A New ANN Approach for Time Series Analysis

Najmeh Neshat[a*], Hashem Mahlooji[b], Murat Kaya[c]

[a]Department of Industrial Engineering, Meybod University, Iran
[b]Industrial Engineering Department, Sharif University of Technology, Iran
[c]Industrial Engineering Program, Sabanci University, Turkey

Abstract:

Time series analysis and accurate forecasting of energy prices are critical for both policymakers and market participants. In the practical analysis of price time series, the coefficients play vital roles; however, their accurate estimation is a challenging issue as they are affected by external factors. This study proposes a new modeling approach for Artificial Neural Networks (ANNs) models based on fuzzy logic. For this purpose, we reformulated an ANN model as a fuzzy nonlinear regression model to capture the advantages of both fuzzy regression and ANN methodologies. This clear-box model can be applied to not only uncertain, ambiguous, and complex environments, but it is also capable of modeling nonlinear patterns. To illustrate the capability of the proposed approach, we report a case study of liquefied natural gas (LNG) prices in Japan's market (the world's largest natural gas importer). The results support that the performance of the proposed approach is acceptable; moreover, it can deal with uncertain and complex environments as a clear-box model.

***Key words**: Time series; Natural gas price; Artificial Neural Networks (ANNs); Fuzzy logic*

1. Introduction

Analysis and forecasting of energy prices is at the heart of energy management, which enables policymakers to manage many frequent tactical decisions such as balancing energy demand and risk management. Price fluctuations influence the distribution and flow of various resources in the energy market and have substantial economic leverage; however, it is difficult to forecast price time series accurately due to their complicated and uncertain nature [1]). Accordingly, energy price analysis and forecasting has been the center of attention from both academic and practical points of view.

Different time series forecasting approaches, from statistical to computational intelligence have been applied in the past [2]. Based on the number of techniques, the related literature is typically divided into two main categories: stand-alone and hybrid methods, consisting of one and more than one technique, respectively. Based on the type of the underlying technique, stand-alone methods can be divided into three classes, namely Statistical, Causal and Computational Intelligence methods. Statistical methods model the dynamic relationship between lagged values of determinants and the forecasted price based on historical data. Examples include autoregressive (AR) and double seasonal Holt-Winter (DSHW) models [3], autoregressive with exogenous inputs (ARX) models [4], threshold ARX (TARX) models [5], generalized autoregressive conditional heteroscedasticity (GARCH) based models [6]–[9], autoregressive integrated moving average (ARIMA) models [10], [11], semi/nonparametric models [12], Seasonal autoregressive integrated moving average (SARIMA) [13], transfer function (TF) models [6], and Grey models [14]–[16]. Certain hybrid versions of the mentioned methods are also suggested, e.g., wavelet-based models [14], [17].

Causal methods focus on formulating the dynamic relationships between causal variables (determinants) and the forecasted energy price. These models may utilize the least-squares method[1] to determine the forecasted energy price in terms of its determinants and lagged data [18]. Different causal methods, e.g., linear regression (LR) [19], nonlinear regression (NLR) [20]–[23] and Logistic or logit regression (LoR) [24] have been widely utilized in the literature.

Energy price forecasting, similar to many real-life problems, has a nonlinear nature. Statistical and causal models, on the other hand, are usually linear in nature [25]; hence they may

[1] For further information, see [10].

not perform well with financial data considering some features such as leptokurtosis, volatility clustering and leverage effects. In addition, the nonlinear behavior of energy price might become too complicated to predict [25]. However, statistical and causal models offer the advantage of having a clear internal logic; this is why they are referred to as "white-box"[2] methods.

Computational Intelligence methods such as Artificial Neural Network (ANN) and Support Vector Machine (SVM) are widely used in this research area due to their ability in handling the hidden features of data as well as their nonlinear modeling capability (e.g., [18], [26]–[28]. ANNs are examples of flexible regression approaches but they are different from the standard methods in some aspects such as: 1) they do not require a prior assumption of the model form in the model building process, 2) they provide better solutions for modeling complex non-linear relationships compared to the parametric models, and 3) they have high robustness when the scope of the appraisal is widened to include aspects such as outliers, non-linearity and other kinds of dependence among data. However, ANNs are not capable of handling the problems caused by uncertain situations, and they are identified as “black box” techniques. Uncertain conditions often arise due to the rapid development of new technologies, or due to imprecise and inadequate data in energy pricing.

Fuzzy regression models are suitable to address such uncertainties, however, these models do not map the function with nonlinear behavior. A number of studies have used fuzzy time series to forecast energy prices in Australia and Singapore electricity markets [29], and Ontario and New England markets [30]; however, recent studies indicate that hybrid models of ANN and fuzzy systems outperform standalone fuzzy regression models [25], [31], [32]. These models of ANNs and fuzzy systems correspond to a fuzzy model of Takagi–Sugeno, wherein the weights of the neural network model are similar to the parameters of the fuzzy system, and hence, they behave as black box models. On the other hand, in clear-box models in which the coefficients are identified, in addition to the magnitude of the correlation between a certain response variable and the determinants, one can also infer the effects of each response variable.

In the light of the above-mentioned considerations, we propose a reformulated ANN model as a nonlinear fuzzy time series capable of dealing with uncertain situations and nonlinear functions. The proposed model brings together the advantages of fuzzy linear regression and

[2] "White-box models are the type of models which one can clearly explain how they behave, how they produce predictions and what the influencing variables are."

ANNs models, while mitigating the shortcomings of these two models. As we illustrate with a case study on the Japanese LNG market price, the flexibility of the model allows it to be applied to uncertain, ambiguous, or complex environments.

The remainder of this paper is organized as follows. In the remaining subsections of Section 1, the techniques embedded in the proposed approach are explained. Section 2 introduces the proposed modeling approach and Section 3 is devoted to experiments. Conclusions and final remarks are presented in Section 4.

1.1. Regression Modelling

Generally, let x_{mn} and y_m represent the n^{th} independent variable and the dependent variable, respectively, in the m^{th} observation, β_i be the parameter associated with the i^{th} independent variable (i=1,2,.., n), and ε_m be the error term associated with the m^{th} observation. The classical regression model can be stated as Equation 1:

$$y_m = \beta_0 + \beta_1 x_{m1} + \ldots.. + \beta_n x_{mn} + \varepsilon_m \qquad m = 1,2,\ldots,k \tag{1}$$

The regression parameter β_i must usually be estimated from sample data. Let β_i be estimated under uncertainty in the adequacy of independent variables or sample data and $X_0 = 1$, then β_i falls into the category of fuzzy regression analysis as Equation 2:

$$\tilde{Y} = \tilde{\beta}_0 + \tilde{\beta}_1 X_1 + \tilde{\beta}_2 X_2 + \ldots. + \tilde{\beta}_n X_n = \sum_{i=0}^{n} \tilde{\beta}_i X_i = X'\tilde{\beta} \tag{2}$$

where $\tilde{Y}$ is the fuzzy output, $X' = [X_0, X_1, X_2, \ldots, X_n]^T$ is the real-valued input vector, and $\tilde{\beta} = \{\tilde{\beta}_0, \tilde{\beta}_1, \ldots., \tilde{\beta}_n\}$ is a set of fuzzy numbers which we want to determine according to some criteria of goodness of fit. In the form of triangular fuzzy numbers, for $\tilde{\beta}_i$, we have Equation 3 as follows:

$$\mu_{\tilde{\beta}_i} = \begin{cases} 1 - \dfrac{|\alpha_i - \beta_i|}{c_i} & \text{if } \alpha_i - c_i \le \beta_i \le \alpha_i + c_i \\ 0 & \text{Otherwise;} \end{cases} \tag{3}$$

where $\mu_{\tilde{\beta}_i}$ is the membership function[3] of $\tilde{\beta}_i$, α_i is the center of the fuzzy number, and c_i is the width or spread around the center of the fuzzy number. According to the extension principle, the membership function of the fuzzy numbers $\tilde{y}_m = x_m^{'}\tilde{\beta}$ can be indicated using pyramidal fuzzy parameter β as Equation 4 [4] :

$$\mu_{\tilde{Y}} = \begin{cases} 1 - \dfrac{|y_m - x_m.\alpha|}{c'|x_m|} & \text{for } x_m \neq 0 \\ 1 & \text{for } x_m = 0; y_m = 0 \\ 0 & \text{for } x_m = 0; y_m \neq 0 \end{cases} \qquad (4)$$

where α and c denote vectors of model values and spreads, all model parameters, respectively. The problem of finding $\tilde{\beta}_i$ is formulated as a linear programming problem (Equation 5)[5]:

Minimize $S = \sum_{m=1}^{k} c'|x_m|$

$$\text{Subject to} \begin{cases} x'_m \cdot \alpha + (1-h)c'|x_m| \geq y_m \\ x'_m \cdot \alpha - (1-h)c'|x_m| \leq y_m \\ c \geq 0 \end{cases} \qquad (5)$$

where the objective function is minimizing the total vagueness (S), which defined as the sum of individual spreads of the fuzzy parameters of the model, k is the number of observations, and h is the threshold[6] of the membership function of $\tilde{Y}$.

1.2. Neural Network Models

[3] This member function is generally developed according to [33]; however, here m= {1, 2, … k}.
[4] A more general definition and further details can be found in [34].
[5] For the first time, [34] developed an LP to find pyramidal fuzzy parameters.
[6] Fuzzy thresholds support only multi-valued logic. Mathematical functions that yield sigmoidal curves can be used to implement fuzzy thresholds. This approach allows establishing points on continuum where decisions are triggered. Basic operations and functions on fuzzy logic can be found in [35].

Neural Network (NN) models are a class of flexible nonlinear models that can find patterns adaptively from data. Through processing experimental data, these systems transfer the knowledge or rules hidden in data to the network's structure. They can predict a phenomenon's future by taking into account its history.

The relationship between the output y_m and the inputs $x_{m1}, x_{m2}, ... x_{mn}$ in a neural network model has the following mathematical representation as Equation 6:

$$y_m = f(W_0 + \sum_{q=1}^{Q} w_{i,q}.g(W_q + \sum_{i=1}^{n} w_{i,q}.x_{mi})) = f(W_0 + \sum_{q=1}^{Q} w_{i,q}.X_{mq}) = f(\sum_{q=1}^{Q} W_q.X_{mq}) \qquad (6)$$

where $w_{i,q}(i = 1,2,....,n; q = 1,2,3.,...,Q)$ and $W_q(q = 0,1,2,....,Q)$ are model parameters called connection weights; n is the number of input nodes; $X_{mq} = g(W_q + \sum_{i=1}^{n} w_{i,q}.x_{mi})$; Q is the number of hidden nodes, and f and g are the transfer functions that are often used as a logistic function[7]. In this sense, the neural network is equivalent to a nonlinear multiple regression model. In practice, a network consisting of one hidden layer that has a small number of hidden nodes generally works well in out-of-sample forecasting [35]. This may be due to the overfitting effect typically found in the neural network modeling process. An over-fitted model has a good fit to the sample used in the model building process but has poor ability of generalization of data out of the sample[35].

To improve the performance of artificial neural networks, certain data mining approaches are often applied in the network training process, namely, Dynamic, Multiple, Prune, Exhaustive prune data-mining methods, as well as Networks with Radial Basis Function (RBFN) [36][8].

Dynamic Method: In this method, first an initial topology for the network is created and then during the training process this topology is modified by adding or deleting hidden parts of the network. This method is very suitable when there are nonlinear relationships between variables. Under these conditions, the dynamic model has a great deal of flexibility in regression applications (for prediction) to explain nonlinear behaviors of response variable. In addition, the performance of this model is evaluated in situations where the observations are time series.

---

[7] Transfer functions are used to model the behavior of intended functions and can be chosen based on LR or trial and error.

Multiple Method: This method first creates several different topologies of possible scenarios (although their number depends on the data selected for training) and puts each of them in parallel in the training process. After the training phase, the model with the least squared errors is chosen as the final model.

Prune Method: The main idea of this type of network is that the efficiency of the network can be increased by eliminating some connections between neurons (weights). These networks consider a selection criterion, which is usually "the inverse of the network error" on a portion of the training data. If this value is increased, it indicates the correct progress of the network, otherwise the network is growing to no avail and the algorithm stops. On the other hand, after adding new network neurons using this selection criterion, one of the weights is removed if this value is increased, which indicates that the decision is correct, and otherwise the weight is returned. This method usually takes a long execution time, but improves the model by removing redundant variables and results in increased model generalization.

Exhaustive Prune Data-mining Method: Similar to the prune method, the process begins with a very large network and then the weakest units are removed from the input and hidden layers during the training process.

2. Formulating the Proposed Model

In this section, we explain how we reformulate an ANN model as a fuzzy nonlinear regression model that is capable of dealing with uncertain situations and nonlinear functions. This proposed model brings together the advantages of both fuzzy regression and ANN models while their mentioned deficiencies are decreased.

The steps are summarized as follows:

Step 1. Replace the crisp parameters $w_{i,q}(i = 0,1,2,....,n; q = 1,2,3.,...,Q)$ and $W_q(q = 0,1,2,....,Q)$ with fuzzy parameters in the form of triangular fuzzy numbers $\tilde{w}_{i,q}(i = 0,1,2,....,n; q = 1,2,3.,...,Q)$ and $\tilde{w}_q(q = 0,1,2,....,Q)$ according to Equation (2) as Equation 7:

$$\tilde{y}_m = f(\tilde{w}_0 + \sum_{q=1}^{Q} \tilde{w}_q . g(\tilde{w}_{0,q} + \sum_{i=1}^{n} \tilde{w}_{i,q} . x_{mi})) = f(\tilde{w}_0 + \sum_{q=1}^{Q} \tilde{w}_q . \tilde{X}_{mq}) = f(\sum_{q=0}^{Q} \tilde{W}_q . \tilde{X}_{mq}) \quad (7)$$

Step 2. Determine the membership function of the fuzzy parameters $\tilde{w}_{i,q} = (a_{iq}, b_{iq}, c_{iq})$ and $\tilde{w}_q = (d_q, e_q, f_q)$ in the form of triangular fuzzy numbers as shown in Equations 8 and 9, respectively[9]:

$$\mu_{\tilde{w}_{i,q}} = \begin{cases} \dfrac{(w_{iq} - a_{iq})}{(b_{iq} - a_{iq})} & \text{If } a_{iq} \le w_{i,q} \le b_{iq} \\ \dfrac{(w_{iq} - c_{iq})}{(b_{iq} - c_{iq})} & \text{If } b_{iq} \le w_{i,q} \le c_{iq} \\ 0 & \text{Otherwise} \end{cases} \quad (8)$$

$$\mu_{w_q} = \begin{cases} \dfrac{(w_q - d_q)}{(e_q - d_q)} & \text{If } d_q \le w_q \le e_q \\ \dfrac{(w_q - f_q)}{(e_q - f_q)} & \text{If } e_q \le w_q \le f_q \\ 0 & \text{Otherwise} \end{cases} \quad (9)$$

Step 3. Calculate membership function of $\tilde{X}_{mq}$ in terms of $\mu_{\tilde{w}_{i,q}} (i = 0,1,2,...., n)$; $(q = 0,1,2,...., Q)$ and $\mu_{\tilde{w}_q} (q = 0,1,2,...., Q)$ through applying the extension principle [38]; [39] given as Eq. (5) (Equation 10):

$$\mu_{\tilde{X}_{mq}} = \begin{cases} \dfrac{X_{mq} - \sum_{i=0}^{n} a_{iq} \cdot x_i}{\sum_{i=0}^{n} b_{iq} \cdot x_i - \sum_{i=0}^{n} a_{iq} \cdot x_i} & \text{If } \sum_{i=0}^{n} a_{iq} \cdot x_i \le X_{mq} \le \sum_{i=0}^{n} b_{iq} \cdot x_i \\ \dfrac{X_{mq} - \sum_{i=0}^{n} c_{iq} \cdot x_i}{\sum_{i=0}^{n} b_{iq} \cdot x_i - \sum_{i=0}^{n} c_{iq} x_i} & \text{If } \sum_{i=0}^{n} b_{iq} \cdot x_i \le X_{mq} \le \sum_{i=0}^{n} c_{iq} \cdot x_i \\ 0 & \end{cases} \quad (10)$$

[9] a<b<c are triangular fuzzy numbers and membership(b)=1, while *b* need not be in the "middle" of *a* and *c*. Basic operations and functions on these fuzzy logic can be found in [37].

Otherwise;

The transfer functions (*f* and *g*) are assumed to be saturated linear transfer functions (Satlin), and the connected weights between the input and the hidden layers are considered to be of a crisp form for simplicity.

Step 4. Similar to Step 3, the membership function of $\tilde{y}_m$ is given as Equation 11 and is modified considering Equation 4 as Equation 12:

$$\mu_{\tilde{y}_m} = \begin{cases} \dfrac{y_m - \sum_{q=0}^{Q} d_q \cdot X_{mq}}{\sum_{q=0}^{Q} e_q \cdot X_{mq} - \sum_{q=0}^{Q} d_q \cdot X_{mq}} & \text{If } \sum_{q=0}^{Q} d_q \cdot X_{mq} \le y_m \le \sum_{q=0}^{Q} e_q \cdot X_{mq} \\ \dfrac{y_m - \sum_{q=0}^{Q} f_q \cdot X_{mq}}{\sum_{q=0}^{Q} e_q \cdot X_{mq} - \sum_{q=0}^{Q} f_q \cdot X_{mq}} & \text{If } \sum_{q=0}^{Q} e_q \cdot X_{mq} \le y_m \le \sum_{q=0}^{Q} f_q \cdot X_{mq} \end{cases} \tag{11}$$

$$\mu_{\tilde{y}_m} = \begin{cases} 1 - \dfrac{\left| y_m - \sum_{q=0}^{Q} e_q \cdot X_{mq} \right|}{\sum_{q=0}^{Q} c_q \cdot \left| X_{mq} \right|} & \text{for } X_{mq} \neq 0 \\ 0 & \text{Otherwise;} \end{cases} \tag{12}$$

Step 5: Finally, the criteria of minimizing total vagueness, *S* (the sum of individual spreads of the fuzzy parameters) is used according to Equation 5 to develop the linear programming problem as follows[10] :

Minimize $\quad S = \sum_{m=1}^{k} \sum_{q=0}^{Q} (e_q - d_q) \cdot \left| X_{mq} \right|$

[10] *h* is the threshold of the membership function of $\tilde{Y}$

$$
\text{Subject to}\begin{cases} \sum_{q=0}^{Q} e_q.X_{mq} + (1-h)(\sum_{q=0}^{Q}(e_q - d_q)\cdot\left|X_{mq}\right|) \geq y_m & m = 1,2,....,k \\ \sum_{q=0}^{q} e_q.X_{mq} - (1-h)(\sum_{q=0}^{Q}(e_q - d_q)\cdot\left|X_{mq}\right|) \leq y_m & q = 0,1,2,....,Q \end{cases} \tag{13}
$$

3. Experiments

In recent decades, the popularity of natural gas has increased considerably because it causes no local pollutants (that is, being a “clean fuel”), and because it generates relatively lower carbon emissions compared to coal. Rapid rise in global gas demand has turned the gas markets from a “buyer’s market” into a “seller’s market”. Therefore, the study and forecasting of gas prices based upon actual data and market information, using efficient modeling and forecasting approaches has become a timely and important area of study.

In this section, we study the Liquified Natural Gas (LNG) price in the Japanese market to test our proposed forecasting model. Japan is the world’s largest LNG importer and one of the three largest gas markets worldwide. Nonlinearity and complexity of the LNG price function and the uncertainty due to lack of data on influencing variables cause LNG price forecasting to be a suitable candidate for using the fuzzy logic approach [40].

Based on the related literature [18], [41]–[44], 15 direct and indirect (auxiliary) determinants were identified which are provided in Table 1 together with their statistical summary information. The related data was collected from February 2008 to February 2020 in monthly terms from International Energy Consortium (IEC) database[11]. The data from 2008 to 2017 was used for training and the rest of the dataset was used for testing. Outliers was removed from the training data set.

To estimate the weights of the related ANN model, five well-known data mining approaches (dynamic, multiple, prune, exhaustive prune, and RFBN) were applied. To reach the best network configuration using the Neuro Solution toolbox in MATLAB software, different networks including 1-3 hidden layers, variable number of neurons in hidden layers, and different activation functions were tested. Finally, the network with the least Mean Square Error (MSE) for each approach was selected as the optimum network.

[11] https://www.energy.gov/

The values of the training termination criteria are: $1*E(-5)$ for Mean Square Error; 200 for the maximum number of each epoch; and $1*E(-10)$ for the minimum error gradient of each epoch. Table 2 provides further information to compare the relative importance of the input factors for each neural network model.

The values of performance indicators Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and Absolute Fraction of Variance (R^2) for the five networks are provided in Table 3, and one of these models is chosen to examine the proposed approach.
where MAPE shows the mean ratio between the error and the actual values are calculated as Equation 14:

$$MAPE = \frac{1}{m}\sum_{i=1}^{m}\left|100\frac{(y_i - \hat{y}_i)}{y_i}\right| \tag{14}$$

where $\hat{y}_i$ is the predicted value by the ANN model y_i is the actual value of the response process and m is the number of points in the data set. The RMSE is calculated by Equation 15:

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2} \tag{15}$$

Finally, R^2 or the coefficient of determination as a statistical criterion which can be applied to multiple regression analysis, is calculated by Equation (16):

$$R^2 = 1 - \left(\frac{\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{m}(y_i - \bar{y})^2}\right) \tag{16}$$

R^2 takes a value between zero and one, where values closer to one indicate a good fit.

Exhaustive prune network is chosen with 4 input variables which are "Brent oil price", “mean oil price of Oman and Dubai”, “gasoline price”, and “mean price of Japan’s imported petroleum”. As “Brent oil price” and “mean price of Japan’s imported oil” variables have a correlation coefficient of 67.8%, “Brent oil price” was eliminated and the variable of “the related year” was added to the input variables considering the performance of the network.

The best configuration of the Exhaustive Pruned Network is determined to be $N^{(4-4-1)}$ which is shown in Figure 1.[12] The weights and biases of the mentioned network are given in Table 4.

As mentioned, the input weights and biases of the Exhausted Prune Model are considered of a crisp form for simplicity. To obtain the center and the interval width for hidden weights (i.e., c_q and e_q) as unknown parameters with the assumption of *h*=0.5, the proposed method is formulated as the following linear programming problem:

Minimize $S = \sum_{m=1}^{70}\sum_{q=0}^{4} c_q |X_{mq}|$

Subject to $\begin{cases} \sum_{q=0}^{4} e_q.X_{mq} + (0.5)(\sum_{q=0}^{4} c_q |X_{mq}|) \geq y_m \\ \sum_{q=0}^{4} e_q.X_{mq} - (0.5)(\sum_{q=0}^{4} c_q |X_{mq}|) \leq y_m \end{cases}$ $\quad m = 1,2,....,70 \quad q = 0,1,2,....,4$

where *m* is the number of trained data and *q* is the number of neurons in the hidden layer of Exhaustive prune networks. The center and width of hidden weights (c_q and e_q) obtained using GAMS Software are presented in Table 5.

To evaluate the performance of the proposed model, the center of hidden weights (c_q) were considered as weights (W_q) for calculating y_m according to the Equation 17 using the test data.

$$y_m = f(0.478 + \sum_{q=1}^{4} c_q.X_{mq}) = Satlin(\sum_{q=0}^{4} c_q.X_{mq}) = \sum_{q=0}^{4} c_q.X_{mq} \tag{17}$$

The results of performance evaluation of the proposed models using 30 test data are provided in Table 6.

Note that considering the center of hidden weights (c_q) leads to the crisp y_m[13].

[12] In a $N^{(x-y-z)}$ network topology, x, y and z stand for the first, the second and the third layer`s neurons, respectively. $N^{(4-4-1)}$ refers to a network consisting of 4 neurons for the input layer, 4 neurons for the hidden layer, and one neuron for the output layer.

Comparing the values of performance indicators in Table 6, we observe the proposed model to achieve an improvement in R^2 of 0.26%, a decrease in MAPE of 3%, and a decrease in RMSE of 2 USD compared with the Exhausted Prune ANN Model. These results validate that reformulating the ANN model based on the fuzzy regression model tends to improve the performance of the ANN model. The ANN model, in turn, outperforms both the linear and the fuzzy regression models. The proposed model also offers the advantage of addressing uncertainty and complexity as a clear box model.

4. Conclusion and Final Remarks

Emerging computational intelligent approaches has led to significant advances in forecasting energy price time series. In this regard, Artificial Neural Network (ANN) models have been extensively used; however, these models are identified as black box techniques and they are not capable of handling problems that involve uncertainty. Fuzzy regression models, on the other hand, are a good choice for addressing uncertainty; however, they cannot map the function with nonlinear behavior.

In this study, we reformulated an ANN model as a fuzzy nonlinear regression model that can deal with both uncertain situations and nonlinear functions. This novel modeling approach can be applied to uncertain, ambiguous, and complex environments due to its flexibility. As a case study, the price function of liquefied natural gas (LNG) in the Japanese market was formulated based on the proposed modeling approach. The performance of the proposed model was compared to those of conventional approaches such as the selected neural network, fuzzy regression, and linear auto regression models. Based on a number of performance indicators, the proposed model is shown to have the best performance.

To add up, the proposed approach is applicable to complex, nonlinear and ambiguous environments due to embedded ANN and fuzzy mechanisms. It is robust to inconsistency and noise in data as well as the presence of high dimensionality and collinearity. Besides, it performs as a clear box model and provides higher generalization capability. The ability to deal with

---

[13] Satlin is a neural transfer function. This transfer function calculates a layer's output from its net input. It clips the input to [-1, 1]. This function is chosen based on trial and error.

uncertain, limited, and non-crisp data would allow this approach to be preferred over conventional alternatives.

References:

[1] Chen, M.-Y. and Chen, B.-T. “A hybrid fuzzy time series model based on granular computing for stock price forecasting,” *Information Sciences*, vol. 294, pp. 227–241, Feb. 2015, doi: 10.1016/j.ins.2014.09.038.

[2] Lu, H. Ma, X. Ma, M. and Zhu, S. “Energy price prediction using data-driven models: A decade review,” *Computer Science Review*, vol. 39, p. 100356, Feb. 2021, doi: 10.1016/j.cosrev.2020.100356.

[3] Ryu, S. Noh, J. and Kim, H. “Deep Neural Network Based Demand Side Short Term Load Forecasting,” *Energies*, vol. 10, no. 1, pp. 1–20, 2016.

[4] Ziel, F. and Steinert, R. “Probabilistic mid- and long-term electricity price forecasting,” *Renewable and Sustainable Energy Reviews*, vol. 94, pp. 251–266, Oct. 2018, doi: 10.1016/j.rser.2018.05.038.

[5] Lago, J. De Ridder, F. and De Schutter, B. “Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms,” *Applied Energy*, vol. 221, pp. 386–405, Jul. 2018, doi: 10.1016/j.apenergy.2018.02.069.

[6] Nogales, F.J. and Conejo, A.J. “Electricity price forecasting through transfer function models", Journal of the Operational Research Society, 57(4), 350-356.

[7] Keynia, F. and Heydari, A. “A new short-term energy price forecasting method based on wavelet neural network,” *IJMOR*, vol. 14, no. 1, p. 1, 2019, doi: 10.1504/IJMOR.2019.096975.

[8] Kumar, V. Singh, N. Singh, D. and Mohanty, S. “Short-Term Electricity Price Forecasting Using Hybrid SARIMA and GJR-GARCH Model,” in *Networking Communication and Data Knowledge Engineering*, Singapore, 2018, pp. 299–310. doi: 10.1007/978-981-10-4585-1_25.

[9] Yang, Z. Ce, L. and Lian, L. “Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods,” *Applied Energy*, vol. 190, pp. 291–305, Mar. 2017, doi: 10.1016/j.apenergy.2016.12.130.

[10] Lin, L. Jiang, Y. Xiao, H. and Zhou, Z. “Crude oil price forecasting based on a novel hybrid long memory GARCH-M and wavelet analysis model,” *Physica A: Statistical Mechanics and its Applications*, vol. 543, p. 123532, Apr. 2020, doi: 10.1016/j.physa.2019.123532.

[11] Zhao, Z. Wang, C. Nokleby, M. and Miller, C. “Improving Short-Term Electricity Price Forecasting Using Day-Ahead LMP with ARIMA Models,” Jan. 2018.

[12] Trespalacios, A. Cortés, L. and Perote, J. “Uncertainty in electricity markets from a semi-nonparametric approach,” *Energy Policy*, vol. 137, p. 111091, Feb. 2020, doi: 10.1016/j.enpol.2019.111091.

[13] Zhang, J. Tan, Z. and Wei, Y. “An adaptive hybrid model for short term electricity price forecasting,” *Applied Energy*, vol. 258, p. 114087, Jan. 2020, doi: 10.1016/j.apenergy.2019.114087.

[14] Chen, Y. Zhang, C. He, K. and Zheng, A. “Multi-step-ahead crude oil price forecasting using a hybrid grey wave model,” *Physica A: Statistical Mechanics and its Applications*, vol. 501, pp. 98–110, Jul. 2018, doi: 10.1016/j.physa.2018.02.061.

[15] Lei, M. and Feng, Z. “A proposed grey model for short-term electricity price forecasting in competitive power markets,” *International Journal of Electrical Power & Energy Systems*, vol. 43, no. 1, pp. 531–538, Dec. 2012, doi: 10.1016/j.ijepes.2012.06.001.

[16] Lei, M. and Feng, Z. *A Novel Grey Model to Short-Term Electricity Price Forecasting for NordPool Power Market*. 2009, p. 4352. doi: 10.1109/ICSMC.2009.5346948.

[17] Shrivastava, N. and Panigrahi, B. “A hybrid wavelet-ELM based short term price forecasting for electricity markets,” *International Journal of Electrical Power & Energy Systems*, vol. 55, pp. 41–50, Feb. 2014, doi: 10.1016/j.ijepes.2013.08.023.

[18] Su, M. Zhang, Z. Zhu, Y. and Zha, D. “Data-Driven Natural Gas Spot Price Forecasting with Least Squares Regression Boosting Algorithm,” *Energies*, vol. 12, no. 6, Art. no. 6, Jan. 2019, doi: 10.3390/en12061094.

[19] Saxena, A. Bansal, R. and Saini, D. *Electricity Price Forecasting by Linear Regression and SVM*. 2016. doi: 10.1109/ICRAIE.2016.7939509.

[20] Atef, S. and Eltawil, A. “A Comparative Study Using Deep Learning and Support Vector Regression for Electricity Price Forecasting in Smart Grids,” in *2019 IEEE 6th International Conference on Industrial Engineering and Applications (ICIEA)*, Apr. 2019, pp. 603–607. doi: 10.1109/IEA.2019.8715213.
[21] Kuo, P.-H. and Huang, C.-J. “An Electricity Price Forecasting Model by Hybrid Structured Deep Neural Networks,” *Sustainability*, vol. 10, p. 1280, Apr. 2018, doi: 10.3390/su10041280.
[22] Misiorek, A. Trueck, S. and Weron, R. “Point and Interval Forecasting of Spot Electricity Prices: Linear vs. Non-Linear Time Series Models,” *Studies in Nonlinear Dynamics & Econometrics*, vol. 10, pp. 1362–1362, Feb. 2007, doi: 10.2202/1558-3708.1362.
[23] Zareipour, H. Bhattacharya, K. and Canizares, C. “Forecasting the hourly Ontario energy price by multivariate adaptive regression splines,” in *2006 IEEE Power Engineering Society General Meeting*, Jun. 2006, p. 7 pp.-. doi: 10.1109/PES.2006.1709474.
[24] Tahir, A. Khan, Z. Javaid, N. Hussain, Z. A. Rasool, and S. Aimal, “Load and Price Forecasting Based on Enhanced Logistic Regression in Smart Grid,” in *Advances in Internet, Data and Web Technologies*, Cham, 2019, pp. 221–233. doi: 10.1007/978-3-030-12839-5_21.
[25] Itaba, S. and Mori, H. “A Fuzzy-Preconditioned GRBFN Model for Electricity Price Forecasting,” *Procedia Computer Science*, vol. 114, pp. 441–448, Jan. 2017, doi: 10.1016/j.procs.2017.09.010.
[26] Meier, J.H., Schneider, S., Schmidt, I., Schüller, P., Schönfeldt, T. and Wanke, B., 2018, May. "ANN-based electricity price forecasting under special consideration of time series properties". *In International Conference on Information and Communication Technologies in Education, Research, and Industrial Applications* (pp. 262-275). Springer, Cham.
[27] Gao, G. Lo, K. and Fan, F. “Comparison of ARIMA and ANN models used in electricity price forecasting for power market,” *Energy and Power Engineering*, vol. 9, no. 4B, Art. no. 4B, Apr. 2017, doi: 10.4236/epe.2017.94B015.
[28] Yadav, A. Sahay, A. Yadav, M. Bhandari, S. Yadav, A. and Sahay, K. “One hour Ahead Short-Term Electricity Price Forecasting Using ANN Algorithms,” in *2018 International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE)*, Oct. 2018, pp. 1–4. doi: 10.23919/ICUE-GESD.2018.8635596.
[29] Safarinejadian, B. Gharibzadeh, M. and Rakhshan, M. “An optimized model of electricity price forecasting in the electricity market based on fuzzy timeseries,” *Systems Science & Control Engineering*, vol. 2, no. 1, pp. 677–683, Dec. 2014, doi: 10.1080/21642583.2014.970733.
[30] Gao, W. Sarlak, V. Parsaei, M. and Ferdosi, M. “Combination of fuzzy based on a meta-heuristic algorithm to predict electricity price in an electricity markets,” *Chemical Engineering Research and Design*, vol. 131, pp. 333–345, Mar. 2018, doi: 10.1016/j.cherd.2017.09.021.
[31] Alshejari, A. Kodogiannis, V. and Leonidis, S. “Development of Neurofuzzy Architectures for Electricity Price Forecasting,” *Energies*, vol. 13, no. 5, Art. no. 5, Jan. 2020, doi: 10.3390/en13051209.
[32] Azadeh, A. Moghaddam, M. Khakzad, M. and Ebrahimipour, V. “A flexible neural network-fuzzy mathematical programming algorithm for improvement of oil price estimation and forecasting,” *Computers & Industrial Engineering*, vol. 62, no. 2, pp. 421–430, Mar. 2012, doi: 10.1016/j.cie.2011.06.019.
[33] Zadeh, L.A., Klir, G.J. and Yuan, B., 1996. "Fuzzy sets, fuzzy logic, and fuzzy systems": selected papers (Vol. 6). World Scientific.
[34] Ishibuchi, H. Nozaki, K. and Tanaka, H. “Distributed representation of fuzzy rules and its application to pattern classification,” *Fuzzy Sets and Systems*, vol. 52, no. 1, pp. 21–32, Nov. 1992, doi: 10.1016/0165-0114(92)90032-Y.
[35] Khashei, M. Rafiei, F. and Bijari, M. “Hybrid Fuzzy Auto-Regressive Integrated Moving Average (FARIMAH) Model for Forecasting the Foreign Exchange Markets,” *International Journal of Computational Intelligence Systems*, vol. 6, no. 5, pp. 954–968, Sep. 2013, doi: 10.1080/18756891.2013.809937.
[36] Sahin, Y. and Duman, E. “Detecting credit card fraud by ANN and logistic regression,” in *2011 International Symposium on Innovations in Intelligent Systems and Applications*, Jun. 2011, pp. 315–319. doi: 10.1109/INISTA.2011.5946108.
[37] Zadeh, L. Klir, G. and Yuan, B. *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers*. World Scientific, 1996.
[38] Souza, P. “Regularized Fuzzy Neural Networks for Pattern Classification Problems,” vol. 13, no. 5, p. 7, 2018.
[39] Torabi, A. Pashapour Nazari, S. and Neshat, N. “A Novel Approach of Artificial Neural Networks Modeling Based on Fuzzy Regression Approach for Forecasting Purposes: The case of liquid gas price in Japan’s

market," *Advances in Industrial Engineering*, vol. 47, no. 1, pp. 15–24, Mar. 2013, doi: 10.22059/jieng.2013.35507.
[40] Atil, A. Lahiani, A. and Nguyen, D. "Asymmetric and nonlinear pass-through of crude oil prices to gasoline and natural gas prices," *Energy Policy*, vol. 65, pp. 567–573, Feb. 2014, doi: 10.1016/j.enpol.2013.09.064.
[41] Hailemariam, A. and Smyth, R. "What drives volatility in natural gas prices?," *Energy Economics*, vol. 80, pp. 731–742, May 2019, doi: 10.1016/j.eneco.2019.02.011.
[42] Jianglin, H. "Comparing the Natural Gas Pipeline Pricing between Europe and America and the Revelation to China," *Energy Procedia*, vol. 5, pp. 659–663, Jan. 2011, doi: 10.1016/j.egypro.2011.03.116.
[43] Maxwell, D. and Zhu, Z. "Natural gas prices, LNG transport costs, and the dynamics of LNG imports," *Energy Economics*, vol. 33, no. 2, pp. 217–226, Mar. 2011, doi: 10.1016/j.eneco.2010.06.012.
[44] Nick, S. and Thoenes, S. "What drives natural gas prices? — A structural VAR approach," *Energy Economics*, vol. 45, pp. 517–527, Sep. 2014, doi: 10.1016/j.eneco.2014.08.010

Najmeh Neshat received her PhD degree in Industrial Engineering, in 2015, from Tarbiat Modaress University, Tehran, Iran. She now works as Assistant Professor at the Industrial Engineering Department of the Meybod University. She is the author and co-author of more than 40technical papers and the author of 2 books on the topics in the area of Energy Systems. Recently, she has directed her research interest towards Agent Based Modelling and Artificial Neural Networks.

**Hashem Mahlooji received his PhD degree in Industrial Engineering, in 1984, from Texas A&M, USA. He is now Full Professor at the Sharif University of Technology. He is the author and co-author of more than 60 technical papers and the author of 5 books on the topics in the area of Simulation and Probability. His research area is focused on the Simulations, Monte Carlo, and Probability and Statistics.

**Murat Kaya is currently working as an assistant professor at the Industrial Engineering Program of Sabancı University, Istanbul, Turkey. He holds M.S. and Ph.D. degrees from Stanford University Management Science and Engineering (MS&E) Department. During his Ph.D. study, Dr. Kaya has worked for several projects at Hewlett Packard Research Laboratories in Palo Alto, USA. Dr. Kaya's recent interest is directed towards applied decision & optimization problems in energy supply chains.

Figure 1. Network structure of Exhaustive prune model

Table 1. Statistical information on influencing variables of LNG price in the Japanese market

Table 2. Comparison of the relative importance of input variables for each neural network model

Table 3. Numerical results of performance indicators R^2, RMSE, and MAPE for ANN models

Table 4. The weights and biases of the Exhaustive Prune Model

Table 5. Center and width of hidden weights w_q

Table 6. Numerical results of performance indicators R^2, RMSE, and MAPE for the proposed model and the alternative models

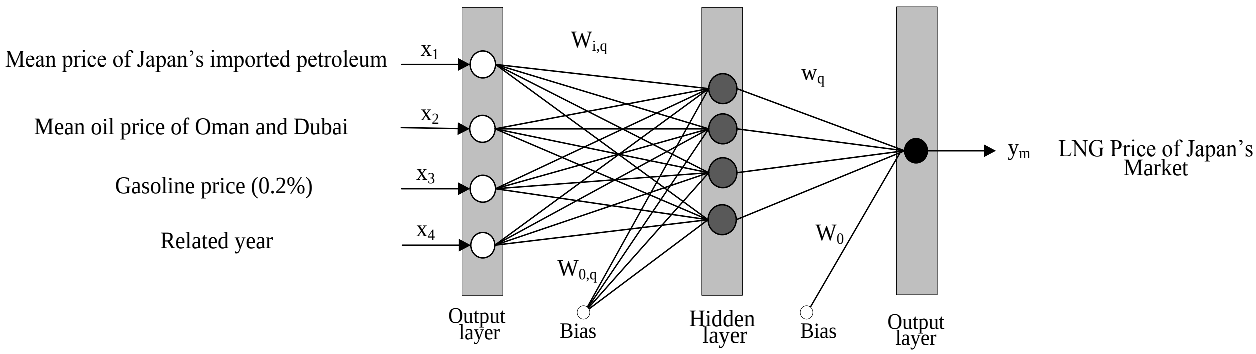

Figure 1. **Network structure of Exhaustive prune model**

Table 1. Statistical information on influencing variables of LNG price in the Japanese market

Influencing variables	Unit	Mean value	Std.dev	Interval		Mean growth rate
				Max	Min	
British ICE natural gas price	MMBtu/$	4.95	2.80	14.33	1.66	0.10
Henry Hub natural gas price	MMBtu/$	4.33	2.81	13.71	1.41	0.13
WTI oil price	MMBtu/$	6.65	4.46	23.16	1.90	0.11
Brent oil price	MMBtu/$	6.45	4.46	22.96	1.70	0.12
Mean price of Japan's imported oil	MMBtu/$	6.65	4.65	22.78	1.94	0.16
Mean oil price of Oman and Dubai	MMBtu/$	6.44	4.65	22.57	1.73	0.16
Price of liquified petroleum gas (LPG)	MMBtu/$	6.96	3.90	19.18	2.17	0.09
Oil fired in oven (1%)	MMBtu/$	4.63	3.02	17.37	1.48	0.11
Oil fired in oven (3.5%)	MMBtu/$	4.09	2.80	15.79	1.32	0.10
Gasoline price (0.2%)	MMBtu/$	7.90	5.58	28.18	2.16	0.13
Oil fired oven (180)	MMBtu/$	3.91	2.71	14.89	1.09	0.12
International gas consumption	TCM	0.210	0.020	0.25	0.18	0.02
International gas production	BCM	0.213	0.023	0.25	0.18	0.025
International oil consumption	MMbl/d	6.543	0.454	7.32	5.79	0.017
International oil production	MMbl/d	6.359	0.402	7.01	5.68	0.015

Table 2 Comparison of the relative importance of input variables for each neural network model

Input Variables	Dynamic	Multiple	Prune	Exhaustive prune	RFBN
British ICE natural gas price	0.098	0.185	0.056	-	0.109
Henry Hub natural gas price	0.158	0.505	0.434	-	0.117
WTI oil price	0.042	0.042	0.068	-	0.110
Brent oil price	0.085	0.056	0.102	0.164	0.112
Mean price of Japan's imported petroleum	0.177	0.174	0.183	0.285	0.115
Mean oil price of Oman and Dubai	0.021	0.089	-	0.152	0.116
Price of liquid petroleum gas (LPG)	0.032	0.054	0.079	-	0.110
Oil fired oven (1%)	0.114	0.055	0.118	-	0.112
Oil fired oven (3.5%)	0.234	0.182	0.138	-	0.129
Gasoline price (0.2%)	0.009	0.174	-	0.549	0.116
Oil fired oven (180)	0.024	0.115	-	-	0.140
International gas consumption	0.021	0.179	0.159	-	0.140
International gas production	0.071	0.103	-	-	0.138
International oil consumption	0.078	0.149	-	-	0.137
International oil production	0.0175	0.194	0.16	-	0.115

Table 3. Numerical results of performance indicators R^2, RMSE, and MAPE for ANN models

Indicator / Network	R^2	RMSE	MAPE
Dynamic	0.8441	0.017	1.812
Multiple	0.8743	0.012	1.211
Prune	0.8722	0.012	1.287
Exhausted Prune	0.8675	0.016	1.521
REBN	0.8125	0.019	1.803

Table 4. The weights and biases of the Exhaustive Prune Model

Input weights				Hidden weights	Biases	
$w_{i,1}$	$w_{i,2}$	$w_{i,3}$	$w_{i,4}$	w_q	$w_{0,j}$	w_0
-2.5677	10.3961	-2.0112	16.0023	-2.6790	-13.2756	-1.6221
1.3113	17.2834	8.5331	-0.2344	0.0402	3.0111	
4.6701	-1.8766	-11.9012	14.7558	14.6723	-8.3783	
0.2349	20.9123	-2.7663	5.03777	5.2310	-5.2311	

Table 5. Center and width of hidden weights w_q

	e_q	c_q
w_0	-1.6221	0.478
w_1	-2.6790	0.882
w_2	0.0402	0.016
w_3	14.6723	2.583
w_4	5.2310	1.771

Table 6. Numerical results of performance indicators R2, RMSE, and MAPE for the proposed model and the alternative models

Model / Indicator	Proposed	Exhausted Prune ANN	Fuzzy regression	Linear regression
R2	0.8701	0.8675	0.7901	0.7733
MAPE	1.491	1.521	2.2988	2.4611
RMSE	0.014	0.016	0.019	0.021