Int J Adv Manuf Technol (2011) 53:645–660
DOI 10.1007/s00170-010-2862-5

ORIGINAL ARTICLE

An integrated fuzzy regression–analysis of variance algorithm for improvement of electricity consumption estimation in uncertain environments

A. Azadeh · O. Seraj · M. Saberi

Received: 12 August 2009 / Accepted: 27 July 2010 / Published online: 14 September 2010

Abstract This study presents an integrated fuzzy regression–analysis of variance (ANOVA) algorithm to estimate and predict electricity consumption in uncertain environment. The proposed algorithm is composed of 16 fuzzy regression models. This is because there is no clear cut as to which of the recent fuzzy regression model is suitable for a given set of actual data with respect to electricity consumption. Furthermore, it is difficult to model uncertain behavior of electricity consumption with conventional time series and proper fuzzy regression could be an ideal substitute for such cases. The algorithm selects the best model by mean absolute percentage error (MAPE), index of confidence (IC), distance measure, and ANOVA for electricity estimation and prediction. Monthly electricity consumption of Iran from 1992 to 2004 is considered to show the applicability and superiority of the proposed algorithm. The unique features of this study are threefold. The proposed algorithm selects the best fuzzy regression model for a given set of uncertain data by standard and proven methods. The selection process is based on MAPE, IC, distance to ideal point, and ANOVA. In contrast to previous studies, this study presents an integrated approach because it considers the most important fuzzy regression approaches, MAPE, IC, distance measure, and ANOVA for selection of the preferred model for the given data. Moreover, it always guarantees the preferred solution through its integrated mechanism.

Keywords Fuzzy regression · Fuzzy mathematical programming · Electricity consumption · Analysis of variance · Uncertainty · Mean absolute percentage error · Index of confidence

A. Azadeh (✉) · O. Seraj
Department of Industrial Engineering and Center of Excellence for Intelligent-Based Experimental Mechanics, College of Engineering, University of Tehran,
P.O. Box 11365-4563, Tehran, Iran
e-mail: aazadeh@ut.ac.ir

M. Saberi
Department of Industrial Engineering, University of Tafresh, Tafresh, Iran
e-mail: mortezza_saber@yahoo.com

M. Saberi
Institute for Digital Ecosystems and Business Intelligence, Curtin University of Technology,
Perth, Australia
e-mail: M.Saberi@curtin.edu.au

1 Significance

This paper presents an intelligent integrated algorithm to model non-stationary time series data with respect to electricity consumption. This is the first study that integrates fuzzy regression, mean absolute percentage error (MAPE), index of confidence (IC), distance measure, and ANOVA for modeling electricity consumption. A unique feature of this study is consideration of 16 fuzzy regression models as the important proposed fuzzy regressions. Using ANOVA together with Duncan's multiple range test (DMRT) increases the reliability of the proposed integrated algorithm. This is also the first study that applies 16 appropriate fuzzy regressions in area of electricity consumption. Four proper criteria are proposed in the present study in contrast to similar studies in which only one simple criterion is proposed.

2 Introduction

Electricity, as a resource of energy, with its ever growing role in world economy, and its multi-purpose application in production and consumption has gained special attention. Through the development of societies and growth of economical activities, electricity becomes more effective on corporations and their services. Corporations use electricity as a production factor. Also, families directly or indirectly rely on electricity. Thus, energy consumption determines their and the society's economical welfare.

A major topic applicable in this criterion is estimation of the electricity consumption, which reveals the consumption growth in the forthcoming years. The most important goal is to find an applicable method with less error to estimate and predict this consumption.

In the real world, sometimes the data cannot be considered and modeled precisely, for example, the water level of a river and the temperature in a room cannot be measured precisely because of their fluctuations [5–7, 39]. The electricity consumption is also not able to be measured in an exact way because of the similar reason [1–4].

Regression analysis is one of the most used statistical tools to explain the variation of a dependent variable Y in terms of the variation of explanatory variables X as: $Y=\mathrm{f}(X)$, where $\mathrm{f}(X)$ is a linear function. It refers to a set of methods by which estimates are made for the model parameters from the knowledge of the values of a given input–output data set. The goal of the regression analysis is:

1. To find an appropriate mathematical model, and
2. To determine the best fitting coefficients of the model from the given data.

The use of statistical regression is bounded by some strict assumptions about the given data. This model can be applied only if the given data are distributed according to a statistical model and the relation between X and Y is crisp. Overcoming such limitations, fuzzy regression is introduced which is an extension of the classical regression and is used in estimating the relationships among variables where the available data are very limited and imprecise, and variables are interacting in an uncertain, qualitative, and fuzzy way. Since the concepts of fuzzy sets are introduced in the literature, the applications of considering fuzzy data to the regression models have been proposed [40].

"Fuzzy regression methods have been developed to deal with fuzzy observations represented as linguistic terms [18]". Fuzzy regression models have been successfully applied to various problems such as forecasting [37, 38], engineering [17, 23], and financial [19]. Chan et al. [17] have used genetic programming-based fuzzy regression to model manufacturing processes. Chen at al. (2009) have used fuzzy regression approach in order to manage cash balance in construction firms. In the present study, it will be shown that it can also be applied for energy forecasting problems. The goal of fuzzy regression analysis is to find a regression model that fits all observed fuzzy data within a specified fitting criterion. Different fuzzy regression models are obtained depending on the fitting criterion used. In general, there are two approaches of fuzzy regression due to different fitting criterions [30]. The first approach is based on minimizing fuzziness as an optimal criterion which first proposed by Tanaka [36]. Different researchers used Tanaka's approach to minimize the total spread of the output [14]. As pointed out by Wang and Tsaur [37, 38], the advantage of this approach is its simplicity in programming and computation, but it has been criticized to provide too wide ranges in estimation which could not give much help in application [38] and not to utilize the concept of least-squares [15]. The second approach uses least-squares of errors as a fitting criterion to minimize the total square error of the output. Different aspects of this approach were investigated by Celmins [12, 13], Diamond [20], Savic and Pedrycz [33], and Chang and Ayyub [15] Celmins [12, 13] defines a compatibility measure between fuzzy data and a model, and uses this measure as a model-fitting criterion. Diamond [20] developed a fuzzy least-square method by using the compact α-level sets. Savic and Pedrycz [33] proposed a combined approach for fuzzy least-square regression analysis (FLSRA) by integrating minimum fuzziness criterion into the ordinary least-squares regression. Chang and Ayyub [15] discussed reliability issues of FLSRA such as standard error and correlation coefficient. This approach, though providing narrower range, costs too much of computation time (2000). Therefore, a natural extension of fuzzy regression would be the integration of the least-squares' concept into fuzzy regression.

The concept of random fuzzy sets, which implies the fuzzy random variables into the linear regression model, is introduced by Nather [25, 27], Nather and Albrecht [26], and Korner and Nather [53]. A doubly linear adaptive fuzzy regression model, which is based on both of core and spread regression model, is proposed by D'Urso and Gastaldi [21].

For our estimation of electricity consumption, a model is used in which minimizing the total squares errors of the spread value is used as the fitting criterion, and a mathematical programming approach is developed such that the predictability of first approach can be improved and the computation complexity of the second method can be decreased. The used model will be introduced more completely in the next section.

In the next section, the most well-known fuzzy regression models and their advantages and shortcomings are discussed. Moreover, 16 well-known fuzzy regression models are discussed. They are the most well-known and cited fuzzy regression methods in the literature. That is why they are used as the cores of this paper.

3 Manufacturing technology and electricity consumption

Exploring literature in field of energy consumption shows that there is a positive relationship between gross domestic production (GDP) and electricity consumption, for example, Narayan et al. [43] have shown that "except for USA, electricity consumption has a statistically significant positive impact on real GDP." Improving some of the main features of manufacturing technology is directly related to electricity consumption. Also, the limitations of energy resources and strictly increasing electricity consumption trend show the need to design accurate devices for consumption of electricity in manufacturing sector in particular. Hence, there is a need to focus on trend of electricity consumption in the future, particularly in manufacturing sector, for example, if the future electricity consumption is accurately estimated and forecasted with high and acceptable confidence, then the needed power plants can be designed and constructed in timely manner for higher efficiency and productivity.

Moreover, the industrial models have classified the industries in two general groups: (1) high energy-consuming industries and (2) low energy-consuming industries. Three kinds of industries out of all industries are placed in high energy-consuming industries which are chemicals, basic metals, and non metal minerals. Also, low energy-consuming industries consist of food-, textile-, paper-, and machinery-producing industries. The variables affecting the demand for electricity are different for each group. The electricity consumption in high energy-consuming industries has more heterogeneity and changes than low energy-consuming industries. As shown in Fig. 1, electricity consumption in high energy-consuming industries has uncertain structure which makes fuzzy regression an ideal candidate for forecasting estimations.

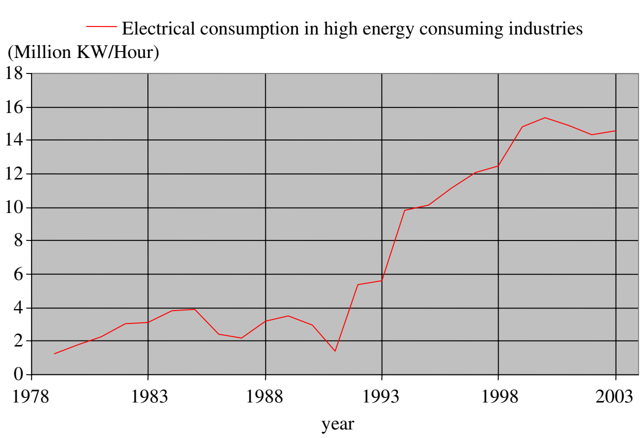


Fig. 1 The trend of electricity consumption in high energy-consuming manufacturing sectors in Iran

As mentioned, forecasting electricity consumption in high energy-consuming manufacturing sector requires advanced intelligent tools such as fuzzy regression. In manufacturing sector, electricity is used both in building components for cooling, heating, and lighting which varies according to the workforce increase, building extension and weather condition, and in the operational process for mechanical and electronic processing. One of the overriding characteristics in the manufacturing sectors is the heterogeneity of industries, products, equipment, technologies, processes, and energy uses. This creates a new problem in forecasting electricity consumption in manufacturing sectors. Even though, this study is concerned with total electricity consumption, the structure and results of this study may be easily extended to manufacturing sectors particularly in high energy-consuming sectors.

4 Fuzzy regression models

Fuzzy linear regression (FLR) was introduced by Tanaka et al. [36] to decide a fuzzy linear relationship by:

$$\widehat{Y} = A_0X_0 + A_1X_1 + \ldots + A_kX_k,$$

where regression coefficients A_j, $j=0\ldots K$, were supposed to be a symmetric triangular fuzzy number, with center α_j, having membership function equal to one, and spreads c_j, $c_j \geq 0$.

Two conditions are studied in this model:

1. Independent variables (x) are crisp, and dependent variable (y) is a fuzzy number.
2. Independent variables (x) are fuzzy numbers, and dependent variable (y) is also a fuzzy number.

The input information are n sets of variables (y_i, x_{i0}, $x_{i1}, \ldots, x_{ik}$), $i=1,2,\ldots n$, $n \geq k+1$, where $x_{i0}=1$. The response variable y_i is assumed to be a symmetric triangular fuzzy number with central value y_i and spreads e_i, where $e_i \geq 0$. In condition 2, independent variables values x_{ij}, $i=1, 2\ldots n$, $j=1,2,\ldots,k$, is also supposing being a symmetric triangular fuzzy number with a center x_{ij} and spreads f_{ij}.

We assigned a symmetric triangular fuzzy number for response variable y_i. If we are just interested in that membership function value of y_i that has at least H, $0 \leq H \leq 1$, we should consider the interval $[\overline{y}_i - (1-H) \times e_i\ \overline{y}_i + (1-H) \times e_i]$. This interval is illustrated in Fig. 2.

Here, H shows the minimum acceptable degree of precision, and we will make reference to this interval as H-certain observed interval. Similarly, suppose that the independent variables x_j, $j=1, 2,\ldots k$, have certain values and regression coefficient Aj, $j=1,2,\ldots k$, are assume to be symmetric triangular fuzzy numbers, the estimated interval corresponding to a input set of independent variables X (x_{i0},

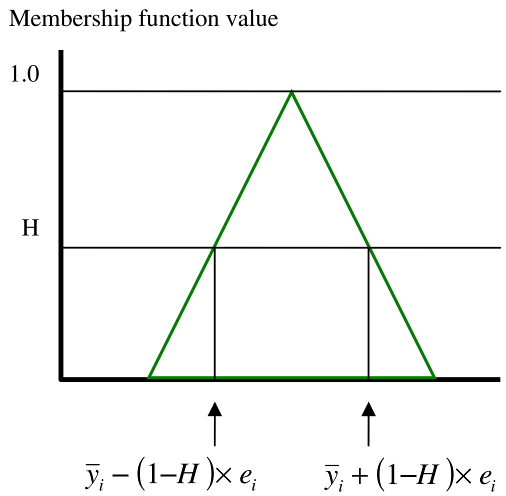


Fig. 2 An H-certain observed interval

$x_{i1},\ldots, x_{ik}$) having membership function value of at least H is: $\left[\sum_{j=0}^{k}\left(\alpha_j-(1-H)\times c_j\right)\times x_{ij}\quad \sum_{j=0}^{k}\left(\alpha_j+(1-H)\times c_j\right)\times x_{ij}\right]$; we will refer to this distance as H-certain estimated interval [22].

For case 1, Tanaka et al. [36] introduced the following linear programming formulation to predict Aj, j=1, 2,...k:

$$\text{minimize}\quad c_0+c_1+c_2+\cdots+c_k \tag{1}$$

subject to :

$$\begin{aligned}
&\sum_{j=0}^{k}\left(\alpha_j+(1-H)\times c_j\right)\times x_{ij}\geq \bar{y}_i+(1-H)\times e_i \qquad i=1,\cdots,n,\\
&\sum_{j=0}^{k}\left(\alpha_j-(1-H)\times c_j\right)\times x_{ij}\leq \bar{y}_i-(1-H)\times e_i \qquad i=1,\cdots,n,\\
&\alpha_j=\text{free},\ c_j\geq 0,\quad j=0,\cdots,k.
\end{aligned}$$

Note that in the above model, c_js are supposed to be non-negative because the fuzziness in estimated intervals usually increases for larger values of independent variables x_j [22]. The formulation of fuzzy regression model of Tanaka et al. [34] is:

$$\text{minimize}\quad \sum_{i=1}^{n}\sum_{j=0}^{k}c_j x_{ij} \tag{2}$$

subject to :

$$\begin{aligned}
&\sum_{j=0}^{k}\left(\alpha_j+(1-H)\times c_j\right)\times x_{ij}\geq \bar{y}_i+(1-H)\times e_i \qquad i=1,\cdots,n,\\
&\sum_{j=0}^{k}\left(\alpha_j-(1-H)\times c_j\right)\times x_{ij}\leq \bar{y}_i-(1-H)\times e_i \qquad i=1,\cdots,n,\\
&\alpha_j=\text{free},\ c_j\geq 0,\quad j=0,\cdots,k.
\end{aligned}$$

Sakawa and Yano studied case 2 [32]. First, depending upon the presumed range of values of coefficients A_j, Sakawa and Yano would categorize the independent variables into three classes:

$$\begin{aligned}
J_1=&\ \text{those variables } j, j=0,\cdots,k,\\
&\ \text{which will have } \alpha_j-(1-H)\times c_j\geq 0,\\
J_2=&\ \text{those variables } j, j=0,\cdots,k,\\
&\ \text{which will have } \alpha_j-(1-H)\times c_j 0,\\
&\ \text{and } \alpha_j+(1-H)\times c_j\geq 0,\\
J_3=&\ \text{those variables } j, j=0,\cdots,k,\\
&\ \text{which will have } \alpha_j+(1-H)\times c_j 0.
\end{aligned}$$

Then, the fuzzy regression model of this approach will be formulated as follows:

$$\text{minimize}\quad \sum_{i=1}^{n}\left(\widehat{y}_{iU}-\widehat{y}_{iL}\right) \tag{3}$$

subject to :

$$\begin{aligned}
&\sum_{j\in J_1\cup J_2}\left(\alpha_j+(1-H)\times c_j\right)\times\left(\bar{x}_{ij}+(1-H)\times f_{ij}\right)\\
&\qquad+\sum_{j\in J_3}\left(\alpha_j+(1-H)\times c_j\right)\times\left(\bar{x}_{ij}-(1-H)\times f_{ij}\right)=\widehat{y}_{iU},\\
&i=1,\cdots,n,\\
&\widehat{y}_{iU}\geq\bar{y}_i-(1-H)\times e_i,\quad i=1,\cdots,n,\\
&\sum_{j\in J_1}\left(\alpha_j-(1-H)\times c_j\right)\times\left(\bar{x}_{ij}-(1-H)\times f_{ij}\right)\\
&\qquad+\sum_{j\in J_2\cup J_3}\left(\alpha_j-(1-H)\times c_j\right)\times\left(\bar{x}_{ij}+(1-H)\times f_{ij}\right)=\widehat{y}_{iL},\\
&i=1,\cdots,n,\\
&\widehat{y}_{iL}\leq\bar{y}_i+(1-H)\times e_i,\quad i=1,\cdots,n,\\
&\alpha_j=\text{free},\ c_j\geq 0,\quad j=0,\cdots,k.
\end{aligned}$$

Sakawa and Yano also proposed the following problem, which is similar to formulation of Tanaka et al [36]:

$$\text{minimize}\quad \sum_{i=1}^{n}\left(\widehat{y}_{iU}-\widehat{y}_{iL}\right) \tag{4}$$

subject to :

$$\begin{aligned}
&\sum_{j\in J_1\cup J_2}\left(\alpha_j+(1-H)\times c_j\right)\times\left(\bar{x}_{ij}+(1-H)\times f_{ij}\right)\\
&\qquad+\sum_{j\in J_3}\left(\alpha_j+(1-H)\times c_j\right)\times\left(\bar{x}_{ij}-(1-H)\times f_{ij}\right)=\widehat{y}_{iU},\\
&i=1,\cdots,n,\\
&\widehat{y}_{iU}\geq\bar{y}_i-H\times e_i,\quad i=1,\cdots,n,\\
&\sum_{j\in J_1}\left(\alpha_j-(1-H)\times c_j\right)\times\left(\bar{x}_{ij}-(1-H)\times f_{ij}\right)\\
&\qquad+\sum_{j\in J_2\cup J_3}\left(\alpha_j-(1-H)\times c_j\right)\times\left(\bar{x}_{ij}+(1-H)\times f_{ij}\right)=\widehat{y}_{iL},\\
&i=1,\cdots,n,\\
&\widehat{y}_{iL}\leq\bar{y}_i+H\times e_i,\quad i=1,\cdots,n,\\
&\alpha_j=\text{free},\ c_j\geq 0,\ j=0,\cdots,k.
\end{aligned}$$

Sakawa and Yano introduced an interactive method to find the suitable value of H, by balancing the increase in the value of H, versus the increase in the objective function's value [22, 32].

Peters [30] considered case 1. His fuzzy regression model is a little complex to explain. Assume that y_{iU}, y_i, and y_{iL} be the upper, center, and lower values of ith observed interval, and let $\widehat{y}_{iU}$ and $\widehat{y}_{iL}$ be the upper and lower values of the ith estimated interval. This model permits $\widehat{y}_{iL}$ to be greater than y_{iL} but smaller than y_{iU}, and $\widehat{y}_{iU}$ to be smaller than y_{iU} but greater than y_{iL}. In fact, the mean of all deviations of $\widehat{y}_{iU}$ from y_i, if $\widehat{y}_{iU}y_i$, and $\widehat{y}_{iL}$ from y_i, if $\widehat{y}_{iL}y_i$, is minimized. This objective function is balanced against the total spreads of estimated intervals equation. By changing Tanaka et al. model [34] into an objective and converting it as a constraint, the formulation of Peters [30] model is:

$$\text{maximize} \quad \overline{\lambda} \tag{5}$$

subject to :

$$\begin{aligned}
&\sum_{j=0}^{k}\left(\alpha_j + c_j\right) \times x_{ij} \geq \overline{y}_i - (1-\lambda_i) \times e_i \qquad i = 1, \cdots, n,\\
&\sum_{j=0}^{k}\left(\alpha_j - c_j\right) \times x_{ij} \leq \overline{y}_i + (1-\lambda_i) \times e_i \qquad i = 1, \cdots, n,\\
&\overline{\lambda} = (\lambda_1 + \lambda_1 + \cdots \lambda_1)/n,\\
&\sum_{i=1}^{n}\sum_{j=0}^{k} c_j x_{ij} \leq P_0 \times \left(1 - \overline{\lambda}\right),\\
&0 \leq \lambda_i \leq 1, \quad i = 1, \cdots, n, \quad \overline{\lambda} \geq 0,\\
&\alpha_j = \text{free},\ c_j \geq 0, \quad j = 0, \cdots, k.
\end{aligned}$$

The following formulation was considered with only one membership function to which the result belongs to the set of good solution λ for all constraints. Therefore, a new FLR formulation is proposed by Peters and applied to this case study:

$$\text{maximize} \quad \lambda \tag{6}$$

subject to :

$$\begin{aligned}
&\sum_{i=1}^{n}\sum_{j=0}^{k} c_j x_{ij} \leq P_0(1-\lambda)\\
&\sum_{j=0}^{k}\left(\alpha_j + c_j\right) x_{ij} \geq \overline{y}_i - (1-\lambda) e_i \quad i = 1, \ldots, n,\\
&\sum_{j=0}^{k}\left(\alpha_j - c_j\right) x_{ij} \geq \overline{y}_i + (1-\lambda) e_i \quad i = 1, \ldots, n,\\
&0 \leq \lambda \leq 1\ ,\ \mathrm{c_j} \geq 0
\end{aligned}$$

Moreover, Wang and Tsaur [37] proposed a new formulation by minimization of the central values with eases the relations of constraint from Peters formulation [30]:

$$\text{maximize} \quad \lambda \tag{7}$$

subject to :

$$\begin{aligned}
&(1-\lambda) \times p - \sum_{j=0}^{N}\left(\sum_{i=1}^{M} c_j \times x_{ij} - e_i\right)^2 \geq -d_0,\\
&(1-\lambda) \times p + \sum_{j=0}^{N} \alpha_j \times x_{ij} + \sum_{j=0}^{N} c_j \times \left|x_{ij}\right| \geq \overline{y}_i + e_i \qquad \forall i = 1, 2, \cdots M,\\
&(1-\lambda) \times p - \sum_{j=0}^{N} \alpha_j \times x_{ij} + \sum_{j=0}^{N} c_j \times \left|x_{ij}\right| \geq -\overline{y}_i + e_i \qquad \forall i = 1, 2, \cdots M,\\
&0 \leq \lambda \leq 1, \qquad \forall i = 1, \cdots, M, \quad x_{i0} = 1,\\
&\alpha_j = \text{free}, \quad c_j \geq 0, \quad j = 0, \cdots, k.
\end{aligned}$$

This model ensures that the optimal value of λ is obtained by the maximum value of λs that satisfies all of the constraints of this model. Therefore, λ is treated in all constraints equally [37, 38]. Ozelkan and Duckstein [29] introduced a similar model to Peters [30], but have not needed the estimation intervals to divide the observed intervals. The formulation of Ozelkan and Duckstein [29] can be written as follows:

$$\text{minimize} \quad \sum_{i=1}^{n}\left(d_{iU} + d_{iL}\right) \tag{8}$$

subject to :

$$\begin{aligned}
&\sum_{j=0}^{k}\left(\alpha_j + (1-H) \times c_j\right) \times x_{ij} \geq \overline{y}_i + (1-H) \times e_i - d_{iU} \quad i = 1, \cdots, n,\\
&\sum_{j=0}^{k}\left(\alpha_j - (1-H) \times c_j\right) \times x_{ij} \leq \overline{y}_i - (1-H) \times e_i + d_{iL} \quad i = 1, \cdots, n,\\
&\sum_{i=1}^{n}\sum_{j=0}^{k} c_j x_{ij} \leq \upsilon,\\
&d_{iL}, d_{iU} \geq 0, \qquad i = 1, \cdots, n,\\
&\alpha_j = \text{free}, c_j \geq 0, \quad j = 0, \cdots, k,
\end{aligned}$$

where υ is a parameter and which should be diversified over all possible amounts of total spreads of estimated intervals, and d_{iU} and d_{iL}, i=1..., n, are upper and lower shift variables.

Hojati et al. [22] introduced a simple goal programming-like method to select the fuzzy regression coefficients such that the total deviation of upper values of H-certain estimated and corresponded observed intervals and deviation of lower values of H-certain estimated and related observed intervals are minimized. This can be obtained by using the following formula:

$$\text{minimize} \quad \sum_{i=1}^{n}\left(d_{iU}^{+} + d_{iU}^{-} + d_{iL}^{+} + d_{iL}^{-}\right) \tag{9}$$

subject to :

$$\begin{aligned}&\sum_{j=0}^{k}\left(\alpha_j+(1-H)\times c_j\right)\times x_{ij}-d_{iU}^{-}\geq y_i\\&\qquad+(1-H)\times e_i-d_{iU}^{+}\qquad i=1,\cdots,n,\\&\sum_{j=0}^{k}\left(\alpha_j-(1-H)\times c_j\right)\times x_{ij}-d_{iL}^{-}\leq y_i\\&\qquad-(1-H)\times e_i+d_{iL}^{+}\qquad i=1,\cdots,n,\\&\sum_{i=1}^{n}\sum_{j=0}^{k}c_jx_{ij}\leq\upsilon,\\&d_{iU}^{+},d_{iU}^{-},d_{iL}^{+},d_{iL}^{-}\geq 0,\qquad i=1,\cdots,n,\\&\alpha_j=\text{free},\ c_j\geq 0,\qquad j=0,\cdots,k,\end{aligned}$$

Note that for each of the indices i=1...N, at most, one of d_{iU}^{+} or d_{iU}^{-} and d_{iL}^{+}or d_{iL}^{-} would be positive [22].

In case 2, Hojati et al.[22] select the fuzzy regression coefficients so that the total difference between upper values of estimated and related observed intervals is minimized at both lower values and upper values of each of the independent variable.

Minimize

$$\sum_{i=1}^{n}\left(d_{ilU}^{+}+d_{ilU}^{-}+d_{ilL}^{+}+d_{ilL}^{-}+d_{irU}^{+}+d_{irU}^{-}+d_{irL}^{+}+d_{irL}^{-}\right)\tag{10}$$

subject to :

$$\begin{aligned}&\sum_{j=0}^{l}\left(\alpha_j+(1-H)\times c_j\right)\times\left(\bar{x}_{ij}-(1-H)\times f_{ij}\right)-d_{ilU}^{-}\\&\qquad=\bar{y}_i+(1-H)\times e_i-d_{ilU}^{+},\\&i=1,\cdots,n,\\&\sum_{j=0}^{l}\left(\alpha_j+(1-H)\times c_j\right)\times\left(\bar{x}_{ij}+(1-H)\times f_{ij}\right)-d^{-}{}_{irU}\\&\qquad=\bar{y}_i-(1-H)\times e_i-d_{irU}^{+},\\&i=1,\cdots,n,\\&\sum_{j=0}^{l}\left(\alpha_j-(1-H)\times c_j\right)\times\left(\bar{x}_{ij}-(1-H)\times f_{ij}\right)-d_{ilL}^{-}\\&\qquad=\bar{y}_i+(1-H)\times e_i-d_{ilL}^{+},\\&i=1,\cdots,n,\\&\sum_{j=0}^{l}\left(\alpha_j-(1-H)\times c_j\right)\times\left(\bar{x}_{ij}+(1-H)\times f_{ij}\right)-d_{irL}^{-}\\&\qquad=\bar{y}_i-(1-H)\times e_i-d_{irL}^{+},\\&i=1,\cdots,n,\\&\sum_{i=1}^{n}\sum_{j=0}^{k}c_jx_{ij}\leq\upsilon,\\&d_{ilU}^{+},d_{ilU}^{-},d_{ilL}^{+},d_{ilL}^{-},d_{irU}^{+},d_{irU}^{-},d_{irL}^{+},d_{irL}^{-}\geq 0,\qquad i=1,\cdots,n,\\&\alpha_j=\text{free},\ c_j\geq 0,\quad j=0,\cdots,k,\end{aligned}$$

where the indices l and r refer, respectively, to the lower and upper values for the intervals of the independent variable; moreover, U refers to the upper value and L refers to the lower value of the observed and estimated intervals [22]. Chen et al. [16] proposed a fuzzy least-squares regression model. The formulation of this model is:

minimize

$$\sum_{i=1}^{n}\left(y_i^m-a^t\times x_i\right)^2+\sum_{i=1}^{n}\left[\left(e_i-c^t\times x_i\right)^2+\left(e_i-c^t\times x_i\right)^2\right]\tag{11}$$

subject to :

$$\begin{aligned}&y_i^m-(1-H)\times e_i\geq a^t\times x_i-c^t\times x_i\qquad i=1,\cdots,n,\\&y_i^m+(1-H)\times e_i\leq a^t\times x_i+c^t\times x_i\qquad i=1,\cdots,n,\\&c\geq 0,x_i\geq 0,\end{aligned}$$

where H=[0, 1].

Wang et al. [37, 38] also considered least-squares approach to minimize the total vagueness of the given data, such that the membership degree of each observation is greater than the threshold h. This leads to the following nonlinear programming problem:

$$\text{minimize}\quad\sum_{j=0}^{N}\left(\sum_{i=1}^{M}c_jx_{ij}-e_i\right)^2\tag{12}$$

subject to :

$$\begin{aligned}&\sum_{j=0}^{N}\alpha_j\times x_{ij}+(1-h)^{1/2}\times\sum_{j=0}^{N}c_j\times\left|x_{ij}\right|\geq\bar{y}_i+(1-h)^{1/2}\times e_i\\&\forall i=1,\cdots,M,\\&-\sum_{j=0}^{N}\alpha_j\times x_{ij}+(1-h)^{1/2}\times\sum_{j=0}^{N}c_j\times\left|x_{ij}\right|\geq-\bar{y}_i+(1-h)^{1/2}\times e_i\\&\forall i=1,\cdots,M,\\&\alpha_j\in R,c_j\geq 0,\quad j=1,2,\cdots,N,x_{i0}=1,\quad 0\leq\lambda\leq 1.\end{aligned}$$

Chang et al. [15] introduced a FLR model. The input and output data are crisp, but coefficients are triangle fuzzy numbers. The formulation of this model can be written as follows:

$$\text{minimize}\quad\sum_{i=1}^{n}\sum_{j=0}^{k}c_jx_{ij}\tag{13}$$

subject to :

$$\begin{aligned}&\sum_{j=0}^{k}\left(\alpha_j+(1-H)\times c_j\right)\times x_{ij}\geq\bar{y}_i\qquad i=1,\cdots,n,\\&\sum_{j=0}^{k}\left(\alpha_j-(1-H)\times c_j\right)\times x_{ij}\leq\bar{y}_i\qquad i=1,\cdots,n,\\&\alpha_j=\text{free},\ c_j\geq 0,\quad j=0,\cdots,k.\end{aligned}$$

In Tanaka et al. [34] approach, the objective is minimizing the total spread of fuzzy coefficient $\widetilde{A}_j$. Thus, the relevance between response variable and independent variables is wanted to be as certain as possible. When studying the minimal total spread, such fuzziness or uncertainty should be considered as $\sum_{i=1}^{n}\sum_{j=0}^{k}\left(c_j x_{ij}-e_i\right)$, which is illustrated in Fig. 2. This model is proposed by Lai and Chang [23].

$$\text{Minimize}\quad \sum_{j=0}^{N}\left(\sum_{i=1}^{M} c_j x_{ij}-e_i\right) \tag{14}$$

subject to :

$$\begin{aligned}
&\sum_{j=0}^{N}\alpha_j \times x_{ij}+(1-h)\times\sum_{j=0}^{N}c_j\times\left|x_{ij}\right|\geq y_i+(1-h)\times e_i\\
&\qquad\forall i=1,\cdots,M,\\
&-\sum_{j=0}^{N}\alpha_j \times x_{ij}+(1-h)\times\sum_{j=0}^{N}c_j\times\left|x_{ij}\right|\geq -y_i+(1-h)\times e_i\\
&\qquad\forall i=1,\cdots,M,\\
&\alpha_j\in R, c_j\geq 0,\quad j=1,2,\cdots,N, x_{i0}=1,\quad 0\leq\lambda\leq 1.
\end{aligned}$$

As mentioned by Modarres et al. [24] and Nasrabadi et al. [28], a quadratic programming could be formulated as follows:

$$\text{minimize}\quad D(h)=\sum_{i=1}^{m}\ (e_i-e_i)^2 \tag{15}$$

subject to :

$$\begin{aligned}
&\widetilde{Y}_{i,L}(h)\leq\widetilde{Y}_{i,R}(h),\quad \widetilde{Y}_{i,R}(h)\geq\widetilde{Y}_{i,L}(h),\quad i=1,\cdots,m,\\
&e_i\geq 0,\quad i=1,\cdots,m,
\end{aligned}$$

where the h points to the level of fitness of the predicted FLR model and should be selected by DM.$\widetilde{Y}_{i,L}(h)$,$\widetilde{Y}_{i,R}(h)$ and $\widetilde{Y}_{i,L}(h)$, $\widetilde{Y}_{i,R}(h)$are, respectively, the lower and upper points of the h-certain intervals of the $\widetilde{Y}_i$ and $\widetilde{Y}_i$ fuzzy numbers [28]. As mentioned by Peters [30] and Ozelkan and Duckstein [29], outliers can be formulated by considering soft constraints to the FLR model. To achieve the maximum degree of fitness of the predicted FLR model, h is set to be one. Then, the model can be constructed as follows:

$$\text{minimize}\quad D^2(h)=\sum_{i=1}^{m}(e_i-e_i)^2, \tag{16}$$

$$\text{minimize}\quad E^2(h)=\sum_{i=1}^{m}\left(\varepsilon_{i,L}^2-\varepsilon_{i,R}^2\right)^2,$$

$$\text{subject to :}\quad \begin{aligned}
&y_i-\overline{y}_i\leq\varepsilon_{i,L}, && i=1,\cdots,m,\\
&\overline{y}_i-y_i\leq\varepsilon_{i,R}, && i=1,\cdots,m,\\
&e_i\geq 0,\quad \varepsilon_{i,L},\varepsilon_{i,R}\geq 0, && i=1,\cdots,m,
\end{aligned}$$

where E^2 (h) is the difference from outliers, and $\varepsilon_{i,L}$ and $\varepsilon_{i,R}$ can be regarded as softening variables [28]. In weighting objectives approach, to generate the efficient solutions to the above multi-objective fuzzy linear regression (MOFLR) model, the following mathematical programming problem is considered:

minimize

$$Z(h)=\omega\sum_{i=1}^{m}(e_i-\overline{e}_i)^2+(1-\omega)\sum_{i=1}^{m}\left(\varepsilon_{i,L}^2-\varepsilon_{i,R}^2\right)^2, \tag{17}$$

$$\text{subject to :}\quad \begin{aligned}
&y_i-\overline{y}_i\leq\varepsilon_{i,L}, && i=1,\cdots,m,\\
&\overline{y}_i-y_i\leq\varepsilon_{i,R}, && i=1,\cdots,m,\\
&e_i\geq 0,\quad \varepsilon_{i,L},\varepsilon_{i,R}\geq 0, && i=1,\cdots,m.
\end{aligned}$$

where ω is defined by decision maker. When $0<\omega<1$, the optimal solution of this problem is an effective result for the above MOFLR model [28].

The summarized fuzzy regression models have been presented in Table 1. Examination of this table shows that different versions of Tanaka's model have been developed since emergence of fuzzy regression.

5 Application of time series in energy modeling: literature review

Two basic approaches could be used in energy modeling: time series and regression. Although some variables such as population, gross domestic production (GDP), and price have important effect on electricity consumption, they are not used in this study because of lack of data. Moreover, data for these indicators are usually available annually, and this study concentrates on monthly consumption. On the other hand, exploring the literature shows that using time series in monthly electricity forecasting is a common approach. Some of the most notable works in energy forecasting are discussed and cited next.

Kulshreshtha and Parikh [45] have proposed a vector autoregressive models (2000) to predict the demand for coal in India. The annual time series data from 1970 to 1995 have been used in their study. Time series forecasting model based on Box–Jenkins has been integrated by combined econometric model to forecast monthly peak system load by Uri [47]. Petroleum products demand in India has been forecasted with a translog econometric model based on time series by Rao and Parikh [48]. In their study, various petroleum products have been predicted until 2010. Autoregressive integrated moving average has been used by Gonzales Chavez [49] to forecast energy production and consumption in two European countries (Austria and Spain). Sfetsos [50] has studied intelligent and traditional forecasting methods to

Table 1 Comparison between fuzzy regression models

Regression model	Features
Tanaka et al. [36] (Eq. 1)	The most original model. Assumes a fuzzy linear function. Minimizing fuzziness as an optimal criterion.
Tanaka and Watada [35] (Eq. 5)	Assumes the crisp input and output data, while the relation between the input and output data is defined by a fuzzy function. Each value of the dependant variable can be estimated as a fuzzy number. It uses extension principle. Measure of compatibility between data and a regression model. By applying the extension principle, it derives the membership function of estimated value. The solution of model is x_j-scale dependant and many c_js may be 0.
Tanaka et al. [34] (Eq. 7)	It rectifies the previous problems of Tanaka models. Each h-certain estimated interval is needed to involve the corresponding h-certain observed interval. This affects in large coefficient spreads c_j if any dependent variable has large spreads e_j or if there are outliers.
Özelkan [54], Özelkan and Duckstein [55], Duckstein [29] (Eq. 9)	Capable of solving the problems of FLR with data outliers. Although having several advantages compared with other FLR models, it has some drawbacks such as the central tendency property does not exist fully and explicitly in the model.
Chang and Ayyub [15] (Eq. 10)	The input and output data are crisp but coefficients are triangle fuzzy numbers. Outliers can be modeled by introducing soft boundaries to the model.

forecast mean hourly wind speed. Sfetsos used autoregressive moving average model, adaptive network-based fuzzy inference system (ANFIS), and neural network in the study. Sfetses [51] has applied time series techniques to forecast wind speed. Time series model has been used by Poggi et al [52] to estimate and predict the wind speed in Corsica. Twelve time series models are compared in short-term spot price forecasting of auction-type electricity markets by Weron and Misiorek [56]. Autoregression (AR) models, spike preprocessed, threshold, and semiparametric ARs are used methods. Erzgräber et al. [46] have analyzed the electricity system price of the Nord Pool spot market in which different time scale analysis tools are assessed. Electricity demand in Sri Lanka has been estimated with time series techniques by Amarawickrama and Hunt [44]. They have forecasted the electricity demand until 2025. Azadeh et al. [8, 11] have presented an integrated algorithm consumption based on artificial neural network, computer simulation, and design of experiments for forecasting monthly electrical consumption in Iran from March 1994 to February 2005. They have used time series data in their study. Azadeh et al. [9] have proposed a hybrid ANFIS, computer simulation, and time series algorithm to estimate and predict electricity consumption estimation in Iran from March 1994 to February 2005. An integrated fuzzy regression and time series framework to estimate and predict electricity demand for seasonal and monthly changes in electricity consumption have also been proposed by Azadeh et al. [10].

6 Method: the integrated algorithm

The proposed algorithm uses the most important relative error estimation method which is mean absolute percentage error (MAPE). All error estimation methods except MAPE have scaled output. As input data used for the model estimation have different scales, MAPE method is the preferred method to estimate relative errors.

$$\mathrm{MAPE} = \frac{\sum_{t=1}^{n} \left| \frac{x_t - \hat{x}_t}{x_t} \right|}{n} \tag{18}$$

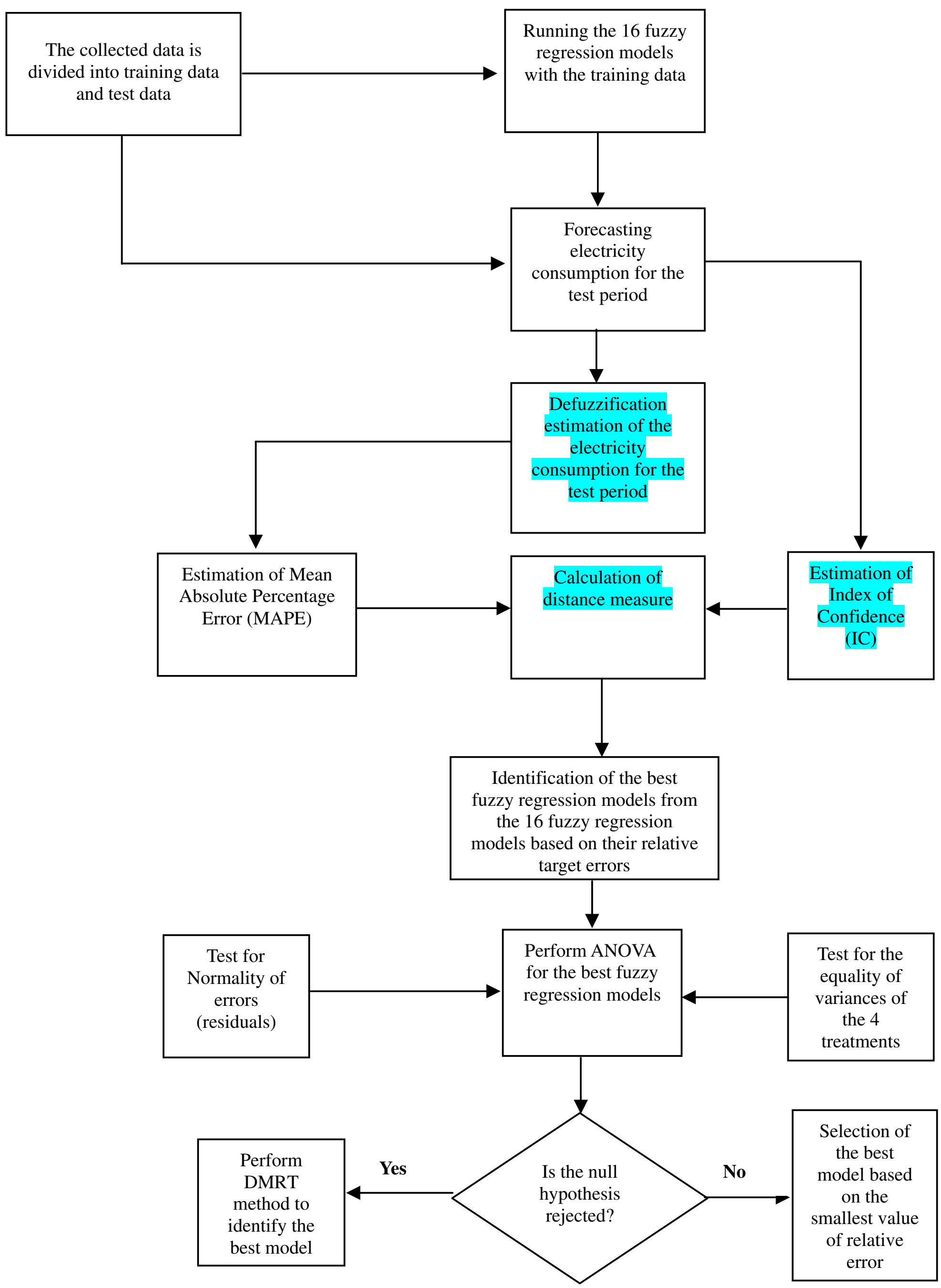
The proposed algorithm may be used to estimate electricity consumption in the future by the selected fuzzy regression model. The collected data are divided into training data and test data. Then, the 16 fuzzy regression models are developed and estimated with the training data set. Next, electricity consumption is forecasted for the test period for the 16 models. Also, a target MAPE is defined. This may be achieved through expert judgment and managerial insights. Furthermore, the best fuzzy regression models are identified from the 16 fuzzy regression models based on the predefined MAPE and IC values. This is because direct conduct of ANOVA for all regression methods may lead to rejection of null hypothesis (that they are equal). Furthermore, selecting the best method becomes a cumbersome and time-consuming task.

The proposed algorithm uses distance measure and ANOVA to select the best fuzzy regression models for electricity estimation and forecasting. Furthermore, if the null hypothesis in ANOVA is rejected, Duncan's multiple range test (DMRT) is used to identify which model is closer to actual data at α level of significance. It also uses MAPE when the null hypothesis in ANOVA is accepted to select the ideal fuzzy regression model. The significance of the proposed algorithm is threefold. First, it is flexible and identifies the best model based on the results of ANOVA and MAPE, whereas previous studies identify the best fitted fuzzy regression model based on relative error results. Second, the proposed algorithm tests the most important fuzzy regression models and consequently identifies the most ideal candidate for electricity consumption, whereas previous studies assume that a typical fuzzy regression model provides the best solutions and estimations. Third, it diagnoses the most well-known fuzzy regression model. Figure 3 depicts the steps of the integrated algorithm of this study. The reader should note that all steps of the integrated algorithm are based on standard and scientific methodologies which are the most advanced fuzzy regression models MAPE, index of confidence (IC), distance measure, and ANOVA. The unique features of the integrated algorithm are discussed in the next sections.

6.1 Defuzzification methods

In some situations, outputs of a fuzzy process need to be converted to their corresponding single scalar values by a procedure. This procedure is called defuzzification. In fact, conversion of a fuzzy quantity to a crisp quantity is a

Fig. 3 The integrated fuzzy regression–ANOVA algorithm for electricity consumption estimation with uncertain data sets



defuzzification process. There are several defuzzification approaches in the literature. In this section, some well-known defuzzification methods which are popular for defuzzifying fuzzy outputs are summarized. Let $\widetilde{A}$ be a fuzzy output, and its support (decision vector) is $x \in X$, where X is the related universe set, and the crisp output (defuzzified value of $\widetilde{A}$) is x^* [31].

- Max-membership**:** the principle of this method is given by the algebraic expression $\mu_{\widetilde{A}(x^*)} \geq \mu_{\widetilde{A}(x)}$ for $x \in X$ [31]. In fact, in the max-membership method, we must find a single point in the support (the range of decision vector) of the fuzzy output ($\widetilde{A}$) which has the highest value of membership function ($\mu_{\widetilde{A}(x)}$). It is clear that in some situations, there may be more than one scalar in the decision vector that has the maximum membership value. In such situations, more than one scalar can be selected as defuzzified value of $\widetilde{A}$.
- Center of area method: the most common interesting defuzzification approach is the center of area, which is expressed by following equation:

  $$x^* = \frac{\int \mu_{\widetilde{A}(x)} x \, d_x}{\int \mu_{\widetilde{A}(x)} \, d_x}$$

  [31].

  In center of area method, at first, the expectation of $\widetilde{A}$ (fuzzy output) in its support should be calculated (the numerator of above equation). After that, the whole area which $\widetilde{A}$ has covered in the universe set (X) must be computed. This computation is shown in the denominator of above equation.
- Mean of maxima method: this approach is the same as the max-membership procedure, but as mentioned above in some situations, there may be more than one scalar in the decision vector that has the maximum membership value. Thus, in max-membership method, more than one scalar may be selected as defuzzified value of $\widetilde{A}$, but this method is able to yield a single scalar point with maximum membership function value [31].

6.2 Index of confidence

For the estimated fuzzy outputs, we can consider the formula below as the total sum of squares which considers the measure of the variation between upper bound and lower bound of the prediction h-certain interval:

$$\mathrm{SST} = \sum_{t=1}^{m} \left(y_i - \left(\sum_{j=0}^{n} \left(\alpha_j - (1-\mathrm{H}) c_j \right) x_{ij} \right) \right)^2 + \sum_{t=1}^{m} \left(\left(\sum_{j=0}^{n} \left(\alpha_j + (1-\mathrm{H}) c_j \right) x_{ij} \right) - y_i \right)^2 . \tag{19}$$

Moreover, the variation of prediction interval with respect to the center regression line can be considered as the regression sum of squares as follows:

$$\mathrm{SSR} = \sum_{t=1}^{m} \left(\left(\sum_{j=0}^{n} \alpha_j x_{ij} \right) - \left(\sum_{j=0}^{n} \left(\alpha_j - (1-\mathrm{H}) c_j \right) x_{ij} \right) \right)^2 + \sum_{t=1}^{m} \left(\left(\sum_{j=0}^{n} \left(\alpha_j + (1-\mathrm{H}) c_j \right) x_{ij} \right) - \left(\sum_{j=0}^{n} \alpha_j x_{ij} \right) \right)^2 . \tag{20}$$

Similarly, the error sum of squares can be formulated as:

$$\mathrm{SSE} = 2 \times \sum_{t=1}^{m} \left(y_i - \left(\sum_{j=0}^{n} \alpha_j \, x_{ij} \right) \right)^2 , \tag{21}$$

where

$$\mathrm{SSE} = \mathrm{SST} - \mathrm{SSR}. \tag{22}$$

Considering the concept of the above formulas, an index of confidence (IC) has been introduced by Wang and Tsaur [38] as:

$$\mathrm{IC} = \frac{\mathrm{SSR}}{\mathrm{SST}} = 1 - \left(\frac{\mathrm{SSE}}{\mathrm{SST}} \right). \tag{23}$$

IC [0, 1] is similar to the determinant confidence (R^2) in classical regression. The higher value of IC (closer to 1) implies a better estimation of y_i.

6.3 Distance measure

Zeleny [41] suggested the "least resistance principle" in multi-criteria problems to select the most satisfactory solution which has the shortest distance to the ideal solution. Since several criteria are mentioned to evaluate the fitness of FLR models, this method is applied to identify the closest FLR model to the ideal solution. As mentioned before, to define the ideal solution, the maximum value of positive criteria and minimum value of negative criteria among all alternatives should be selected. In order to consider the overall distance of all criteria, the distance measures below can be applied:

$$D_p(w) = \sum_{k=1}^{K} \left(w_k^p (1 - d_k)^p \right)^{\frac{1}{p}}, \tag{24}$$

where K is the number of criteria, p is the level of distance where p [1, ∞], and d_k denotes the distance ratio of the current solution of kth criterion to its related optimal (ideal) solution. For positive criteria, d_k can be calculated through formula $d_k = \frac{\text{current solution}}{\text{optimal solution}}$, and for negative criteria, we have $d_k = \frac{\text{optimal solution}}{\text{current solution}}$, where w_k represents the importance of the kth criterion at level p, where we have

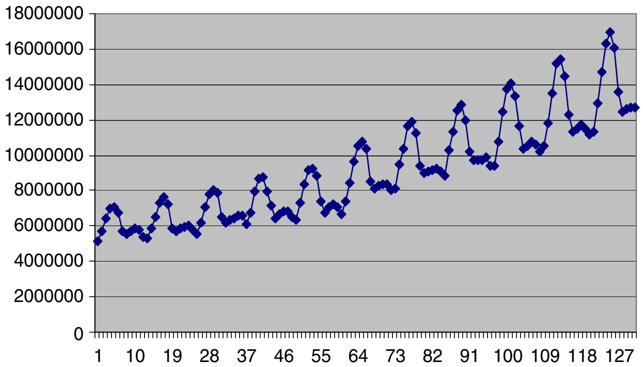


Fig. 4 The trend of raw data

$\sum_{k=1}^{K} w_k = 1$. Several distance measures can be achieved via considering different values for parameter p, which this parameter usually sets to 1, 2, and ∞. For p=1, p=2, and p=∞, the distance measure below can be defined:

$$D_{p=1}(w) = 1 - \sum_{k=1}^{K} (w_k d_k) \qquad p = 1 \tag{25}$$

$$D_{p=2}(w) = \sqrt{\sum_{k=1}^{K} w_k^2 (1 - d_k)^2} \quad p = 2 \tag{26}$$

$$D_{p=\infty}(w) = \max_k \{w_k (1 - d_k)\} \quad p = \infty \tag{27}$$

Therefore, using these distance measures can help us to find a FLR method which is closest to the ideal solution.

6.4 Analysis of variance

The estimated results of fuzzy regression, conventional regression, and actual data are compared by ANOVA F-test. The experiment should be designed such that variability arising from extraneous sources can be systematically controlled. Time is the common source of variability in the experiment that can be systematically controlled through blocking. Therefore, a blocked design of ANOVA may be applied. The hypotheses are:

$$\begin{aligned} &H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 \\ &H_1 : \mu_i \neq \mu_j \quad , i,j = 1,2,3,4; i \neq j. \end{aligned} \tag{28}$$

The reader should note that μ_1 to μ_4 are the average estimation values obtained from the best four fuzzy regression models. If the null hypothesis is accepted, then the preferred model is the one which has the lowest MAPE. Otherwise, if the null hypothesis is rejected, DMRT is used to compare treatment means and to select the preferred model. Furthermore, it is identified as to which pairs caused the rejection of the null hypothesis. To perform DMRT, Eq. 29 is used where $S_{yi.}$ is the mean square error obtained from ANOVA, r_α (p, f) is obtained

Table 2 The estimated MAPE, IC, and distance measure for the 16 fuzzy regression models

Number of models	Fuzzy regression model	MAPE	IC	Distance measure for p=1	Distance measure for p=2	Distance measure for p=∞
1	[36]	0.07759	0.99	0.101	0.101	0.1012
2	[34]	0.06305	0.98	0.014	0.011	0.0093
3	[32]	0.06700	0.98	0.043	0.039	0.0382
4	[32]	0.06280	0.97	0.017	0.012	0.0101
5	[30]	0.06188	0.98	0.005	0.005	0.0051
6	[30]	0.07873	0.97	0.117	0.107	0.1070
7	[37][a]	No feasible solution	–	–	–	–
8	[29]	0.06464	0.01	0.516	0.495	0.4949
9	[22]	0.06520	0.96	0.041	0.030	0.0255
10	[22]	0.06464	0.01	0.516	0.495	0.4949
11	[16]	1.20	0.79	0.575	0.485	0.4742
12	[38]	0.06445	0.99	0.020	0.020	0.0199
13	[15]	0.06600	0.75	0.152	0.125	0.1212
14	[23][b]	Unbounded Solution	–	–	–	–
15	[28]	0.06595	0.95	0.051	0.037	0.0309
16	[28]	0.06620	0.1	0.482	0.451	0.4495

[a] There is no feasible solution for this model in this case study with different rates of parameter P

[b] This model has an unbounded solution

Table 3 The estimated models and coefficients of the 16 fuzzy regression models

Model	Fuzzy regression model	Estimated model	Parameter of models
1	[36]	$\widehat{y}_i = (693467.3577 \quad 0) + (0 \quad 0)x_{i1} + (0.9012 \quad 4.5252)x_{i2}$ [a]	H=0.95
2	[34]	$\widehat{y}_i = (205069.1299 \quad 7429496.8877) + (0 \quad 0)x_{i1} + (0.9879 \quad 0.3770)x_{i2}$	H=0.6
3	[32]	$\widehat{y}_i = (388819.9380 \quad 0) + (0 \quad 0)x_{i1} + (0.9539 \quad 2.1132)x_{i2}$	H=0.95
4	[32]	$\widehat{\Upsilon}_i = (225119.3395 \quad 10027558.6745) + (0 \quad 0)x_{i1} + (0.9879 \quad 0.5027)x_{i2}$	H=0.7
5	[30]	$\widehat{y}_i = (753665.9620 \quad 4548306.5817) + (0 \quad 0)x_{i1} + (0.9425 \quad 0.2742)x_{i2}$	P=873,500,000
6	[30]	$\widehat{y}_i = (0 \quad 8189458.9484) + (0 \quad 0)x_{i1} + (1.0477 \quad 0.0859)x_{i2}$	P=2,000,000,000
7	[37]	No feasible solution	No feasible solution
8	[29]	$\widehat{y}_i = (355994.8508 \quad 0) + (0 \quad 0)x_{i1} + (0.9655 \quad 0)x_{i2}$	$H = 1 U = 10,000,000$
9	[22]	$\widehat{y}_i = (360255.1050 \quad 5852210.0000) + (0 \quad 0)x_{i1} + (0.9631 \quad 0)x_{i2}$	H=0.95
10	[22]	$\widehat{y}_i = (355994.8508 \quad 0) + (0 \quad 0)x_{i1} + (0.9655 \quad 0)x_{i2}$	H=1
11	[16]	$\widehat{y}_i = (6900296.4154 \quad 7362654.7738) + (0.8438 \quad 0.7738)x_{i1} + (0.8465 \quad 0.7764)x_{i2}$	H=0.2
12	[38]	$\widehat{y}_i = (0 \quad 18025118.3045) + (0 \quad 2.1487)x_{i1} + (0.9977 \quad 2.1952)x_{i2}$	H=0.99
13	[15]	$\widehat{y}_i = (92703.0439 \quad 266221.3463) + (0 \quad 0.0182)x_{i1} + (0.9837 \quad 0.1444)x_{i2}$	H=0.1
14	[23]	Unbounded solution	Unbounded solution
15	[28]	$\widehat{y}_i = (448264.9555 \quad 0) + (0 \quad 0)x_{i1} + (0.9525 \quad 0.4192)x_{i2}$	H=0.9
16	[28]	$\widehat{y}_i = (542926.0909 \quad 410079.3011) + (0 \quad 0)x_{i1} + (0.9433 \quad 0)x_{i2}$	$\omega = 0.1 \qquad H = 0.9$

[a] $\widehat{y}_i = \widetilde{A}_0 x_{i0} + \widetilde{A}_1 x_{i1} + \widetilde{A}_2 x_{i2}$ $\widetilde{A} = (\text{centeralvalue} \quad \text{spreadoftrianglefuzzynumber}); x_{i0} = 1$

from Duncan tables, α is the level of significance, and f is the degree of freedom.

$$R_p = r_\alpha(p, f) S_{yi.} \tag{29}$$

7 Experiment: the case study

The proposed algorithm is applied to 130 data, which are the monthly electricity consumption in Iran from April 1992 to February 2004 (Appendix). Next, the data are divided into two sets. One data set is dedicated for estimating the preferred model (referred to as the train data set). The second data set is dedicated for evaluating the validity of the estimated model (referred to as test data set). Usually, train data set contains 70% to 90% of all data, and the remaining data are used for test data set [42]. The 130 monthly data are divided into 118 train data (about 90% of all data) and 12 test data sets (about 10% of all data, 1 year).

This is because the training sample must be relatively large to provide accurate estimates. Second, the test sample is usually chosen around ten to be compared with other methods. Also, in this study, it is shown that n=12 provide robust results for the test period.

Figure 4 shows the trend of raw data. Moreover, the raw data has a trend. As removing the trend is needed for more precise estimation and also there are uncertainties and ambiguousness in the trend of data, fuzzy regression models seem to be the ideal candidates for the given data.

Table 4 ANOVA for the best four fuzzy regression models

Source	Degree of freedom	Sum square	Mean square	F_{value}	P_{value}
Fuzzy regression models (treatment)	3	7.92896E+10	2.64299E+10	10.58	0.0001
Blocks (year)	11	1.71700E+14	1.56090E+13		
Error	33	8.24084E+10	2.49723E+09		
Total	47	1.71861E+14			

8 Results and analysis

In this section, the result of solving each fuzzy regression model will be presented. The target MAPE is defined as 0.065 by expert judgment. Furthermore, only the methods which have MAPE values of lower than 0.065 are selected for ANOVA. The Lingo software was used to solve each of the 16 models. Table 2 presents the calculated MAPE errors, IC, and distance measure for the 16 fuzzy regression methods. Table 3 shows the estimated models and coefficients for the 16 models.

By noting the target MAPE=0.065, models 5, 4, 2, and 12 are selected for ANOVA of the next stage. Moreover, these FLR models have minimum values of distance measures to ideal points based on estimated MAPE and IC. In order to perform the ANOVA, a randomized complete block design, which has four treatments and 12 blocks, is performed (Table 4). Because the null hypothesis is rejected in ANOVA, all pairs of treatment means are compared using DMRT to foresee which pairs resulted in rejection of H_0 (Table 5).

As mentioned in the proposed algorithm, two assumptions need to be tested prior to ANOVA. Moreover, data normality and equality of variances need to be checked. To check whether the probability distribution of data is normal, Anderson–Darling test is applied (See Fig. 5).

Since $P_{\text{value}}=0.02$ is greater than $\alpha=0.01$, we can conclude that the probability distribution of the data is normal. In Fig. 6, residuals versus the fitted values are plotted. Since there is no trend in this diagram, we can conclude that variances are equal. Thus, the randomized complete block design could be performed. The results of this statistical analysis are summarized in Table 4. As mentioned, Table 4 reveals that the null hypothesis is rejected with a P value of 0.0001. Consequently, DMRT needs to be conducted to foresee the reason for rejection. The results of DMRT show that there is significant difference between means of Peters' model and means of other three models (see Table 5).

It can be revealed from Table 5 that Peters, Wang, Sakawa, and Tanaka models do not differ significantly; however, Peters should be chosen for the given set of actual data because we showed it has the smallest MAPE (Table 3). This model provides the smallest relative error and should be the selected model for estimating and forecasting monthly electricity consumption in Iran. The proposed algorithm may be used to estimate electricity consumption in various regions with uncertain and ambiguous data.

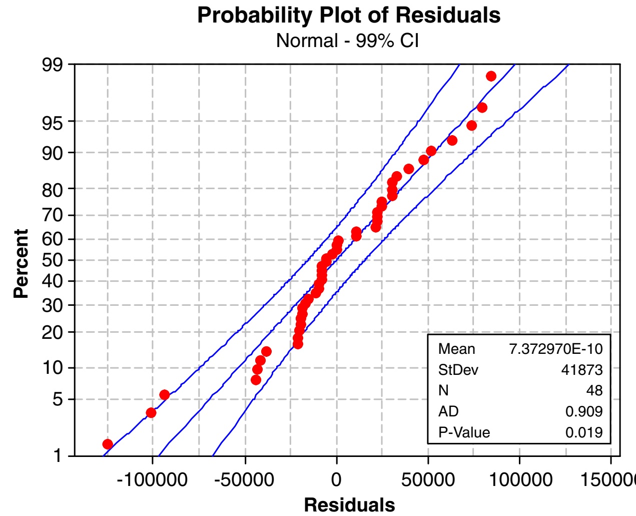


Fig. 5 Probability plot of residuals for normal distribution

9 Conclusion

The proposed algorithm may be used to estimate electricity consumption with ambiguous data by the selected fuzzy regression model. The collected data are divided into training data and test data. Then, the 16 fuzzy regression models are developed and estimated with the training data set. Next, electricity consumption is forecasted for the test period for the 16 models. Furthermore, the best fuzzy regression models are identified from the 16 fuzzy regression models based on their predefined target error (MAPE), IC, and distance measure. The proposed algorithm uses ANOVA to select the best fuzzy regression model (from the selected fuzzy regression methods of last stage based on their target MAPE) for electricity consumption estimation and forecasting. Furthermore, if the null hypothesis in ANOVA is rejected, DMRT is used to identify which model is closer to actual data at α level of significance. It also uses relative error method (MAPE) when the null hypothesis in ANOVA is accepted to select the ideal fuzzy regression. The significance of the proposed algorithm is threefold. First, it is flexible and identifies the best model based on the results of MAPE, IC, distance measure, and ANOVA, whereas previous studies consider the best fitted fuzzy regression model based on relative error results. Second, the proposed algorithm tests the most important fuzzy regression models and consequently

Table 5 The results of DMRT

Pair of models	$S_{\bar{y}_{i0}}$	Difference	R_4	R_3	R_2	Significance
Sakawa versus Wang	14,425.76	93,749	60,011.2	*	*	Yes
Sakawa versus Peters	14,425.76	86,007	*	58,568.6	*	Yes
Sakawa versus Tanaka	14,425.76	20,050	*	*	56,116.2	*
Tanaka versus Wang	14,425.76	73,699	*	58,568.6	*	Yes
Tanaka versus Peters	14,425.76	65,957	*	*	56,116.2	Yes
Peters versus Wang	14,425.76	7,742	*	*	56,116.2	*

* (null)

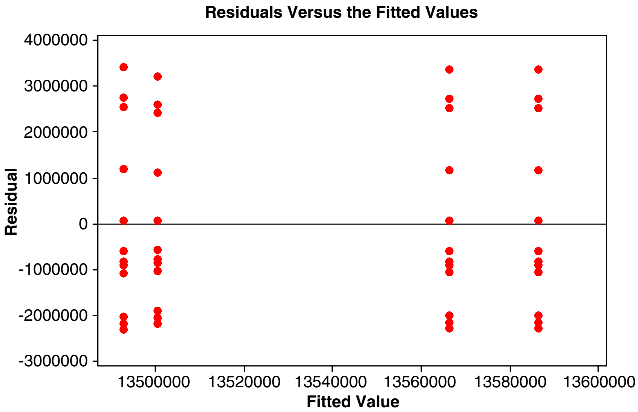


Fig. 6 Residuals versus fitted values to test for equal variances

identifies the most ideal candidate for electricity consumption estimation. Third, it provides a comprehensive assessment with respect to fuzzy regression ANOVA, IC, and MAPE.

As mentioned, the proposed algorithm is composed of 16 fuzzy regression models. This is because there is no clear cut as to which of the recent fuzzy regression model is suitable for a given set of actual and ambiguous data with respect to electricity consumption. Furthermore, it is difficult to model uncertain behavior of electricity consumption estimation with conventional time series, and proper fuzzy regression could be an ideal substitute for such cases. Monthly electricity consumption of Iran from 1992 to 2004 is considered to show the applicability and superiority of the proposed algorithm. We showed that Peters model is the superior model for the case study based on running the 16 fuzzy regression models, distance measure and ANOVA. The proposed algorithm may be used for other real cases to identify the optimum model with lowest error and statistical noise.

Acknowledgement The authors are grateful for the valuable comments and suggestions from the respected reviewers. Their valuable comments and suggestions have enhanced the strength and significance of our paper.

Appendix: Case Study Data

Table 6 Monthly electricity consumption in Iran from 1992 to 2004

1	5,112,500	Train data	66	10,331,172	Train data
2	5,681,521	Train data	67	8,496,857	Train data
3	6,395,367	Train data	68	8,079,716	Train data
4	6,975,157	Train data	69	8,302,567	Train data
5	7,110,100	Train data	70	8,355,663	Train data
6	6,735,385	Train data	71	8,334,044	Train data
7	5,736,692	Train data	72	8,065,338	Train data
8	5,521,386	Train data	73	8,143,074	Train data
9	5,675,558	Train data	74	9,499,280	Train data
10	5,855,284	Train data	75	10,356,647	Train data
11	5,811,723	Train data	76	11,662,210	Train data
12	5,412,274	Train data	77	11,913,777	Train data
13	5,343,047	Train data	78	11,226,582	Train data
14	5,890,343	Train data	79	9,437,329	Train data
15	6,498,841	Train data	80	8,962,981	Train data
16	7,281,385	Train data	81	9,100,005	Train data
17	7,598,830	Train data	82	9,151,585	Train data
18	7,225,359	Train data	83	9,252,917	Train data
19	5,896,872	Train data	84	9,118,551	Train data
20	5,666,248	Train data	85	8,822,002	Train data
21	5,826,754	Train data	86	10,307,161	Train data
22	5,951,504	Train data	87	11,330,227	Train data
23	6,011,416	Train data	88	12,506,292	Train data
24	5,751,556	Train data	89	12,879,931	Train data
25	5,536,390	Train data	90	12,004,267	Train data
26	6,217,736	Train data	91	10,235,825	Train data
27	7,062,810	Train data	92	9,754,226	Train data
28	7,757,156	Train data	93	9,701,765	Train data
29	8,064,807	Train data	94	9,703,783	Train data
30	7,838,773	Train data	95	9,901,824	Train data
31	6,508,814	Train data	96	9,401,967	Train data
32	6,171,428	Train data	97	9,436,829	Train data
33	6,315,113	Train data	98	10,729,817	Train data
34	6,430,886	Train data	99	12,437,144	Train data
35	6,597,923	Train data	100	13,715,487	Train data
36	6,583,064	Train data	101	14,094,341	Train data
37	6,117,896	Train data	102	13,308,948	Train data
38	6,760,571	Train data	103	11,637,033	Train data
39	7,962,063	Train data	104	10,331,066	Train data
40	8,679,644	Train data	105	10,555,722	Train data
41	8,720,364	Train data	106	10,789,612	Train data
42	7,941,711	Train data	107	10,606,500	Train data
43	7,134,022	Train data	108	10,210,820	Train data
44	6,438,431	Train data	109	10,510,355	Train data
45	6,660,092	Train data	110	11,833,496	Train data
46	6,847,702	Train data	111	13,466,023	Train data
47	6,812,296	Train data	112	15,215,272	Train data
48	6,472,784	Train data	113	15,405,851	Train data
49	6,352,603	Train data	114	14,479,378	Train data
50	7,289,726	Train data	115	12,274,190	Train data
51	8,351,912	Train data	116	11,332,188	Train data
52	9,178,630	Train data	117	10,331,172	Train data
53	9,225,747	Train data	118	8,496,857	Train data
54	8,807,836	Train data	119	11,483,643	Test data
55	7,396,819	Train data	120	11,707,494	Test data
56	6,758,709	Train data	121	11,485,285	Test data
57	7,034,106	Train data	122	11,202,808	Test data
58	7,207,062	Train data	123	11,328,070	Test data
59	7,042,722	Train data	124	12,920,315	Test data

60	6,672,616	Train data	125	14,715,280	Test data
61	7,408,035	Train data	126	16,291,332	Test data
62	8,462,697	Train data	127	16,944,113	Test data
63	9,674,038	Train data	128	16,086,172	Test data
64	10,523,297	Train data	129	13,581,415	Test data
65	10,738,586	Train data	130	12,435,256	Test data

References

1. Azadeh A, Ghaderi SF, Anvari M, Saberi M (2007) Performance assessment of electric power generations using an adaptive neural network algorithm. Energy Policy 35:3155–3166
2. Azadeh A, Ghaderi SF, Tarverdian S, Saberi M (2007) Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption. Appl Math Comput 186:1731–1741
3. Azadeh A, Ghaderi SF, Sohrabkhani S (2007) Forecasting electrical consumption by integration of neural network, time series and ANOVA. Appl Math Comput 186:1753–1761
4. Azadeh A, Tarverdian S (2007) Integration of genetic algorithm, computer simulation and design of experiment for forecasting electrical consumption. Energy Policy 35(10):5229–5241
5. Azadeh A, Ghaderi SF, Anvari M, Saberi M (2006a) Measuring performance electric power generations using artificial neural networks and fuzzy clustering. In: Capolino GA, Franquelo LG (eds) Proceedings of the 32nd annual conference of the IEEE Industrial Electronics Society, IECON, Paris, 2006
6. Azadeh A, Ghaderi SF, Anvari M, Saberi M, Izadbakhsh H (2006) An integrated artificial neural network and fuzzy clustering algorithm for performance assessment of decision making units. Appl Math Comput 187(2):584–599
7. Azadeh A, Ghaderi SF, Tarverdian S, Saberi M (2006c) Integration of artificial neural networks and GA to predict electrical energy consumption. In: Capolino GA, Franquelo LG (eds) Proceedings of the 32nd annual conference of the IEEE Industrial Electronics Society–IECON'06, Conservatoire National des Arts & Metiers, Paris, 2006
8. Azadeh A, Ghaderi SF, Sohrabkhani S (2008) A simulated-based neural network algorithm for forecasting electrical energy consumption in Iran. Energy Policy 36(7):2637–2644
9. Azadeh A, Saberi M, Gitiforouz A, Saberi Z (2009) A hybrid simulation-adaptive network based fuzzy inference system for improvement of electricity consumption estimation. Expert Syst Appl 36(8):11108–11117
10. Azadeh A, Saberi M, Seraj O (2010) An integrated fuzzy regression algorithm for energy consumption estimation with non-stationary data: a case study of Iran. Energy 35(6):2351–2366
11. Azadeh A, Saberi M, Ghaderi SF, Gitiforouz A, Ebrahimipour V (2008) Improved estimation of electricity demand function by integration of fuzzy system and data mining approach. Energy Convers Manage 49(8):2165–2177
12. Celmins A (1987) Least squares model fitting to fuzzy vector data. Fuzzy Sets Syst 22:245–269
13. Celmins A (1987) Multidimensional least-squares model fitting of fuzzy models. Math Comput Model 9:669–690
14. Chang PT, Lee ES (1996) A generalized fuzzy weighted least-squares regression. Fuzzy Sets Syst 82:289–298
15. Chang YHO, Ayyub BM (2001) Fuzzy regression methods—a comparative assessment. Fuzzy Sets Syst 119:187–203
16. Chen K, Rys MJ, Lee ES (2006) Modeling of thermal comfort in air conditioned rooms by fuzzy regression analysis. Math Comput Model 43:809–819
17. Chan KY, Kwong CK, Fogarty TC (2010) Modeling manufacturing processes using a genetic programming-based fuzzy regression with detection of outliers. Inf Sci 180(4):506–518
18. Chen LH, Hsueh CC (2009) Fuzzy regression models using the least-squares method based on the concept of distance. IEEE Trans Fuzzy Syst 17(6):1259–1272
19. Chen CW, Wang MHL, Lin JW (2009) Managing target cash balance in construction firms using fuzzy regression approach. Int J Uncertain Fuzziness Knowl-based Syst 17(5):667–684
20. Diamond P (1988) Fuzzy least squares. Inf Sci 46:141–157
21. D'Urso P, Gastaldi T (2000) A least-squares approach to fuzzy linear regression analysis. Comput Stat Data Anal 34:427–440
22. Hojati M, Bector CR, Smimou K (2005) A simple method for computation of fuzzy linear regression. Eur J Oper Res 166:172–184
23. Lai YJ, Chang SI (1994) A fuzzy approach for multi-response optimization: an off-line quality engineering problem. Fuzzy Sets Syst 63:117–129
24. Modarres M, Nasrabadi MM, Nasrabadi E, Mohtashmi GR (2003) Evaluation of fuzzy linear regression models: a mathematical programming approach. In: Proceedings of 4th seminar on fuzzy sets and its applications, Babolsar, pp 129–135.
25. Nather W (1997) Linear statistical inference for random fuzzy data. Statistics 29:221–240
26. Nather W, Albrecht M (1990) Linear regression with random fuzzy observations. Statistics 21:521–531
27. Nather W (2000) On random fuzzy variables of second order and their application to linear statistical inference with fuzzy data. Metrika 51:201–221
28. Nasrabadi MM, Nasrabadi E, Nasrabadi AR (2005) Fuzzy linear regression analysis: a multi-objective programming approach. Appl Math Comput 163:245–251
29. Ozelkan EC, Duckstein L (2000) Multi-objective fuzzy regression: a general framework. Comput Oper Res 27:635–652
30. Peters G (1994) Fuzzy linear regression with fuzzy intervals. Fuzzy Sets Syst 63:45–55
31. Ross TJ (1995) Fuzzy logic with engineering applications. McGraw-Hill, New York
32. Sakawa M, Yano H (1992) Multi-objective fuzzy linear regression analysis for fuzzy input-output data. Fuzzy Sets Syst 47:173–181
33. Savic D, Pedrycz W (1991) Evaluation of fuzzy regression models. Fuzzy Sets Syst 39:51–63
34. Tanaka H, Hayashi I, Watada J (1989) Possibilistic linear regression analysis for fuzzy data. Eur J Oper Res 40:389–396
35. Tanaka H, Watada J (1988) Possibilistic linear systems and their application to the linear regression model. Fuzzy Sets Syst 27:275–289
36. Tanaka H, Uejima S, Asia K (1982) Linear regression analysis with fuzzy model. IEEE Trans Syst Man Cybern 12:903–907
37. Wang HF, Tsaur RC (2000a) Resolution of fuzzy regression model. Eur J Oper Res 126:637–650
38. Wang HF, Tsaur RC (2000b) Insight of a fuzzy regression model. Fuzzy Sets Syst 112:355–369
39. Wu HC (2003) Fuzzy estimates of regression parameters in linear regression models for imprecise input and output data. Comput Stat Data Anal 42:203–217
40. Zadeh LA (1975) The concept of linguistic variable and its application to approximate reasoning I, II and III. Inf Sci 8:199–249, 8, 301–357 and 9, 43–80
41. Zeleny M (1982) Multiple criteria decision making. McGraw-Hill, New York
42. Zhang GP (2001) An investigation of neural networks for linear time-series forecasting. Comput Oper Res 28:1183–1202
43. Narayan PK, Narayan S, Prasad A (2008) A structural VAR analysis of electricity consumption and real GDP: Evidence from the G7 countries. Energy Policy 36:2765–2769
44. Amarawickrama HA, Hunt LC (2008) Electricity demand for Sri Lanka: a time series analysis. Energy 33:724–739

45. Kulshreshtha M, Parikh JK (2000) Modelling demand for coal in India: vector autoregressive models with cointegrated variables. Energy 25:149–168
46. Erzgräber H, Strozzi F, Zaldívar JM, Touchette H, Gutiérrez E, Arrowsmith DK (2008) Time series analysis and long range correlations of Nordic spot electricity market data. Physica A 387:6567–6574
47. Uri ND (1978) Forecasting peak system load using a combined time series and econometric model. Appl Energy 4:219–227
48. Rao RD, Parikh JK (1996) Forecast and analysis of demand for petroleum products in India. Energy Policy 24:583–592
49. Gonzales Chavez S, Xiberta Bernat J, Llaneza Coalla H (1999) Forecasting of energy production and consumption in Asturias (Northern Spain). Energy 24:183–198
50. Sfetsos A (2000) A comparison of various forecasting techniques applied to mean hourly wind speed time series. Renew Energy 21:23–35
51. Sfetses A (2002) A novel approaches for the forecasting of mean hourly wind speed time series. Renew Energy 27:163–174
52. Poggi P, Muselli M, Notton G, Cristofari C, Louche A (2003) Forecasting and simulating wind speed in Corsica by using an autoregressive model. Energy Convers Manage 44:3177–3196
53. Körner R, Näther W (1998) Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least-squares estimates. Inform Sci, 109, pp 95–118
54. Özelkan EC (1997) Multi-objective fuzzy regression applied to the calibration of conceptual rainfall–runoff models, Unpublished PhD Dissertation, Department of Systems and Industrial Engineering, The University of Arizona
55. Özelkan EC, Duckstein L (2000) Multiobjective fuzzy regression: a general framework. Comput Oper Res 27, pp 635–640
56. Weron R, Misiorek A (2008) Forecasting spot electricity prices: a comparison of parametric and semiparametric time series models. Int J Forecasting 24(4):744–763