training set. The predicted index number is obtained by de-normalizing the model's output, and it is subsequently rounded to the nearest integer. The midpoint of the interval that corresponds to the predicted index number is defined as the defuzzified prediction. When the index number exceeds the number of intervals, the midpoint of the final interval is employed as the defuzzified prediction. When the index number is less than one, the midpoint of the first interval is applied as the defuzzified prediction. The model's defuzzified prediction provides predictions of the trend variations, which are used to estimate the prediction of the actual data. Assuming that the trend features of the $(N + 1)$-th granule are predicted, this granule can be translated into specific predictions by:

$$\hat{y}_{N+1}\_L = k^*_{N+1} \times t + a^*_{N+1}, \tag{13}$$

$$\hat{y}_{N+1}\_U = \hat{y}_{N+1}\_L + (b - a)^*_{N+1}, \tag{14}$$

where $t \in [1, T]$, T represents the length of equal-length granulation, a^*_{N+1} denotes the predicted beginning level of the $(N + 1)$-th granule, $(b - a)^*_{N+1}$ indicates the predicted fluctuation range of the $(N + 1)$-th granule, k^*_{N+1} signifies the predicted changing trend of the $(N + 1)$-th granule, $\hat{y}_{N+1}\_L$ represents the lower bound of the prediction and $\hat{y}_{N+1}\_U$ represents the upper bound of the prediction. To evaluate the prediction performance, the final prediction is calculated as follows:

$$\hat{y}_{N+1} = (\hat{y}_{N+1}\_L + \hat{y}_{N+1}\_U)/2. \tag{15}$$

In this way, predictions on the training set can be obtained from individual model. To assess the efficacy of individual model on the training set, the root mean square error (RMSE) is utilized as an evaluation indicator. The models are then ranked according to their RMSE values, where lower values signify higher rankings. Based on these rankings, several models with high rankings are selected to make predictions on the test set.

Step 5: Ensemble predictions.

By combining the prediction results of several component models, the ensemble method improves the overall performance and mitigates the risk associated with model selection. A frequently employed ensemble technique based on a parallel strategy entails aggregating the prediction outcomes of multiple models that forecast time-series data. Let $y = [y_1, y_2, \ldots, y_n]$ represent the test data, and $\hat{y}^i = [\hat{y}^i_1, \hat{y}^i_2, \ldots, \hat{y}^i_n]$ denote the corresponding prediction outcome generated by the i-th model computed by Eqs. (13)–(15). The predictions of the models are linearly weighted as follows:

$$\hat{y}_k = w_1\hat{y}^1_k + w_2\hat{y}^2_k + \ldots + w_p\hat{y}^p_k = \sum_{u=1}^{p} w_u\hat{y}^u_k. \tag{16}$$

Here, p represents the number of selected models, $\hat{y}^u_k$ is the predicted output of each model for the k-th test data, and w_u represents the importance assigned to each model, which is calculated as follows:

$$w_i = \frac{w'_i}{\sum_{u=1}^{p} w'_u}, \tag{17}$$

$$w'_i = \frac{1}{\text{RMSE(i)}}, \tag{18}$$

where RMSE(i) represents the prediction error of the i-th model on the training set. To guarantee that models with relatively minor errors on the training set acquire more weight and vice versa, weight is assigned in inverse proportion to the models' errors. The proposed TIG_FTS_SEL model's predictions come true through the previously described process.

Step 6: Examine the performance of the TIG_FTS_SEL model.

We use the root mean square error (RMSE), mean absolute percentage error (MAPE), and mean absolute error (MAE) to examine the performance of the proposed TIG_FTS_SEL model, and they are calculated as follows:

$$RMSE = \sqrt{\frac{1}{L} \sum_{t=1}^{L} (\hat{y} - y_t)^2}, \tag{19}$$

$$MAPE = \frac{1}{L} \sum_{t=1}^{L} \frac{|\hat{y} - y_t|}{y_t} \times 100, \tag{20}$$

$$MAE = \frac{1}{L} \sum_{t=1}^{L} |\hat{y} - y_t|, \tag{21}$$

where L denotes the number of predicted data, and $\hat{y}$ and y_t represent the predicted and actual values, respectively.

4 Experiments

In this section, the experiments are conducted to examine the performance of the proposed TIG_FTS_SEL model. The experiments are performed on different time-series datasets, including Mackey–Glass time series, Melbourne temperature time series (MT time series), Zuerich monthly sunspot numbers, the minimum daily temperature of Cowichan Lake Forestry of British Columbia (daily temperature time series), standard's and Poor's 500 (stock index time series), monthly mean total sunspot time series, and historical levels of Lake Erie time series. In addition, the proposed model is compared with two types of models: numerical time-series prediction models (AR (Box et al. 1976), MA (Box et al. 1976), ARIMA (Box et al. 1976), NARnet (Benmouiza and Cheknane 2013), linear SVR (Hsia and Lin 2020)) and granular time-series models (LFIGFIS (Yang et al. 2017), IFIGFIS (Yang et al. 2017), TFIGFIS (Yang et al. 2017), Dong and Pedrycz's model (2008), Wang et al.'s model (2015), Feng et al.'s model (2021)). Furthermore, the proposed TIG_FTS_SEL model is compared with the individual component model and the traditional ensemble method based on the involved component models (TIG_FTS_EL). Based on empirical evidence, the performance of the ensemble model tends to reach a state of saturation after employing approximately four or five individual methods (Makridakis and Winkler 1983). This suggests that the prediction performance of the ensemble model remains relatively stable as the number of component models increases. In this work, we take four approaches from different component models and merge them, discarding the other approaches.

4.1 Experiment on Mackey–Glass time series

The following delay differential equation yields the Mackey–Glass time series, which is a classic description of chaotic systems:

$$\frac{Y(t)}{dt} = \frac{0.2Y(t-\tau)}{1 + Y^{10}(t-\tau)} - 0.1Y(t). \tag{22}$$

Let $Y(t) = 0$, $\tau = 1.7$, and a chaotic time series with 1201 values is obtained, as shown in Fig. 2. For the purpose of comparison, the time window length T is set to 13, which converts the time series into 92 information granules and a granular time series $G = \{G_1, G_2, \ldots, G_{92}\}$ is obtained. The experimental test set consists of the final 3 information granules, whereas the first 89 information granules are used as the training set. Further, the trend feature datasets are constructed following the procedure described in Sect. 3, as shown in Table 2. Using the fuzzy C-means algorithm, the unequal-length interval of each
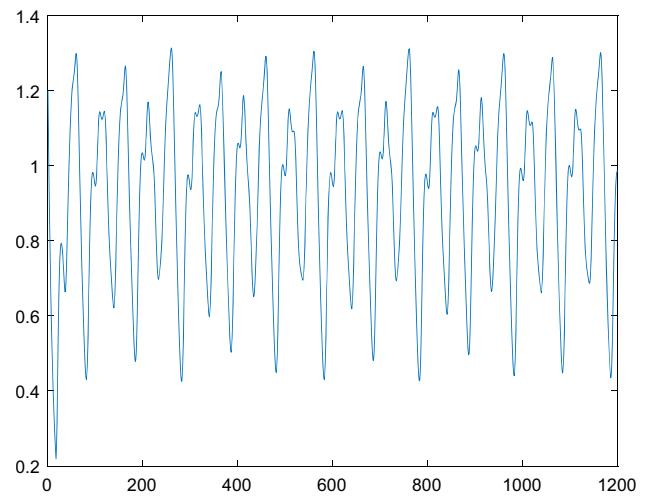

Fig. 2 Mackey–Glass time series

Table 2 Trend feature datasets of the Mackey–Glass time series

TIG	Intercept	Fluctuation range	Slope
G_1	1.1570	0.0662	−0.0686
G_2	0.0504	0.1402	0.0433
G_3	0.8127	0.0257	−0.0120
...	...	...	...
G_{89}	0.7317	0.0358	0.0373

trend feature dataset is obtained by Eqs. (6)–(9). Furthermore, the fuzzy sets A_i, B_i, and C_i ($i = 1, 2, \ldots, 13$) for the intercept dataset, the fluctuation range dataset, and the slope datasets are defined using Eqs. (11) and (12). At the same time, linguistic terms are assigned to every defined fuzzy set to describe each trend feature dataset. The corresponding linguistic descriptions for trend feature datasets are presented in Table 3. In this study, fuzzy relations are determined using BPNN, SVM, LSTM, GRNN, DBN, and FLRG, where the input and output of the models are the index number of the fuzzy set. To ensure the fairness of the comparison experiment, all models are assigned the same lag value of 2, indicating that they have the same order.

Table 3 Linguistic values

TIG	Intercept	Fluctuation range	Slope
G_1	A_{11}	B_{10}	C_1
G_2	A_1	B_{13}	C_{13}
G_3	A_5	B_3	C_5
...	...	...	...
G_{89}	A_4	B_4	C_{12}

Table 4 Ranking results of different models

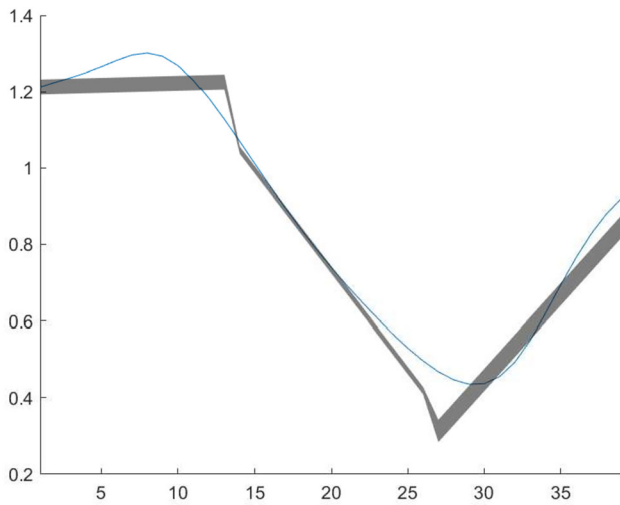| Method	Rank
TIG_FTS_BPNN	5
TIG_FTS_SVM	2
TIG_FTS_RULE	6
TIG_FTS_LSTM	3
TIG_FTS_GRNN	1
TIG_FTS_DBN	4

Fig. 3 Comparison between actual and predicted values

The fitting errors of the individual models on the training set are determined using the RMSE as an evaluation indicator. The models' performances on this dataset are ranked according to the fitting error, as presented in Table 4. The top four models that are selected for combination include the TIG_FTS_GRNN, TIG_FTS_SVM, TIG_FTS_LSTM, and TIG_FTS_DBN models.

Each of the four selected models is utilized for predicting the test data, and their prediction results are combined using a linearly weighted ensemble approach. For the prediction by a single model, TIG_ FTS_ GRNN is selected as an example to illustrate the prediction of test data x_{1159} ($1159 = 89 \times 13 + 2$). The TIG_FTS_GRNN model predicts the index numbers for each parameter of G_{90} as 12, 6, and 8. The obtained defuzzification results are 1.2345, 0.0438, and 0.0050, representing the trend characteristics of G_{90}. The predictions of the lower and upper bounds of x_{1159} are obtained as follows:

$$\hat{x}_{1159}\_L = 0.0050 \times 2 + 1.2345 = 1.2445,$$
$$\hat{x}_{1159}\_U = \hat{x}_{1159}\_L + 0.0438 = 1.2883.$$

To evaluate the performance of the model, the final prediction of the TIG_ FTS_ GRNN model for x_{1159} is calculated as follows:

$$\begin{aligned}\hat{x}_{1159}^{1} &= (\hat{x}_{1159}\_L + \hat{x}_{1159}\_U)/2 \\ &= (1.2445 + 1.2883)/2 \\ &= 1.2664.\end{aligned}$$

Similarly, the predictions $\hat{x}_{1159}^{2}$, $\hat{x}_{1159}^{3}$, and $\hat{x}_{1159}^{4}$ of x_{1159} are obtained using the TIG_FTS_SVM, TIG_ FTS_LSTM, and TIG_FTS_DBN models.

Finally, the weighted linear ensemble method is used to combine the four models, and weight is assigned to each model according to its performance on the training set, which is calculated as follows:

Table 5 RMSE, MAPE, and MAE comparisons for Mackey–Glass time series

Methods	RMSE	MAPE	MAE
LFIGFIS (Yang et al. 2017)	0.0622	6.1067	0.0516
IFIGFIS (Yang et al. 2017)	0.1132	10.5025	0.0882
TFIGFIS (Yang et al. 2017)	0.1241	12.3763	0.1001
NARnet (Benmouiza and Cheknane 2013)	0.2280	22.4482	0.1756
AR (Box et al. 1976)	0.2771	26.1058	0.2230
Linear SVR (Hsia and Lin 2020)	0.3717	38.5961	0.3269
TIG_FTS_BPNN	0.0969	10.8549	0.0780
TIG_FTS_SVM	0.0994	11.5768	0.0859
TIG_FTS_RULE	0.0974	9.4247	0.0725
TIG_FTS_LSTM	0.0850	10.1496	0.0703
TIG_FTS_GRNN	0.0880	10.1755	0.0696
TIG_FTS_DBN	0.0918	11.1766	0.0290
TIG_FTS_EL	0.0582	6.0966	0.0423
TIG_FTS_SEL	**0.0531**	**5.6749**	**0.0413**

Table 6 RMSE, MAPE, and MAE comparisons for MT time series

Methods	RMSE	MAPE	MAE
LFIGFIS (Yang et al. 2017)	**4.0175**	15.3606	3.1819
IFIGFIS (Yang et al. 2017)	4.4525	17.4581	3.4980
TFIGFIS (Yang et al. 2017)	4.5914	16.2509	3.4525
NARnet (Benmouiza and Cheknane 2013)	6.9910	36.2732	6.1815
AR (Box et al. 1976)	13.0316	53.7938	11.4057
Linear SVR (Hsia and Lin 2020)	19.5443	82.1745	18.3119
TIG_FTS_BPNN	4.2622	16.1335	3.1499
TIG_FTS_SVM	4.2293	16.1408	3.1144
TIG_FTS_RULE	5.7965	24.4047	4.5573
TIG_FTS_LSTM	4.4619	15.5374	3.1724
TIG_FTS_GRNN	4.2293	16.1408	3.1144
TIG_FTS_DBN	4.5424	15.3726	3.1785
TIG_FTS_EL	4.3560	**15.3599**	3.1101
TIG_FTS_SEL	4.2228	15.4663	**3.0507**

Table 7 RMSE, MAPE, and MAE comparisons for Zuerich monthly sunspot numbers time series

Methods	RMSE	MAPE	MAE
LFIGFIS (Yang et al. 2017)	33.7351	132.9506	25.9583
IFIGFIS (Yang et al. 2017)	40.6512	241.4182	32.0404
TFIGFIS (Yang et al. 2017)	43.8859	317.4153	35.8767
NARnet (Benmouiza and Cheknane 2013)	61.1764	305.4631	53.9146
AR (Box et al. 1976)	56.3509	177.5745	47.4319
Linear SVR (Hsia and Lin 2020)	91.8851	191.0395	74.6757
TIG_FTS_BPNN	32.5604	33.0844	26.6579
TIG_FTS_SVM	30.8467	38.1216	26.8845
TIG_FTS_RULE	45.2856	49.1418	36.0895
TIG_FTS_LSTM	30.8257	36.8834	26.8241
TIG_FTS_GRNN	30.4531	37.0047	26.4539
TIG_FTS_DBN	30.8257	36.8834	26.8241
TIG_FTS_EL	31.3345	**31.7139**	25.7735
TIG_FTS_SEL	**30.2872**	33.6894	**25.7340**

$$w = [w_1, w_2, w_3, w_4]$$

$$= \left[\frac{w'_1}{\sum_{i=1}^4 w'_i}, \frac{w'_2}{\sum_{i=1}^4 w'_i}, \frac{w'_3}{\sum_{i=1}^4 w'_i}, \frac{w'_4}{\sum_{i=1}^4 w'_i} \right]$$

$$= [0.3034, 0.2459, 0.2335, 0.2172],$$

where

$$w' = [w'_1, w'_2, w'_3, w'_4]$$

$$= \left[\frac{1}{0.0693}, \frac{1}{0.0856}, \frac{1}{0.0901}, \frac{1}{0.0969} \right]$$

$$= [14.4300, 11.6822, 11.0988, 10.3199].$$

The final prediction $\hat{x}_{1159}$ of x_{1159} obtained by the ensemble method is calculated as follows:

$$\hat{y}_{1159} = w_1 \hat{x}_{1159}^1 + w_2 \hat{x}_{1159}^2 + w_3 \hat{x}_{1159}^3 + w_4 \hat{x}_{1159}^4 = 1.2136.$$

Predictions of the other test data can be obtained by conducting the aforementioned procedures. The comparison of the prediction results of the TIG_FTS_SEL model and the actual data is presented in Fig. 3, where it can be seen that the trend of the predicted values aligns with that of the actual data. In addition, the TIG_FTS_SEL model is compared with the traditional linear ensemble model involving all component models (TIG_FTS_EL), granular models, and numerical models, as well as component models, including TIG_FTS_BPNN, TIG_FTS_SVM, TIG_FTS_LSTM, TIG_FTS_GRNN, TIG_FTS_DBN, and TIG_FTS_RULE. The models' performances are evaluated using RMSE, MAPE, and MAE. In Table 5, the prediction performances of the TIG_FTS_SEL model and comparison models on this dataset are presented. It is generally

Table 8 RMSE comparisons for daily temperature time series

Methods	RMSE
AR (Box et al. 1976)	5.081
Dong and Pedrycz's model (1) (2008)	3.584
Dong and Pedrycz's model (2) (2008)	3.473
Wang et al.'s model (2015)	3.469
TIG_FTS_BPNN	3.427
TIG_FTS_SVM	2.846
TIG_FTS_RULE	**2.348**
TIG_FTS_LSTM	3.045
TIG_FTS_GRNN	3.289
TIG_FTS_DBN	3.289
TIG_FTS_EL	2.963
TIG_FTS_SEL	2.844

Table 9 MAPE comparisons for stock index time series

Methods	MAPE
ABFO based model (Majhi et al. 2009)	2.409
BFO based model (Majhi et al. 2009)	2.531
PSO based model (Majhi et al. 2009)	2.385
GA based model (Majhi et al. 2009)	2.617
Wang et al.'s model (2015)	2.437
TIG_FTS_BPNN	3.699
TIG_FTS_SVM	1.903
TIG_FTS_RULE	4.280
TIG_FTS_LSTM	2.939
TIG_FTS_GRNN	3.976
TIG_FTS_DBN	1.695
TIG_FTS_EL	0.824
TIG_FTS_SEL	**0.714**

Table 11 RMSE comparisons for historical levels of Lake Erie time series

Methods	RMSE
AR (Box et al. 1976)	0.3191
MA (Box et al.1976)	0.3346
ARIMA (Box et al.1976)	0.2727
NARnet (Benmouiza and Cheknane 2013)	0.3006
Feng et al.'s model (2021)	0.2591
TIG_FTS_BPNN	0.2212
TIG_FTS_SVM	0.2125
TIG_FTS_RULE	0.2255
TIG_FTS_LSTM	0.2087
TIG_FTS_GRNN	0.2706
TIG_FTS_DBN	0.2206
TIG_FTS_EL	0.2004
TIG_FTS_SEL	**0.1999**

Table 10 RMSE comparisons for monthly mean total sunspot time series

Methods	RMSE
AR (Box et al. 1976)	63.6328
MA (Box et al. 1976)	68.4253
ARIMA (Box et al. 1976)	99.3976
NARnet (Box et al. 2013)	68.6114
Feng et al.'s model (2021)	52.0902
TIG_FTS_BPNN	29.2143
TIG_FTS_SVM	20.9874
TIG_FTS_RULE	16.7313
TIG_FTS_LSTM	22.7097
TIG_FTS_GRNN	22.1148
TIG_FTS_DBN	20.9873
TIG_FTS_EL	16.5135
TIG_FTS_SEL	**15.7458**

believed that the smaller the values of RMSE, MAPE, and MAE are, the better the performance of the prediction model is. Based on the results presented in Table 5, the TIG_FTS_SEL model exhibits the lowest RMSE, MAPE, and MAE values among all the models. This suggests that the TIG_FTS_SEL model outperforms the other models in terms of prediction effectiveness. In general, both the proposed selection-based ensemble model (TIG_FTS_-SEL) and the traditional ensemble model based on the involved component models (TIG_FTS_EL) exhibit superior performance compared to the individual models. In addition, the proposed TIG_FTS_SEL model contributes to improving the prediction performance of the TIG_FTS_EL model. This can be explained by the fact that not all models yield accurate predictions for a specific time series. Thus,

the ensemble long-term forecasting approach with a selection strategy might make the model perform better in making predictions.

4.2 Experiment on MT time series

The dataset utilized in this experiment comprises a temporal sequence of maximum daily temperatures in Melbourne, Australia, spanning from 1981 to 1990, including 3650 data points. In this experiment, the time-series data are divided into 40 time-domain windows, each of which

contains 91 data points. The training set covers the previous 38 temporal windows for prediction purposes, while the last 2 are reserved for testing. The model order is fixed at two throughout the experiment. Table 6 shows the prediction errors in terms of RMSE, MAPE, and MAE obtained by the TIG_FTS_SEL model and comparative models on the test samples. As shown in Table 6, the TIG_FTS_SEL model outperforms all component models except for the component model TIG_FTS_DBN, which has a lower MAPE. Compared to the related numerical and granular models, the TIG_FTS_SEL model exhibits superior performance in terms of MAE and is second only to the LFIGFIS model in terms of RMSE and MAPE. Although the TIG_FTS_EL model demonstrates sound prediction performance in terms of MAPE, the TIG_FTS_SEL model outperforms it when evaluating the results using RMSE and MAE.

4.3 Experiment on Zuerich monthly sunspot numbers time series

The Zuerich monthly sunspot numbers time series covers 2820 data points spanning from 1749 to 1983. Setting the time-domain window size to a fixed length of 33 results in the generation of 84 information granules. The initial 79 information granules are utilized as the training set for predicting the subsequent 2 information granules. Table 7 presents the evaluation indicator values of the TIG_FTS_SEL model and the other comparison models. The table shows that the TIG_FTS_SEL model has the smallest evaluation indicator values, except when MAPE is used as an evaluation indicator. In that case, the TIG_FTS_SEL model is second to the traditional ensemble model based on all component models (TIG_FTS_EL) and the component model TIG_FTS_BPNN. This suggests that the TIG_FTS_SEL model has the highest prediction performance.

4.4 Experimental summary

Supplementary experiments are conducted on four distinct time-series datasets, including daily temperature time series (minimum daily temperature of Cowichan Lake Forestry of British Columbia recorded from April 1, 1979 to May 30, 1996), stock index time series (Standard's and Poor's 500 (S &P 500) time series from January 3, 1994 to October 23, 2006), monthly mean total sunspot time series (spanning from January 1749 to December 2019), and historical levels of Lake Erie time series (spanning from January 1860 to September 2016). The performances of the TIG_FTS_SEL model, along with those of the other comparative models, in predicting these datasets are shown in Tables 8, 9, 10 and 11. Regarding the RMSE and MAPE

values on these datasets, both the proposed method and the conventional method within the ensemble framework exhibit superior performance compared to their corresponding component models. Also, the proposed TIG_FTS_SEL method, based on the ranking selection model, enhances the prediction performance of the traditional ensemble method (TIG_FTS_EL) that relies on all component models. In addition, compared to the existing granular models and numerical models, the proposed TIG_FTS_SEL model achieves sound evaluation indicator values.

5 Conclusion

This study proposes a long-term forecasting method named TIG_FTS_SEL, which is based on trend information granules, fuzzy time series, and ensemble learning. In the proposed method, the time series is initially granulated to extract the valuable information inherent in the original time series effectively, enabling prediction at the granular level and reducing accumulative errors. Then the trend features captured by information granules are used to construct the trend feature datasets, which are further fuzzified by applying the fuzzy C-means clustering algorithm to enable linguistic descriptions of these features. In addition, the proposed model uses different methods to determine fuzzy relations, and constructs an ensemble model for predictive purposes. Instead of merging all models, the ensemble model just incorporates those that outperform on the training set. Generally, this method exhibits superior performance on datasets compared to its component models and the ensemble method based on all component models, mitigating the potential drawbacks of relying on a single model. Through comparison experiments with other granular models and numerical models on seven available time-series datasets, the validity of the proposed model is confirmed.

Author Contributions YL: conceptualization, methodology, investigation, formal analysis, data collection, software, validation, writing—original draft. LW: supervision, methodology, writing—review and editing.

Data availability The datasets analyzed in this study are publicly available datasets utilized in previous literature, including Mackey-Glass time series (https://ww2.mathworks.cn/help/fuzzy/predict-chaotic-time-series-code.html;jsessionid=5f211830b69e40ff46d8380efa3e), MT time series (https://github.com/FinYang/tsdl), Zuerich monthly sunspot numbers (https://github.com/FinYang/tsdl), Daily temperature time series (https://geographic.org/globalweather/britishcolumbia/cowichanlakeforestry040.html), Stock index time series (https://finance.yahoo.com/quote/), Monthly mean total sunspot time

series (http://www.sidc.be/silso/datafiles), and historical levels of Lake Erie time series (http://www.glisaclimate.org/projects/484/page/2536).

Declarations

Conflict of interest No potential conflict of interest was reported by the author.

References

Adhikari R, Agrawal R (2014) Performance evaluation of weights selection schemes for linear combination of multiple forecasts. Artif Intell Rev 42:529–548

Aladag CH, Basaran MA, Egrioglu E et al (2009) Forecasting in high order fuzzy times series by using neural networks to define fuzzy relations. Expert Syst Appl 36(3):4228–4231

Askari S, Montazerin N (2015) A high-order multi-variable fuzzy time series forecasting algorithm based on fuzzy clustering. Expert Syst Appl 42(4):2121–2135

Bas E, Grosan C, Egrioglu E, Yolcu U (2018) High order fuzzy time series method based on pi-sigma neural network. Eng Appl Artif Intell 72:350–356

Benmouiza K, Cheknane A (2013) Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models. Energy Conv Manag 75:561–569

Box GE, Jenkins GM, Reinsel GC, Ljung GM (1976) Time series analysis: forecasting and control. Holden Bay, San Francisco

Brown RG (1959) Statistical forecasting for inventory control. McGraw-Hill, New York

Chen SM (1996) Forecasting enrollments based on fuzzy time series. Fuzzy Sets Syst 81(3):311–319

Chen SM, Chung NY (2006) Forecasting enrollments using high-order fuzzy time series and genetic algorithms. Int J Intell Syst 21(5):485–501

Chen SM, Jian WS (2017) Fuzzy forecasting based on two-factors second-order fuzzy-trend logical relationship groups, similarity measures and PSO techniques. Inf Sci 391:65–79

Chen SM, Phuong BDH (2017) Fuzzy time series forecasting based on optimal partitions of intervals and optimal weighting vectors. Knowl-Based Syst 118:204–216

Chen SM, Tanuwijaya K (2011) Multivariate fuzzy forecasting based on fuzzy time series and automatic clustering techniques. Expert Syst Appl 38(8):10594–10605

Chen SM, Zou XY, Gunawan GC (2019) Fuzzy time series forecasting based on proportions of intervals and particle swarm optimization techniques. Inf Sci 500:127–139

Cheng S, Chen S, Jian W (2016) Fuzzy time series forecasting based on fuzzy logical relationships and similarity measures. Inf Sci 327:272–287

Dong R, Pedrycz W (2008) A granular time series approach to long-term forecasting and trend forecasting. Physica A 387(13):3253–3270

Fang Z, Yang S, Lv C et al (2022) Application of a data-driven XGBoost model for the prediction of COVID-19 in the USA: a time-series study. BMJ Open 12(7):e056685

Feng G, Zhang L, Yang J, Lu W (2021) Long-term prediction of time series using fuzzy cognitive maps. Eng Appl Artif Intell 102:104274

Gautam SS, Abhishekh Singh S (2018) A new high-order approach for forecasting fuzzy time series data. Int J Comput Intell Appl 17(04):1850019

Goyal G, Bisht DC (2023) Adaptive hybrid fuzzy time series forecasting technique based on particle swarm optimization. Granul Comput 8(2):373–390

Granata F, Di Nunno F (2021) Forecasting evapotranspiration in different climates using ensembles of recurrent neural networks. Agric Water Manage 255:107040

Guo H, Wang L, Liu X, Pedrycz W (2021) Trend-based granular representation of time series and its application in clustering. IEEE T Cybern 52(9):9101–9110

Hao Y, Tian C (2019) A novel two-stage forecasting model based on error factor and ensemble method for multi-step wind power forecasting. Appl Energy 238:368–383

Hinton GE, Osindero S, Teh YW (2006) A fast learning algorithm for deep belief nets. Neural Comput 18(7):1527–1554

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780

Hsia JY, Lin CJ (2020) Parameter selection for linear support vector regression. IEEE Trans Neural Netw Learn Syst 31(12):5639–5644

Huang H, Chen J, Sun R, Wang S (2022) Short-term traffic prediction based on time series decomposition. Physica A 585:126441

Huarng K, Yu TH (2006) The application of neural networks to forecast fuzzy time series. Physica A 363(2):481–491

Iqbal S, Zhang C, Arif M et al (2020) A new fuzzy time series forecasting method based on clustering and weighted average approach. J Intell Fuzzy Syst 38(5):6089–6098

Kaushik S, Choudhury A, Sheron PK et al (2020) AI in healthcare: time-series forecasting using statistical, neural, and ensemble architectures. Front Big Data 3:4

Kumar G, Singh UP, Jain S (2022) An adaptive particle swarm optimization-based hybrid long short-term memory model for stock price time series forecasting. Soft Comput 26(22):12115–12135

Liu T, Wei H, Liu S, Zhang K (2020) Industrial time series forecasting based on improved gaussian process regression. Soft Comput 24:15853–15869

Lu W, Chen X, Pedrycz W et al (2015) Using interval information granules to improve forecasting in fuzzy time series. Int J Approx Reasoning 57:1–18

Maaliw RR, Ballera MA, Mabunga ZP, et al. (2021) An ensemble machine learning approach for time series forecasting of COVID-19 cases. In: 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), IEEE, pp 0633–0640

Majhi R, Panda G, Majhi B, Sahoo G (2009) Efficient prediction of stock market indices using adaptive bacterial foraging optimization (ABFO) and BFO based techniques. Expert Syst Appl 36(6):10097–10104

Makridakis S, Winkler RL (1983) Averages of forecasts: Some empirical results. Manage Sci 29(9):987–996

McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 5:115–133

Panigrahi S, Behera H (2018) A computationally efficient method for high order fuzzy time series forecasting. J Theor Appl Inf Technol 96:7215–7226

Panigrahi S, Behera HS (2020) A study on leading machine learning techniques for high order fuzzy time series forecasting. Eng Appl Artif Intell 87:103245

Pedrycz W, Vukovich G (2001) Abstraction and specialization of information granules. IEEE Trans Syst Man Cybern Part B-Cybern 31(1):106–111

Rumelhart DE, Hinton GE, McClelland JL et al (1986) A general framework for parallel distributed processing. Parallel Distrib Process Explor Microstruct Cogn 1(45–76):26

Song C, Fu X (2020) Research on different weight combination in air quality forecasting models. J Clean Prod 261:121169

Song Q, Chissom BS (1993) Forecasting enrollments with fuzzy time series—Part I. Fuzzy Sets Syst 54(1):1–9

Song Q, Chissom BS (1993) Fuzzy time series and its models. Fuzzy Sets Syst 54(3):269–277

Specht DF et al (1991) A general regression neural network. IEEE Trans Neural Netw 2(6):568–576

Vapnik V (1995) The nature of statistical learning theory. Springer, New York, NY

Wang L, Liu X, Pedrycz W (2013) Effective intervals determined by information granules to improve forecasting in fuzzy time series. Expert Syst Appl 40(14):5673–5679

Wang W, Pedrycz W, Liu X (2015) Time series long-term forecasting model based on information granules and fuzzy clustering. Eng Appl Artif Intell 41:17–24

Yang X, Yu F, Pedrycz W (2017) Long-term forecasting of time series based on linear fuzzy information granules and fuzzy inference system. Int J Approx Reasoning 81:1–27

Zadeh L (1979) Fuzzy sets and information granularity. Adv Fuzzy Set Theory Appl 11:3–18

Zeng S, Chen SM, Teng MO (2019) Fuzzy forecasting based on linear combinations of independent variables, subtractive clustering algorithm and artificial bee colony algorithm. Inf Sci 484:350–366