INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
Online ISSN 1841-9844, ISSN-L 1841-9836, Volume: 19, Issue: 4, Month: August, Year: 2024
Article Number: 6611, https://doi.org/10.15837/ijccc.2024.4.6611

CCC Publications

AGORA UNIVERSITY PRESS

Holiday Peak Load Forecasting Using Grammatical Evolution-Based Fuzzy Regression Approach

Guo Li, Xiang Hu, Shuyi Chen, Kaixuan Chang, Peiqi Li, Yujue Wang

Guo Li
State Grid Commercial Big Data Co., Ltd.
42 Donggexinli, Yongwai, Dongcheng District, Beijing, China
liguo_sg@126.com

Xiang Hu
State Grid Commercial Big Data Co., Ltd.
42 Donggexinli, Yongwai, Dongcheng District, Beijing, China
robin_rhu@126.com

Shuyi Chen*
State Grid Commercial Big Data Co., Ltd.
42 Donggexinli, Yongwai, Dongcheng District, Beijing, China
*Corresponding author: csy_0730@163.com

Kaixuan Chang
State Grid Commercial Big Data Co., Ltd.
42 Donggexinli, Yongwai, Dongcheng District, Beijing, China
853532107@qq.com

Peiqi Li
State Grid Commercial Big Data Co., Ltd.
42 Donggexinli, Yongwai, Dongcheng District, Beijing, China
peiqidev@163.com

Yujue Wang
State Grid Commercial Big Data Co., Ltd.
42 Donggexinli, Yongwai, Dongcheng District, Beijing, China
358639599@qq.com

Abstract

Peak load forecasting plays an important role in electric utilities. However, the daily peak load forecasting problem, especially for holidays, is fuzzy and highly nonlinear. In order to address the nonlinearity and fuzziness of the holiday load behaviors, a grammatical evolution-based fuzzy regression approach is proposed in this paper. The proposed hybrid approach is based on the theorem that fuzzy polynomial regression can model all fuzzy functions. It employs the rules of the grammatical evolution to generate fuzzy nonlinear structures in polynomial form. Then, a two-stage fuzzy regression approach is used to determine the coefficients and calculate the fitness of the fuzzy functions. An artificial bee colony algorithm is used as the evolution system to update the elements of the grammatical evolution system. The process is repeated until a fuzzy model that best fits the load data is found. After that, the developed fuzzy nonlinear model is applied to forecast holiday peak load. Considering that different holidays possess different load patterns, a separate forecaster model is built for each holiday. Test results on real load data show that an averaged absolute percent error less than 2% can be achieved, which significantly outperforms existing methods involved in the comparison.

Keywords: Load forecasting, fuzzy nonlinear model, grammatical evolution, artificial bee colony algorithm.

1 Introduction

Forecasting of electric loads is a key component of the daily operation and planning of an electric utility. Since forecasts have a significant impact on the efficiency of the operation of electric utilities, such as unit commitment, economic load dispatch, planning for energy transactions, maintenance scheduling and security analysis, they have become vital for the electricity industry. Most of the time, daily peak load forecasting is vital to utilities, for it is the first and the most important step of any network planning procedure [1]. Especially, with the development of renewable energy, peak load forecasting has emerged as a key point to reach the maximum efficiency of the system.

Thus, accurate load forecasting method should be developed. However, it is a difficult task, because the load series is complex, and there are many important factors that the relationship between them has been found to be fuzzy and highly nonlinear [2, 3].

Over the past years, many forecasting methods have been tried out and received various degrees of success. These methods can be mainly divided into two categories:

(1) traditional methods, such as linear regression [4], autoregressive models [5], Box and Jenkins method [6], Kalman filtering method [7].

(2) new forecasting methods, mainly artificial intelligence techniques, such as expert systems [8], artificial neural network [9], fuzzy inference [10], and support vector machines [11].

Traditional methods are still very attractive because they can generate load models in physical forms and some physical interpretation may be attached to them, allowing engineers and system operators to understand load behaviors [3]. However, these are basically linear methods, while the load series, especially for holidays, is known to be complex nonlinear functions, therefore, their accuracy is limited.

New forecasting methods have received much attention in recent years, because they are able to approximate complex nonlinear functions and give better performance than traditional methods in load forecasting. However, the performance of them on holidays is not as good as on weekdays. The main reason lies in that they require sufficient data sets to develop models, which are usually not available for holidays. Moreover, the load profile of holidays is more sensitive to special events [12], such as change of the weather, so it is important for engineers and system operators to understand the load behaviors of holidays. But most of these new methods are black-box in nature.

In [13] and [14], fuzzy linear regression methods based on the same day type were used to develop models for holiday load forecasting, and received good accuracy. In this paper, we extend the fuzzy regression methods considering multiple variables, and a fuzzy nonlinear regression model is adopted, namely grammatical evolution-based fuzzy regression (GEFR) method. The proposed approach can address the problems with the characteristics that only a small number of data is available and the relationship between factors is fuzzy and nonlinear. The proposed approach employs grammatical evolution (GE) to generate nonlinear structures of multiple variables, and then a two-stage fuzzy regression method is used to determine the coefficients and calculate the fitness of the fuzzy nonlinear functions. An artificial bee colony algorithm (ABC) algorithm is used to update and find the fuzzy nonlinear function that best fits the data, which is then applied to do peak load forecasting of a holiday. The proposed approach can address the nonlinearity and fuzziness of the load with a small amount of load data, and develop models in fuzzy nonlinear functions forms, which can help engineers and system operators to understand load behavior. The real load data is used to evaluate the effectiveness of the proposed approach and the numerical results show that an accurate performance is achieved.

The rest of this paper is organized as follows. In section 2, a theorem on which the proposed approach is based is introduced. In section 3, the details of the grammatical evolution-based fuzzy regression approach are described. In section 4, the process of the proposed approach for holidays peak load forecasting is presented. Finally, the test results and the validation of the proposed approach are shown in section 5.

2 Overview

Many factors, such as weather conditions, the economic environment, types of holidays, affect the loads of holidays. In [15], past weekday data relevant to a given holiday is also proved to be critical for improving the performance of load forecasting.

There are many uncertainties in these factors. For example, the randomness of weather conditions, the fuzzy character of the system control functions and the customers' demands. So, compared with crisp ones, fuzzy relationships between these factors and loads are more suitable and have been successfully used in many load forecasting problems [2, 14, 15].

Moreover, the relationships between these factors are very complex and nonlinear in nature. In order to forecast the loads accurately, a model that can address both the fuzziness and nonlinearity should be developed.

Theorem 1 shows that fuzzy polynomial regression is a very general method of modeling fuzzy nonlinear functions.

Theorem 1 [16]:Let Δ be a positive integer, $\Delta \geq 2$, Ω be a collection of all fuzzy polynomial functions of degree $\leq \Delta$. Let D be a metric on the fuzzy numbers in $\mathbb{R}$. Given $\varepsilon > 0$, then there is a sufficiently large Δ so that there is a fuzzy polynomial function $\tilde{P}$ in Ω such that

$$D(\tilde{P}(\tilde{X}_1, ..., \tilde{X}_n), \tilde{f}(\tilde{X}_1, ..., \tilde{X}_n; \tilde{K}_1, ..., \tilde{K}_n)) < \varepsilon$$

for all $\tilde{X}_i$ in $[a, b]$, where $\tilde{X}_i$ is triangle fuzzy variable, $\tilde{f}$ is any given fuzzy function, $\tilde{K}_i$ is triangle fuzzy parameter, $[a, b]$ is closed bounded interval.

Theorem 1 indicates that the fuzzy and nonlinear modeling problem can be transformed into finding a fuzzy polynomial function in Ω so that (1) is satisfied.

Stating this in brief, the fuzzy nonlinear functions in the polynomial form can model all fuzzy functions which are the extension principle extensions of continuous real-valued functions.

A fuzzy polynomial function consists of linear, higher order and interaction terms. For example, if there are three independent variables in a fuzzy model, a fuzzy polynomial function form of the model can be written as follows:

$$\tilde{Y} = \tilde{K}_1\tilde{X}_1 + \tilde{K}_2(\tilde{X}_1)^2 * \tilde{X}_3 + \tilde{K}_3\tilde{X}_3 + \tilde{K}_4(\tilde{X}_3)^2$$

where $\tilde{Y}$ is the fuzzy dependent variable.

The general form of a fuzzy polynomial model where both dependent and independent variables are fuzzy numbers can be represented as follows [17]:

$$\tilde{Y} = \tilde{K}_0 + \sum_{i_1=1}^{N} \tilde{K}_{i_1}\tilde{X}_{i_1} + \sum_{i_1=1}^{N}\sum_{i_2=1}^{N} \tilde{K}_{i_1 i_2}\tilde{X}_{i_1}\tilde{X}_{i_2} + \sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\sum_{i_3=1}^{N} \tilde{K}_{i_1 i_2 i_3}\tilde{X}_{i_1}\tilde{X}_{i_2}\tilde{X}_{i_3} + \cdots \sum_{i_1=1}^{N} \cdots \sum_{i_d=1}^{N} \tilde{K}_{i_1 \cdots i_d} \prod_{j=1}^{d} \tilde{X}_{i_j}$$

where N and d are the number of design variables, $\tilde{X}_{i_j}$ is the i_j th independent variable.

Since there are a large number of fuzzy polynomial functions in Ω, especially when multiple variables are considered, an approach to search for a fuzzy polynomial function that best fits the data is developed in the next section.

3 Grammatical evolution-based Fuzzy regression approach

In order to determine the fuzzy nonlinear model that best fits the data, a grammatical evolution-based fuzzy regression approach is proposed in this research.

Grammatical evolution (GE) is a system that employs a robust new mapping process, the end result of which is the ability to generate sentences in an arbitrary language [18]. In order to generate a sentence in an arbitrary language, three key components are needed in the process of grammatical evolution:

(1) Rules that map real-valued variables in optimization process onto a sentence in an arbitrary language.

(2) fitness functions that assign fitness to the generated sentences.

(3) evolution system that find a best sentence through a continuous optimization process.

In the proposed hybrid approach, the rules of the grammatical evolution is used to generate the fuzzy nonlinear structures, and the fuzzy regression is used to determine the coefficients and calculate the fitness of the fuzzy functions. To accelerate the grammatical evolution process and find a near-optimal or optimal solution effectively, an artificial bee colony algorithm is used in the searching process of the evolution system. The process is repeated until a fuzzy polynomial function that best fits the load data is found.

3.1 Rules of the hybrid approach

In this research, rules are used to map real numbers onto structures of fuzzy polynomial functions. Since fuzzy polynomial functions can be always represented by only two arithmetic operations, "+" and "*", the number of elements in GE is set to an odd number [19]. The elements in odd numbers are used to represent the fuzzy variables, and those in even numbers are used to represent the fuzzy arithmetic operations. All the elements are in the range between 0 to 1. Assuming that there are n variables $(\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_n)$ in the fuzzy function, melements $(p_1, p_2, ..., p_m)$ in GE, the rules are defined as follows:

Rule 1: For elements in odd numbers,

IF $0 < p \leq 1/(m+1)$, THEN no variable is represented;
IF $l/(m+1) < p \leq (l+1)/(m+1)$, THEN p represents the lth variable.

Rule 2: For elements in odd numbers,

IF $0 < p \leq 1/2$, THEN p represents the fuzzy arithmetic operation "+";
IF $1/2 < p \leq 1$, THEN p represents the fuzzy arithmetic operation "*".

For example, if there are three variables $(\tilde{X}_1, \tilde{X}_2, \tilde{X}_3)$ in the fuzzy nonlinear function, and nine elements in GE, from the rules above, the elements (0.32, 0.11, 0.43, 0.72, 0.62, 0.93, 0.76, 0.33, 0.88) represent the fuzzy polynomial structure $\tilde{X}_1 + \tilde{X}_1 * \tilde{X}_2 * \tilde{X}_3 + \tilde{X}_3$.

3.2 Fuzzy regression based fitness function

After mapping real numbers onto fuzzy polynomial structures, fuzzy regression is used to determine the coefficients of the fuzzy function. A set of historical data is used as training data in the fuzzy regression process. Then, the total error of the fuzzy regression model, which is defined as the difference between the estimated and observed fuzzy responses, is calculated as the fitness function that assigns fitness to the generated fuzzy polynomial function.

Since Tanaka et al. proposed fuzzy regression model between fuzzy dependent and crisp independent variables firstly in 1982 [20], many methodologies have been developed to improve the model [21–23].

In this research, both dependent and independent variables are fuzzy numbers. So a two-stage approach for formulating fuzzy regression models is used to develop the fuzzy model. It was also shown that this fuzzy regression approach can establish a more accurate fuzzy model with high explanatory power [23].

The fuzzy regression model of this approach contains crisp regression coefficients and a fuzzy adjustment variable. The crisp regression coefficients are used to represent the trend in scale between fuzzy dependent and fuzzy independent variables, and fuzzy adjustment variables are used to deal with the fuzziness from observations to reduce the fuzzy estimation errors. The model is formulated as follows:

$$\tilde{Y} = K_0 + K_1\tilde{X}_1 + K_2\tilde{X}_2 + \cdots K_n\tilde{X}_n + \tilde{\delta}$$

where K_i is crisp coefficient, $\tilde{\delta}$ is the fuzzy adjustment variable.

In the two-stage fuzzy regression approach, the fuzzy observations are first defuzzified into crisp values. The crisp values of the observations are used to determine the relationship between the fuzzy dependent and fuzzy independent variables. Using the classical least-squares method, the crisp regression coefficients can be calculated.

Then, the crisp regression coefficients are treated as the known parameters, and the fuzzy adjustment variables in the fuzzy regression models can be determined. In order to achieve the best explanation for the fuzzy regression model, a mathematical program to minimize the total estimation error based on the Hamming distance concept is used. The program is formulated as follows:

$$\begin{array}{ll} \text{Min} & \sum\limits_{i=1}^{n} \frac{1}{2m} \sum\limits_{k=1}^{m} \left(\left| \left(\hat{\tilde{Y}}_i'\right)_{\alpha_k}^{\mathrm{L}} - \left(\tilde{Y}_i\right)_{\alpha_k}^{\mathrm{L}} \right| + \left(\left| \left(\hat{\tilde{Y}}_i'\right)_{\alpha_k}^{\mathrm{U}} - \left(\tilde{Y}_i\right)_{\alpha_k}^{\mathrm{U}} \right| \right) \right. \\ \text{s.t.} & \left(\hat{\tilde{Y}}_i'\right)_{\alpha_k}^{\mathrm{L}} = b_0 + b_1 \left(\tilde{X}_{i1}\right)_{\alpha_k}^{\mathrm{L}} + b_2 \left(\tilde{X}_{i2}\right)_{\alpha_k}^{\mathrm{L}} + \ldots + b_p \left(\tilde{X}_{ip}\right)_{\alpha_k}^{\mathrm{L}} + (\tilde{\delta})_{\alpha_k}^{\mathrm{L}} \\ & \left(\hat{\tilde{Y}}_i'\right)_{\alpha_k}^{\mathrm{U}} = b_0 + b_1 \left(\tilde{X}_{i1}\right)_{\alpha_k}^{\mathrm{U}} + b_2 \left(\tilde{X}_{i2}\right)_{\alpha_k}^{\mathrm{U}} + \ldots + b_p \left(\tilde{X}_{ip}\right)_{\alpha_k}^{\mathrm{U}} + (\tilde{\delta})_{\alpha_k}^{\mathrm{U}} \\ & i = 1, \ldots, n; \quad k = 1, \ldots, m \end{array}$$

where $\tilde{Y}_i$ is the fuzzy observed response, $\hat{\tilde{Y}}_i$ is the fuzzy estimated response, α_k is the α-cut at the kth level. L is the lower bound and U is the upper bound. b_i is the crisp coefficient calculated in the first step.

For triangular fuzzy numbers, only two α-cuts, α equals to 0 and 1 are required to build up the fuzzy regression model.

3.3 Artificial Bee Colony algorithm based evolution system

As the fitness of each generated fuzzy polynomial function is calculated, a search method is used to find the best solution represented by the elements of GE. So an artificial Bee Colony algorithm is employed to do this.

Artificial bee colony algorithm is an optimization algorithm based on the intelligent behaviors of honey bee swarm [24], and has been used successfully in many optimization problems [25, 26]. It was shown that the performance of this algorithm is better or similar to the other swarm-based optimization algorithms in many tests [27]. In the ABC algorithm, the position of a food source represents a possible solution to the optimization problem, and the colony of artificial bees contains three groups: employed bees, which randomly search for food-source positions; onlookers, which choose good food sources from those founded by the employed bees to do further search, and scouts that abandon founded food sources and search for new ones. The main steps of ABC algorithm are as follows:

(1) Randomly generate solutions in the search spaces and calculate the objective function values, then assign these solutions to employed bees.

(2) The employed bees do the search process, each of them always remembers its previous best position and produces a new one within its neighborhood in its memory.

(3) Each onlooker chooses a food source depending on the information shared by the employed bees, and then does the search process.

(4) A food source is abandoned by the bees if it cannot be further improved through predetermined search cycles. Then a scout randomly determines a new food source and replaces the abandoned one.

(5) Go to step (2) until the termination criterion is satisfied.

3.4 The hybrid approach

The flowchart of the proposed GEFR approach is illustrated (Figure 1).

The steps of the approach are described as follows.

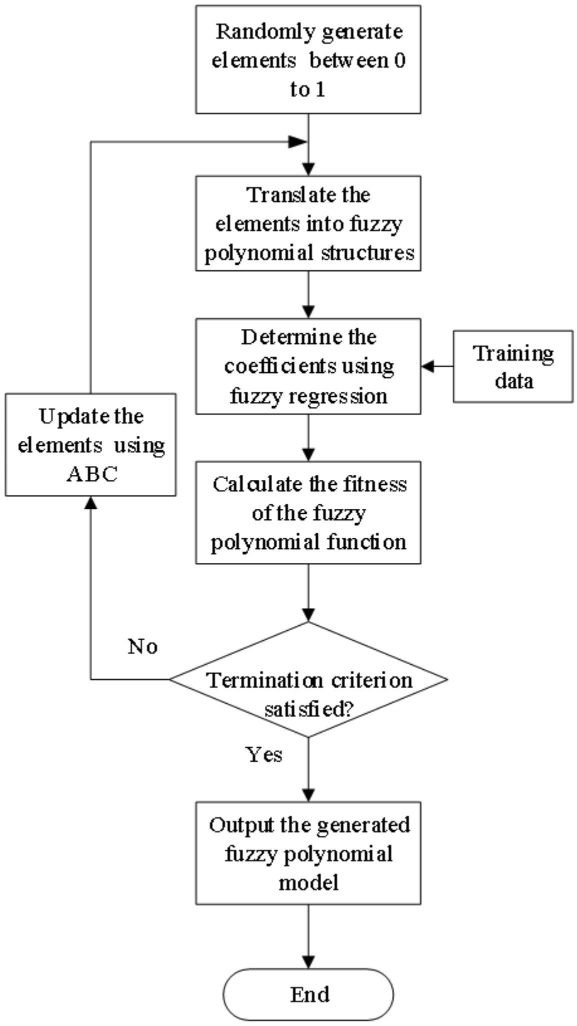

Figure 1: Flowchart of the GEFR approach

(1) Randomly generate a set of elements. In order to make sure the fuzzy polynomial functions represented by these elements be significant, the number of the elements should be an odd number. All the elements are in the range between 0 and 1.

(2) According to rule 1 and rule 2 in Section 3.2, the elements are translated into fuzzy polynomial structures. The elements in odd numbers are translated into fuzzy variables, and those in even numbers are translated into the fuzzy arithmetic operations.

(3) A set of historical data is used as training data to determine the fuzzy polynomial function. As the structures of the fuzzy polynomial functions are known, a two-stage fuzzy regression method, described in Section 3.3, is used to determine the coefficients of the functions.

(4) With the models of the fuzzy polynomial functions treated as the known parameters, the total error of the fuzzy regression model is calculated as fitness of the functions.

(5) If termination criterion is satisfied, stop and the fuzzy polynomial function that best fits the data is found. Otherwise, ABC algorithm is used to update the elements. After new elements are produced, Step (2) to Step (4) is repeated. The iteration continues until termination criterion is satisfied.

4 Fuzzy polynomial regression approach for peak load forecasting

For public holidays, only a small amount of load data is available. For a given holiday, the load profile is different from any other days of the year, that is to say, only one load data sample is available for the given holiday every year. Moreover, the load profile of holidays is more sensitive to special events, such as sudden change of the weather. So it is necessary for engineers and system operators to understand the load behaviors of holidays. To deal with these issues, the fuzzy polynomial regression approach is used for peak load forecasting of holidays.

In the fuzzy polynomial regression approach for peak load forecasting of holidays, three variables, the peak load for the weekday before the forecasted day, the temperature of the weekday before the forecasted day, and the temperature of the forecasted day, are considered to be dependent variables. Because of the fuzziness of the independent variables, such as the uncertainty of the customers' demands, the three dependent variables are presented in fuzzy numbers.

For computational efficiency, symmetric triangular fuzzy numbers are used in this research. A symmetric triangular fuzzy number $\tilde{X}$ denoted by (x, γ)

$$\tilde{X}(t) = \begin{cases} \frac{1}{\gamma}(t - x + \gamma), x - \gamma \leq t \leq x \\ \frac{1}{\gamma}(x + \gamma - t), x \leq t \leq x + \gamma \\ \quad 0 \qquad , otherwise \end{cases}$$

where x is the center of the fuzzy number, γ is the spread.

Let $\tilde{X}_i$: (x_i, γ_i) be one fuzzy independent variable, $D_{1,i}$, $D_{2,i}$, $D_{3,i}$ and $D_{4,i}$ be the load or temperature data of four weekdays before the forecasted day. The fuzzy independent variable is calculated as follows:

$$x_i = \frac{D_{1,i} + D_{2,i} + D_{3,i} + D_{4,i}}{4}$$

$$\gamma_i = \sqrt{\frac{\sum_{j=1}^{4}(D_{j,i} - x_i)^2}{4}}$$

Then, the GEFR approach described in section 3 is used to find a fuzzy polynomial function that best fits the data. In the GEFR, the number of elements reflects the complexity of the generated fuzzy model. The more elements are used, the more complex the model will be. And the training errors will be reduced. However, there are many random characters in the load forecasting problem. A model that is too complex may fit the training data well, but will lead to over-fitting problem and cause high forecasting errors on the test data. It shows the relationship between the elements number, the training error, and the testing error in a load forecasting process (Figure 2).

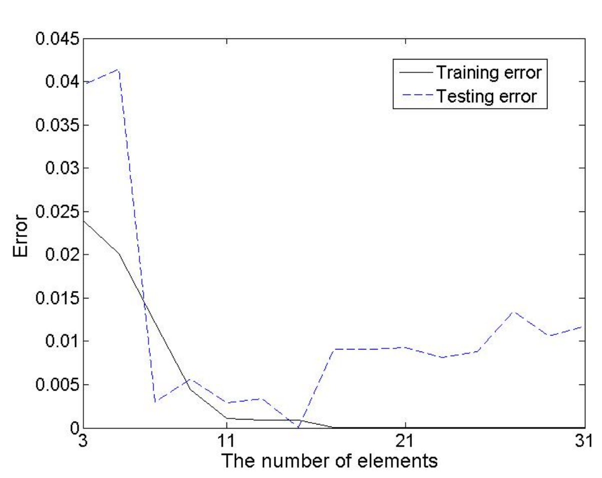

Figure 2: The relationship between the number of elements, the training error, and the testing error

In order to be away from generating too complex model, the empirical risk minimization (ERM) principle is adopted. A penalty term is introduced into the fitness function. The fitness function in the load forecasting problem is defined as follows:

$$Fitness = \frac{1}{N}\sum_{i=1}^{N} L_i + \lambda J(f)$$

Table 1: Models Generated for Independence Day

Algorithm	Fuzzy models
GEFR	$\tilde{P} = 1.08 * 10^5 - 23.67(\tilde{X}_2)^2 + 50.18\tilde{X}_2\tilde{X}_3 - 3.16 * 10^3\tilde{X}_3 + \tilde{44.30}$
FSLR	$\tilde{P} = 9.82 * 10^3 + 0.37\tilde{X}_1 + \tilde{27.78}$
FSPR	$\tilde{P} = -2.63 * 10^5 + 28.53\tilde{X}_1 - 7.23 * 10^{-4}(\tilde{X}_1)^2 + \tilde{372.63}$
FMLR	$\tilde{P} = -2.97 * 10^4 - 0.43\tilde{X}_1 + 442.80\tilde{X}_2 + 333.71\tilde{X}_3 + \tilde{205.22}$

where N is the number of training data, L_i is the error between observed value and estimated value, λ is the coefficient, $J(f)$ is the complexity of the generated model.

The dependent variable of the fuzzy polynomial model generated by GEFR is a fuzzy number. In order to calculate crisp load forecasting results, α-cuts at α=0 is used to defuzzify the dependent variable into crisp values.

5 Test results

In this section, the proposed approach is applied for peak load forecasting on a set of practical utility data. Considering that different holidays possess different load patterns, a separate forecaster model is built for each holiday.

To evaluate the effectiveness of the proposed method, another three fuzzy regression-based load forecasting methods, fuzzy simple linear regression (FSLR), fuzzy simple polynomial regression (FSPR), and fuzzy multiple linear regression (FMLR), are employed as a comparison.

The forecasting performance is represented by the absolute percent error (APE), which is evaluated as follows:

$$APE = \frac{|L_{actual} - L_{forecast}|}{L_{actual}} \times 100\%$$

where L_{actual} and $L_{forecast}$ are the actual load and the forecast load of the holiday, respectively.

For a given holiday, the load profile is different from others of the year, so the training data consist of load data of the same holiday in the past years. That is to say, there are eight training data, one training data each year, for each holiday. Each training data is of three dimensions, the peak load for the weekday before the holiday ($\tilde{X}_1$), the temperature of the weekday before the holiday ($\tilde{X}_2$), and the temperature of the holiday ($\tilde{X}_3$). Using the training data, the fuzzy models of the holidays are generated. Then, based on these models, the peak load of the holidays is calculated.

Taking Independence Day for example, all the models generated by GEFR, FSLR, FSPR and FMLR (see Table 1). From this table, it can be seen that the model generated by FSLR and FMLR are all fuzzy linear models, and the model generated by FSPR is fuzzy polynomial one, but only one variable is considered. The one generated by GEFR is fuzzy nonlinear model that contains second-order and interaction terms, with multiple variables considered.

Using the generated models, the peak load of Independence Day is calculated. After all the models have been generated, independent variables of the forecasted day are used to calculate the holiday peak load. The APE for all the holidays are shown (see Table 2).

It can be seen that compared with the other methods, the GEFR approach obtains the smallest APE (see Table 2). It can be explained that the generated fuzzy nonlinear models are able to capture the nonlinear behaviors of the load.

It should be mentioned that the forecasting result for Memorial Day shows low accuracy compared with the other holidays. It is because the limitations of the method: only a small number of training samples can be used in this method, leading to instability of prediction results. To this kind of days with anomalous load, prior expert knowledge of load profiles is needed to improve the forecasting accuracy [12].

Table 2: Comparison of APE

	GEFR	FSLR	FSPR	FMLR
New Year's Day	1.22%	6.52%	2.42%	1.22%
Martin Luther King Day	0.19%	0.31%	3.59%	0.15%
President's Day	1.31%	4.25%	6.77%	0.39%
Memorial Day	4.50%	13.70%	10.62%	2.59%
Independence Day	1.32%	19.30%	14.46%	2.16%
Labor Day	1.54%	2.96%	0.10%	2.83%
Thanksgiving Day	2.49%	1.23%	1.35%	3.56%
Christmas	1.14%	1.23%	1.40%	1.49%
Average error	1.71%	6.19%	5.09%	1.80%

6 Conclusion

To solve the problem of peak load forecasting for holidays, this paper proposed a grammatical evolution-based fuzzy regression (GEFR) approach to forecast holiday peak load. This approach is based on the theorem that fuzzy polynomial function can model all fuzzy functions which are the extension principle extensions of continuous real-valued functions. In this method, GE is used to generate fuzzy nonlinear structures, and fuzzy regression is used to determine the coefficients and calculate the fitness of each fuzzy nonlinear function.

Compared with the other load forecasting methods, the advantages of the proposed approach are:

(1) It can address the nonlinearity and fuzziness of the holiday peak load.

(2) It can develop models in fuzzy nonlinear functions forms, which are useful for engineers and system operators to understand load behaviors.

(3) Only a small amount of data is needed in this method. With the development of technologies such as large models of artificial intelligence, the method proposed in this paper can be combined with the prediction results of large models to enhance the interpretability and accuracy of the results.

Funding

This work is supported in part by the Research on Key Technologies for Supply-Demand Balance Perception and Optimization Decision Support Based on Source-Grid- Load-Storage Data Coordination (527404239004). The contracting party (Party A) of this project is State Grid Xiongan Financial Technology Group Co., Ltd.

Author contributions

The authors contributed equally to this work.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Maksimovich, S.M.; Shiljkut, V.M. (2009). The Peak Load Forecasting Afterwards Its Intensive Reduction, *IEEE Trans. Power Del.*, 24(3), 1552–1559, 2009.

[2] Nazarko, J.;Zalewski, W. (1999). The fuzzy regression approach to peak load estimation in power distribution systems, *IEEE Trans. Power Syst.*, 14(3), 809–814, 1999.

[3] Negnevitsky, M.; Mandal, P.;Srivastava, A.K. (2009). An overview of forecasting problems and techniques in power systems, In *Proc. 2009 IEEE Power Energy Soc. Gen. Meet.*, Calgary, Canada, 1–4, 2009.

[4] Haida, T.; Muto, S. (1994). *Regression based peak load forecasting using a transformation technique,*, *IEEE Trans. Power Syst.*, 9(4), 1788–1794, 1994.

[5] Huang, S.J.; Shih, K.R. (2003). Short-term load forecasting via ARMA model identification including non-Gaussian process considerations, *IEEE Trans. Power Syst.*, 18(2), 673–679, 2003.

[6] Hagan, M.T.; Behr, S.M. (1987). The Time Series Approach to Short Term Load Forecasting, *IEEE Trans. Power Syst.*, 2(3), 785–791, 1987.

[7] Guan, C.; Luh, P.B.; Michel, L.D.; Chi, Z. (2013). Hybrid Kalman Filters for Very Short-Term Load Forecasting and Prediction Interval Estimation, *IEEE Trans. Power Syst.*, 28(4), 3806–3817, 2013.

[8] Matijaš, M.; Suykens, J.A.; Krajcar, S. (2013). Load forecasting using a multivariate meta-learning system, *Expert Syst. Appl.*, 40(11), 4427–4437, 2013.

[9] Hippert, H.S.; Pedreira, C.E.; Souza, R.C. (2001); Neural networks for short-term load forecasting: a review and evaluation, *IEEE Trans. Power Syst.*, 16(1), 44-55, 2001.

[10] Hinojosa, V.H.; Hoese, A. (2010). Short-Term Load Forecasting Using Fuzzy Inductive Reasoning and Evolutionary Algorithms, *IEEE Trans. Power Syst*, 25(1), 565–574, 2010.

[11] Elattar, E.E.; Goulermas, J.; Wu, Q.H. (2010). Electric Load Forecasting Based on Locally Weighted Support Vector Regression, *IEEE Trans. Syst.*, 40(4), 438–447, 2010.

[12] Arora, S.; Taylor, J.W. (2013); Short-Term Forecasting of Anomalous Load Using Rule-Based Triple Seasonal Methods, *IEEE Trans. Power Syst.*, 28(3), 3235–3242, 2013.

[13] Song, K.B.; Baek, Y.S.; Hong, D.H.; Jang, G. (2005). Short-term load forecasting for the holidays using fuzzy linear regression method, *IEEE Trans. Power Syst.*, 20(1), 96–101, 2005.

[14] Song, K.B.;Ha, S.K., Park, J.W.; Kweon, D.J.; Kim, K.H. (2006). Hybrid load forecasting method with analysis of temperature sensitivities, *IEEE Trans. Power Syst.*, 21(2), 869–876, 2006.

[15] Wi, Y.M.; Joo, S.K.; Song, K.B. (2012). Holiday Load Forecasting Using Fuzzy Polynomial Regression With Weather Feature Selection and Adjustment, *IEEE Trans. Power Syst.*, 27(2), 596–603, 2012.

[16] Buckley, J.J.; Feuring, T. (2000). Linear and non-linear fuzzy regression: Evolutionary algorithm solutions, *Fuzzy Sets Syst.*, 112(3), 381–394, 2000.

[17] Jiang, H.; Kwong, C.K.; Ip, W.H.; Chen, Z. (2013). Chaos-Based Fuzzy Regression Approach to Modeling Customer Satisfaction for Product Design, *IEEE Trans. Fuzzy Syst.*, 21(5), 926–936, 2013.

[18] O'Neill, M.; Ryan, C. (2001). Grammatical evolution, *IEEE Trans. Evol. Comput.*, 5(4), 349–358, 2001.

[19] Chan, K.Y.; Dillon, T.S.; Kwong, C.K. (2011). Modeling of a Liquid Epoxy Molding Process Using a Particle Swarm Optimization-Based Fuzzy Regression Approach, *IEEE Trans. Ind. Informat.*, 7(1), 148–158, 2011.

[20] Tanaka, H.; Uejima, S.; Asai, K. (1982). Linear regression analysis with fuzzy model, *IEEE Trans. Syst.*, Man, Cybern 12(6), 903–907, 1982.

[21] Chen, L.H.; Hsueh, C.C. (2009). Fuzzy Regression Models Using the Least-Squares Method Based on the Concept of Distance, *IEEE Trans. Fuzzy Syst.*, 17(6), 1259–1272, 2009.

[22] Hong, D.H.; Lee, S.; Do, H.Y. (2001). Fuzzy linear regression analysis for fuzzy input–output data using shape-preserving operations, *Fuzzy Sets Syst.*, 122(3), 513–526, 2001.

[23] Chen, L.H.; Hsueh, C.C.; Chang, C.J. (2013). A two-stage approach for formulating fuzzy regression models, *Knowl.-Based Syst.*, 52(1), 302–310, 2013.

[24] Karaboga, D.; Basturk, B. (2007). A powerful and Efficient Algorithm for Numerical Function Optimization: Artificial Bee Colony (ABC) Algorithm, *J. Global Optimiz.*, 39(3), 459–471, 2007.

[25] Abu-Mouti, F.S.; El-Hawary, M.E. (2011). Optimal Distributed Generation Allocation and Sizing in Distribution Systems via Artificial Bee Colony Algorithm, *IEEE Trans. Power Del.*, 26(4), 2090–2101, 2011.

[26] Todorovic, N.; Petrovic, S. (2013). Bee Colony Optimization Algorithm for Nurse Rostering, *IEEE Trans. Syst.,Man, Cybern.,Syst.*, 43(2), 467–473, 2013.

[27] Vural, R.A.; Yildirim, T.; Kadioglu, T.; Basargan, A. (2012). Performance Evaluation of Evolutionary Algorithms for Optimal Filter Design, *IEEE Trans. Evol. Comput.*, 16(1), 135–147, 2012.

Cite this paper as:

Li, G.; Hu, X.; Chen, S.; Chang, K.; Li, P.; Wang, Y. (2024). Holiday Peak Load Forecasting Using Grammatical Evolution-Based Fuzzy Regression Approach, *International Journal of Computers Communications & Control*, 19(4), 6611, 2024.
https://doi.org/10.15837/ijccc.2024.4.6611