

Article

Short-Term Electricity Price Forecasting Based on Similar Day-Based Neural Network

Chun-Yao Lee * and Chang-En Wu

Department of Electrical Engineering, Chung Yuan Christian University, No. 200, Zhongbei Road, Zhongli District, Taoyuan City 320, Taiwan; g9778019@cycu.edu.tw

* Correspondence: cyl@cycu.edu.tw; Tel.: +886-3-2654827

Received: 5 August 2020; Accepted: 24 August 2020; Published: 26 August 2020



Abstract: This paper presents four refined distance models to the application of forecasting short-term electricity price namely Euclidean norm, Manhattan distance, cosine coefficient, and Pearson correlation coefficient. The four refined models were constructed and used to select the days, which are like a reference day in electricity prices and loads, called similar days in this study. Using the similar days, the electricity prices of a forecast day were further obtained by similar day regression (SDR) and similar day based artificial neural network (SDANN). The simulation results of the case of the PJM (Pennsylvania, New Jersey and Maryland) interchange energy market indicate the superiority and availability of the selection 45 framework days and three similar days based on Pearson correlation coefficient model.

Keywords: similar-day method; linear regression; artificial neural network; electricity price

1. Introduction

Under liberalization of electricity industry, the bidding of electricity price is an important operation model of the market operation. The forecast of electricity price is very helpful for the bidding strategies of the market participants [1,2]. There are many factors impact electricity price, in which some factors are more important than others are, and practically, we can only consider those more important factors. In the power market, the load pattern is an effective parameter on the bidding behavior of utility. Therefore, we can consider the historical price and system load as the factors which impact on price [3]. Scholars usually used classical regression analysis theory [4–6] to forecast the price for managing location marginal price (LMP). Nowadays, artificial intelligence methods, such as neural network [7–10] and particle swarm optimization [11–14] are widely used in many fields. Particularly, some modified neural network, such as fuzzy neural [15,16], and some statistical methods, such as time series method [17,18], are used for forecasting the price. However, the historical data and proper training are non-linear, and mass data of load and electricity always decrease the level of forecasting accuracy. Thus, using similar days to forecast the price becomes a new issue. Furthermore, the electricity price forecasting was concerned by the combination of the electricity price and load data to increase the accuracy. And the manner can efficiently improve the computational time under the approximately accuracy. In this paper, we refine four distance models to calculate the similarity between the similar days and reference day, which is the day before the forecasting day. Finally, the four models of choosing similar days, Euclidean norm, Manhattan distance, Cosine coefficient, and Pearson correlation coefficient were discussed to the influence of the forecasting accuracy of the electricity price.

2. Similar Days Selection

2.1. Relation Evaluation

Choosing appropriate data of price and load is important to LMP forecast. We apply the method choosing similar days to increase the performance of the forecasting [19,20]. In the study, the historical data of PJM (Pennsylvania, New Jersey and, Maryland) interchange energy market, including LMP and load were used due to the high degree of reliability and representative. We defined the correlation coefficient (R) between the price and load in the data, shown as (1). The coefficient of determination R^2 is employed to evaluate the correlation. For example, regarding the LMP and load from January to May in 2006, as shown in Figure 1 [10], R^2 is 0.615 obtained by the regression simulation, which means the R is 0.784 and the LMP and load is high correlation, according to Table 1. Thus, using the load to forecast the LMP is reasonable [21–25].

$$R = \frac{\sum_{i=1}^n (P_i - \bar{P})(L_i - \bar{L})}{\sqrt{\sum_{i=1}^n (P_i - \bar{P})^2} \sqrt{\sum_{i=1}^n (L_i - \bar{L})^2}} \quad (1)$$

where

P_i is LMP of reference day at number i (\$/MWh),

\bar{P} is average LMP of reference day (\$/MWh),

L_i is load of reference day at number i (MW),

\bar{L} is average load reference day (MW).

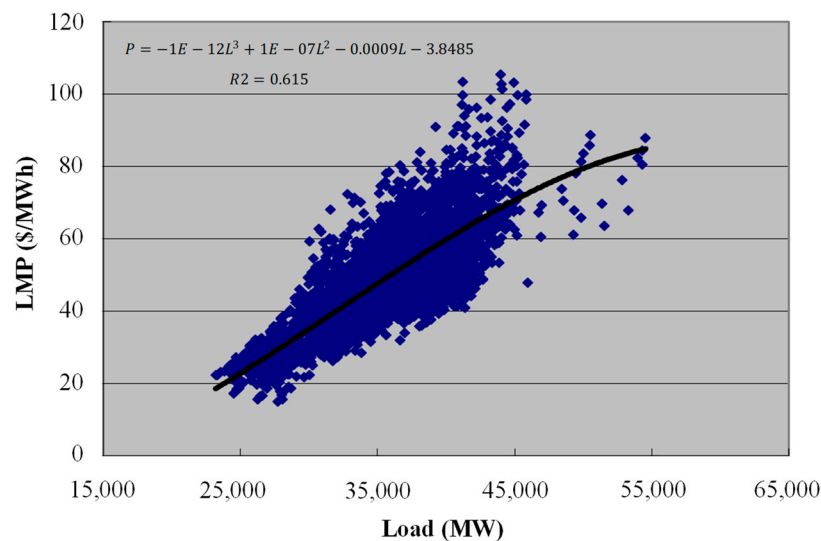


Figure 1. Relationship between LMP and load in the PJM (January–May 2006).

Table 1. Related explanation.

R	Explanation
$R = \pm 1$	Perfect correlation
$0.7 \leq R \leq 1$	High degree of correlation
$0.3 \leq R < 0.7$	Moderate degree of correlation
$0 \leq R < 0.3$	Low degree of correlation
$R = 0$	No correlation

2.2. Similarity Calculation Model

The purpose of this paper is to forecast LMP according to historical loads and prices, called a day before the forecast day as a reference day. The days which are like the reference day in electricity

prices and loads are selected and are namely similar days. The structure of similar days is based on collecting historical data that affect electricity prices and calculate the distance between all historical data and the forecasted day. The correlation between all the data and the forecasted day is evaluated. We adopt four commonly used models in calculating distance, which are Euclidean norm, Manhattan distance, Cosine coefficient and Pearson correlation coefficient. The models with weighted factors are refined, so that obtained model A is the Refined Euclidean norm, model B is the Refined Manhattan distance, mode C is Refined Cosine coefficient, and model D is Refined Pearson correlation coefficient. The models used to evaluate the similarity between the reference days and searched similar days in this paper. Many data science techniques are based on measuring similarity and dissimilarity between objects. Measuring similarity between objects can be performed in a number of ways. Generally, we can divide similarity metrics into two different groups which are similarity-based metrics and distance based metrics. Cosine coefficient and Pearson correlation coefficient are the part of similarity-based metrics and Euclidean distance and Manhattan distance are the part of distance based metrics. We choose those methods because the methods are well-known, relatively easy to use and simple from another method. In our paper, the models of Euclidean distance and Manhattan distance are used to evaluate the data dissimilarity, which means the relativity is higher when the distance value is smaller. The models of Cosine coefficient and Pearson correlation coefficient are used to evaluate the data similarity, which means the relativity is higher when the distance value is bigger. The four refined models are rebuilt as models A, B, C and D, which are described as follows:

2.2.1. Model A: Refined Euclidean Norm

The relativity of Euclidean norm is refined based on [10]. The relativity model involves LMPs and loads of the reference and similar days, shown as (2).

$$\|D\| = \sqrt{[\hat{w}_1 \Delta L_t]^2 + [\hat{w}_2 \Delta L_{t-1}]^2 + [\hat{w}_3 \Delta P_t]^2 + [\hat{w}_4 \Delta P_{t-1}]^2} \quad (2)$$

where

$$\Delta P_{t-1} = P_{t-1} - P_{t-1}^p, \quad \Delta L_t = L_t - L_t^p, \quad \Delta L_{t-1} = L_{t-1} - L_{t-1}^p,$$

$\|D\|$ is distance,

P_t is LMP of reference day at time t (\$/MWh),

P_t^p is LMP of similar day in past at time t (\$/MWh),

L_t is load of reference day at time t (MW),

L_t^p is load of similar day in past at time t (MW),

P_{t-1} is LMP of reference day at time $(t-1)$ (\$/MWh),

P_{t-1}^p is LMP of similar day in past at time $(t-1)$ (\$/MWh),

L_{t-1} is load of reference day at time $(t-1)$ (MW),

L_{t-1}^p is load of similar days in past at time $(t-1)$ (MW),

ΔL_t is load deviation between reference day and similar days at t time (MW),

ΔL_{t-1} is load deviation between reference day and similar days at time $(t-1)$ (MW),

ΔP_t is LMP deviation between reference day and similar days at time t (\$/MWh),

ΔP_{t-1} is LMP deviation between reference day and similar days at time $(t-1)$ (\$/MWh),

\hat{w}_i is weight factor ($i = 1 \sim 4$).

2.2.2. Model B: Refined Manhattan Distance

The relativity of Manhattan distance is refined based on [10]. The relativity model involves LMPs and loads of reference and similar days, shown as (3).

$$\|D\| = |\hat{w}_1 \Delta L_t| + |\hat{w}_2 \Delta L_{t-1}| + |\hat{w}_3 \Delta P_t| + |\hat{w}_4 \Delta P_{t-1}| \quad (3)$$

2.2.3. Model C: Refined Cosine Coefficient

The relativity of Cosine coefficient is refined based on [26]. The relativity model involves LMPs and loads of reference and similar days, shown as (4).

$$\|D\| = \frac{E_1 + E_2 + E_3 + E_4}{\sqrt{(D_1 + D_2 + D_3 + D_4)(D_5 + D_6 + D_7 + D_8)}} \quad (4)$$

where

$$\begin{aligned} D_1 &= (\hat{w}_1 L_t)^2, D_2 = (\hat{w}_2 L_{t-1})^2, D_3 = (\hat{w}_3 P_t)^2, D_4 = (\hat{w}_4 P_{t-1})^2, D_5 = (\hat{w}_1 L_t^p)^2, \\ D_6 &= (\hat{w}_2 L_{t-1}^p)^2, D_7 = (\hat{w}_3 P_t^p)^2, D_8 = (\hat{w}_4 P_{t-1}^p)^2, E_1 = \hat{w}_1^2 (L_t \cdot L_t^p), E_2 = \hat{w}_2^2 (L_{t-1} \cdot L_{t-1}^p) \\ E_3 &= \hat{w}_3^2 (P_t \cdot P_t^p), E_4 = \hat{w}_4^2 (P_{t-1} \cdot P_{t-1}^p). \end{aligned}$$

2.2.4. Model D: Refined Pearson Correlation Coefficient

The relativity of Pearson correlation coefficient is refined based on [26]. The model involves LMPs and loads of reference and similar days, as shown in (5).

$$\|D\| = \frac{L_1 \cdot L_2 + L_3 \cdot L_4 + P_1 \cdot P_2 + P_3 \cdot P_4}{\sqrt{L_1^2 + L_3^2 + P_1^2 + P_3^2} \sqrt{L_2^2 + L_4^2 + P_2^2 + P_4^2}} \quad (5)$$

where

$$\begin{aligned} L_1 &= \hat{w}_1 (L_t - \bar{L}_t), L_2 = \hat{w}_1 (L_t^p - \bar{L}_t^p), L_3 = \hat{w}_2 (L_{t-1} - \bar{L}_{t-1}), L_4 = \hat{w}_2 (L_{t-1}^p - \bar{L}_{t-1}^p), \\ P_1 &= \hat{w}_3 (P_t - \bar{P}_t), P_2 = \hat{w}_3 (P_t^p - \bar{P}_t^p), P_3 = \hat{w}_4 (P_{t-1} - \bar{P}_{t-1}), P_4 = \hat{w}_4 (P_{t-1}^p - \bar{P}_{t-1}^p), \end{aligned}$$

\bar{P}_t is average LMP of reference day at time t (\$/MWh),

\bar{P}_{t-1} is average LMP of reference day at time $(t-1)$ (\$/MWh),

\bar{P}_t^p is average LMP of similar day in past at time t (\$/MWh),

\bar{P}_{t-1}^p is average LMP of similar day in past at time $(t-1)$ (\$/MWh),

\bar{L}_t is average load of reference day at time t (MW),

\bar{L}_{t-1} is average load of reference day at time $(t-1)$ (MW),

\bar{L}_t^p is average load of similar day in past at time t (MW),

\bar{L}_{t-1}^p is average load of similar day in past at time $(t-1)$ (MW).

The average LMP and load mean the value of the LMP and load as the average values from reference day at time t and $t-1$ and also on a similar day in the past at time t and $t-1$.

2.3. Interval of Time Framework

The interval of time framework, d , can be chosen when the reference day is determined, as shown in Figure 2 [26]. For example, assuming that the reference date is point A, the similar days are the d days from the day before a reference day, and past d days before and after the reference day in the previous year. In this study, the d is assumed as 15, 30, 45 and 60.

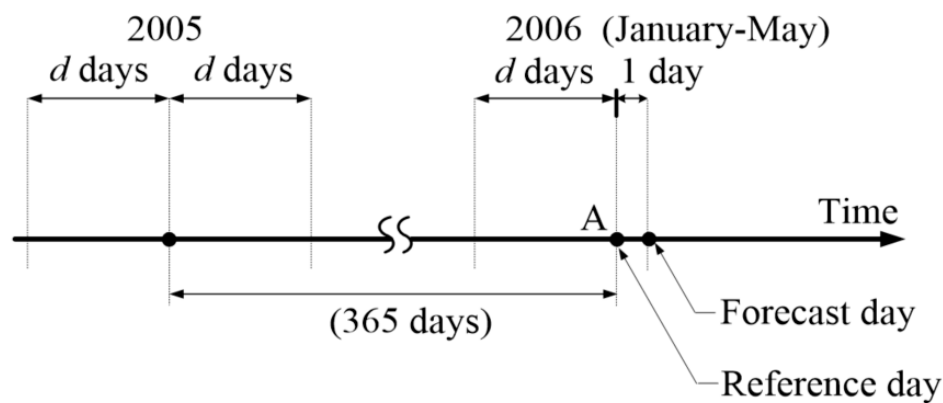


Figure 2. Time framework for the selection of similar days.

2.4. Procedure of Similar Date Selection

There are five processes to choose similar days which are as follows: reference day definition, interval selection, similarity calculation, similarity sorting, and similar day determination, as shown in Figure 3.

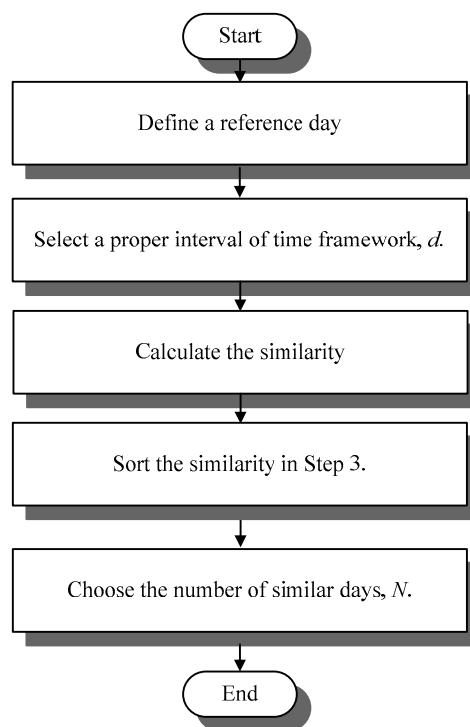


Figure 3. Similar day selection flowchart.

- Step 1.** Define a reference day which is one day before the forecasting day;
- Step 2.** Select a proper interval of time framework, d , based on Figure 2;
- Step 3.** Equations (2)–(5) are used to calculate the similarity between the selecting days and the reference day in LMP and load;
- Step 4.** Sort the similarity in Step 3. Note that the (2) and (3) are dissimilarity. Nevertheless (4) and (5) are similarity;
- Step 5.** Choose the number of similar days, N .

3. Similar Day Regression Model

3.1. Brief Review of Regression Model

The paper uses a typical regression model to forecast the LMPs, as shown in (6), in which the weight factors can be obtained by the least square method [26]. The obtained weight factors in (6) are applied in (2)–(5).

$$P_{t+1} = \hat{w}_0 + \hat{w}_1 L_t + \hat{w}_2 L_{t-1} + \hat{w}_3 P_t + \hat{w}_4 P_{t-1} \quad (6)$$

where

- P_t is LMP of similar day at t time (\$/MWh),
- P_{t-1} is LMP of similar day at $(t - 1)$ time (\$/MWh),
- P_{t+1} is LMP of forecasting day at $(t + 1)$ time (\$/MWh),
- L_t is load of similar day at t time (MW),
- L_{t-1} is load of similar day at $(t - 1)$ time (MW),
- \hat{w}_i is weight factors ($i = 0 \sim 4$).

3.2. Regression Model of Forecasting LMPs

The four steps to forecast LMP by using regression model are variables definition, weight factor analysis, weight factor evaluation, and LMPs forecasting, as shown in Figure 4.

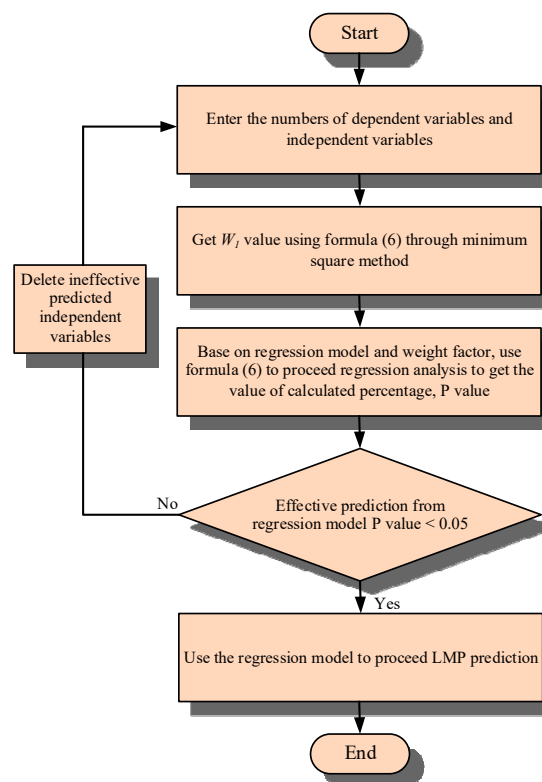


Figure 4. Regression model forecasting procedure.

- Step 1.** Variables definition: In the regression model, the dependent variables are assumed as forecasting P_{t+1} , and the independent variables are as P_t , P_{t-1} , L_t , and L_{t-1} ;
- Step 2.** Weight factor analysis: Base on the LMPs and loads of similar days and the reference day, the weight factors in (6) can be analyzed;
- Step 3.** Weight factor evaluation: The forecast is acceptable when the probability values (p value) of weight factors are smaller than 0.05. Thus, the independent variables with the P bigger than 0.05 should be removed and the remained weight factors are analyzed though step (2);

Step 4. LMPs forecasting: \bar{P}_t and \bar{L}_t in (7) are the average LMP and load data respectively, which are obtained by taking an average of selected N similar days. And the forecasting price P_{t+1} can be calculated by the obtained \bar{P}_t and \bar{L}_t .

$$P_{t+1} = \hat{w}_0 + \hat{w}_1 \bar{L}_t + \hat{w}_2 \bar{L}_{t-1} + \hat{w}_3 \bar{P}_t + \hat{w}_4 \bar{P}_{t-1} \quad (7)$$

4. Neural Network Based Forecasting Models of Similar Day

4.1. Neural Network Model

Figure 5 illustrates the typical structure of the artificial neural network (ANN) [10] which is consisted of three layers namely input layer, hidden layer, and output layer. The ANN based on price forecasting is called SDANN. The input variables of the proposed SDANN structure for price forecasting are mean similar day data \bar{P}_t , which is an average of chosen N similar days, actual hourly load L_t , and actual hourly LMP P_t . The output variable of the structure is the LMP of the forecast day, \hat{P}_{t+24} . N is 3, 5, and 10 in this paper.

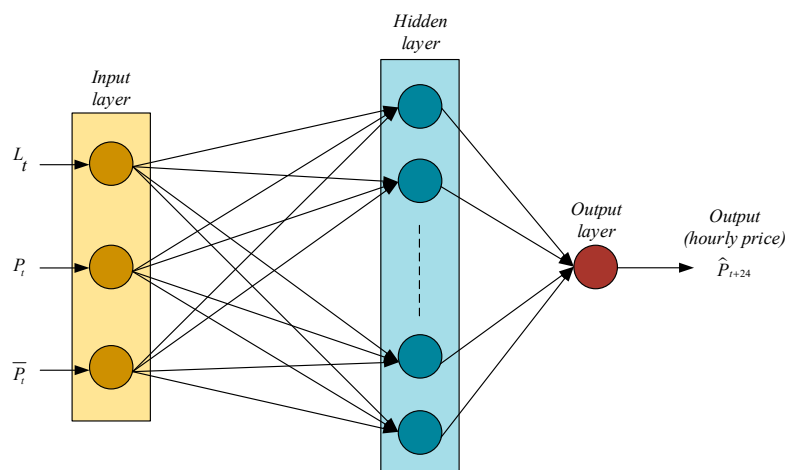


Figure 5. Neural network architecture diagram.

The applied ANN uses Back-Propagation (BP) algorithm and $2n + 1$ hidden neuron is chosen for forecasting where n is the number of input nodes. The weights are adjusted by using the calculated error between forecasting and actual values. The transfer function is sigmoid and relative parameters assumed as the follows: the number of input neurons is 3, the hidden layer is 1, the output neurons is 1, the hidden layer neuron is 7, the learning rate (α) is 0.8, the momentum factor is 0.1 and the iteration is 5000.

4.2. Neural Network Forecasting Model

When using neural network to proceed predicting, firstly, applying neural network proceed LMP forecasting. The historical data which is LMP and load must divide into two section: training data and testing data. Secondly, the training data will be putted into neural network and proceed to repeat training and learning. Finally, the test data will be applied to forecast LMP after finishing training network, as shown in Figure 6 [27].

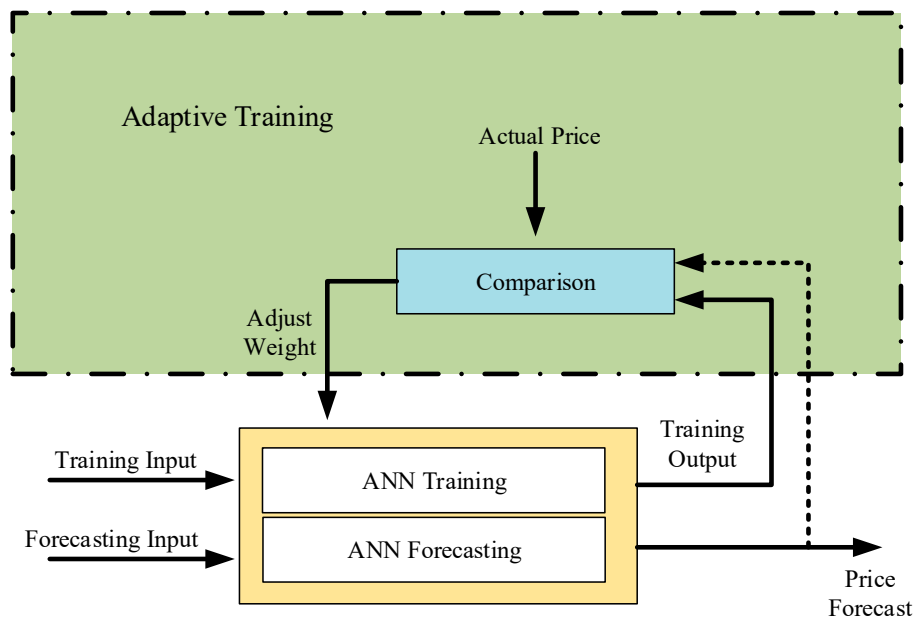


Figure 6. Price forecast based on neural network.

- (a) Training: Selecting similar day interval, d , it is 45 days in the paper, and therefore the load and LMP data have 134 days (excluding the load and LMP of reference day.) as input training data. According to Figure 2, input training data is based on the reference date which is point A. The similar days are the d days from the day before a reference day and past d days before and after the reference day in the previous year. And during training, need to exclude the load and LMP of reference day.
- (b) Testing: Forecasting input data is the load and LMP of the reference day.

4.3. Forecast Accuracy

Forecasting prices which are applied to neural network and regression model must examine forecast accuracy by error value. In forecast accuracy, then mean absolute percentage error (MAPE) and forecast mean square error (FMSE) are the standard to verify the accuracy [26].

- (a) Mean absolute percentage error

Also known as the absolute average relative error, it can show forecast accuracy accurately, shown as (8):

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|P_i^{\text{true}} - P_i^{\text{est}}|}{\bar{P}_i^{\text{true}}} \quad (8)$$

where

P_i^{true} is i -th actual LMP (\$/MWh),

P_i^{est} is i -th forecasting LMP (\$/MWh),

\bar{P}_i^{true} is actual average LMP on prediction day (\$/MWh), which is shown in (9).

$$\bar{P}_i^{\text{true}} = \frac{1}{n} \sum_{i=1}^n P_i^{\text{true}} \quad (9)$$

- (b) Forecasting mean square error

The value of FMSE is $[0, \infty]$. The forecasting accuracy will be lower when the FMSE value is bigger, related to (8) and (9).

$$\text{FMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (p_i^{\text{true}} - p_i^{\text{est}})^2} \quad (10)$$

The forecasting accuracy will be lower when the FMSE value is bigger, related to (8) and (9).

5. Comparison of Simulation Results

5.1. Regression Model Simulation Result and Comparison

To verify the effectiveness of the proposed method in this research, five days as the forecast days are chosen from the historical data of PJM (Pennsylvania, New Jersey and, Maryland) interchange energy market, which are 20 January, 10 February, 5 March, 7 April and 13 May of 2006. The data used is adequate to be applied in the proposed model. The time chosen as intervals d are 15 days, 30 days, 45 days, and 60 days. We find out similar days N (three days, five days, and 10 days.) through models A to D to be the input of regression model and proceed LMP forecasting. MAPE simulation result of regression model LMP forecasting, which is illustrated in Table 2. To average the simulation result of the 5 forecasting days, obviously, the best LMP forecasting accuracy in model D, $d = 45$, and $N = 3$, the MAPE is 9.41%, which is illustrated in Table 3. Therefore, the forecasting performance by using model D to choose similar day is better than the other three models. Moreover, the performance of $d = 45$ in this research is the best when choosing intervals because there would be problems in cross seasonal electricity using habit differences if the chosen interval is too long. However, if the chosen interval is too short makes the data amount too small. Moreover, the performance is better if similar day, $N = 3$, because the data chosen is the most like similar day.

Table 2. Regression model price forecast of MAPE.

20 January 2006												
MAPE	$d = 15$			$d = 30$			$d = 45$			$d = 60$		
ANN	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$
Model A (%)	8.64	8.32	8.23	9.62	8.03	8.46	10.13	14.96	13.51	7.69	8.41	8.33
Model B (%)	8.80	8.18	7.85	12.13	9.97	8.46	13.63	9.02	14.53	9.69	9.86	9.76
Model C (%)	13.11	12.70	10.21	10.91	11.19	9.62	10.16	19.27	13.16	11.31	14.87	13.71
Model D (%)	12.76	14.77	14.13	10.73	11.16	12.57	8.10	8.39	10.57	9.69	8.48	10.76
10 February 2006												
MAPE	$d = 15$			$d = 30$			$d = 45$			$d = 60$		
ANN	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$
Model A (%)	11.06	12.67	10.69	11.51	11.28	11.07	10.48	11.74	10.22	12.12	11.99	11.35
Model B (%)	11.05	11.84	12.27	11.13	11.28	11.01	10.04	11.28	11.19	12.06	11.66	11.55
Model C (%)	11.50	11.66	11.32	10.46	11.51	11.26	9.67	10.20	11.81	10.96	10.89	11.07
Model D (%)	9.19	11.04	13.10	9.37	10.22	10.61	8.93	9.54	11.31	11.46	10.09	11.79

Table 2. Cont.

5 March 2006												
MAPE	$d = 15$			$d = 30$			$d = 45$			$d = 60$		
	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$
ANN												
Model A (%)	7.84	9.65	9.54	9.07	9.44	8.71	9.67	7.68	8.57	6.57	6.66	7.95
Model B (%)	9.99	11.32	9.75	9.34	9.44	8.70	10.67	7.64	8.11	8.03	6.89	7.95
Model C (%)	11.41	12.21	14.81	10.89	12.94	14.24	11.10	9.81	9.21	11.65	9.90	6.62
Model D (%)	8.14	7.24	7.76	8.33	7.31	8.22	6.21	6.41	7.74	6.81	6.63	8.15
7 April 2006												
MAPE	$d = 15$			$d = 30$			$d = 45$			$d = 60$		
	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$
ANN												
Model A (%)	10.17	10.07	9.25	11.47	11.36	11.56	12.68	15.21	12.68	11.95	11.73	12.31
Model B (%)	9.44	10.07	10.62	10.70	11.36	11.29	11.98	11.71	10.90	12.28	10.05	11.37
Model C (%)	9.05	8.36	8.45	9.49	9.73	8.34	10.08	12.67	12.74	11.73	12.93	12.04
Model D (%)	11.78	9.82	10.43	12.56	12.69	11.52	12.54	13.09	13.12	12.03	13.42	13.09
13 May 2006												
MAPE	$d = 15$			$d = 30$			$d = 45$			$d = 60$		
	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$
ANN												
Model A (%)	16.78	17.34	19.70	16.99	18.34	19.07	11.57	14.82	13.51	16.76	18.16	19.54
Model B (%)	14.40	12.44	16.43	16.48	16.98	18.65	11.63	14.81	16.31	15.23	16.23	17.97
Model C (%)	11.78	7.85	9.14	11.59	9.99	9.50	11.63	13.41	14.71	13.62	11.93	17.03
Model D (%)	8.34	9.88	15.37	10.19	11.59	15.24	11.29	13.67	14.08	11.38	13.77	14.17

Table 3. Five-day average MAPE analysis (Regression model).

MAPE	$d = 15$			$d = 30$			$d = 45$			$d = 60$		
RM	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$
Model A (%)	10.90	11.61	11.48	11.73	11.69	11.77	10.91	12.88	11.70	11.02	11.39	11.90
Model B (%)	10.74	10.77	11.38	11.96	11.81	11.62	11.59	10.89	12.21	11.46	10.94	11.72
Model C (%)	11.37	10.56	10.79	10.67	11.07	10.59	10.53	13.07	12.33	11.85	12.10	12.09
Model D (%)	10.04	10.55	12.16	10.24	10.59	11.63	9.41	10.22	11.36	10.27	10.48	11.59

5.2. Neural Network Simulation Result and Comparison

In this study, five days as the forecast days are chosen, which are 20 January, 10 February, 5 March, 7 April and 13 May of 2006. The times chosen as interval d are 15 days, 30 days, 45 days, and 60 days. The similar days N (three days, five days, and 10 days) through models A until D to be the input of the neural network and proceed LMP forecasting. The simulation result is illustrated in Table 4. To average the simulation results of the five forecasting days, obviously, the best forecasting accuracy, which illustrated in Table 5 is from model D ($d = 45, N = 3$).

Table 4. Neural network for forecasting electricity price of MAPE.

20 January 2006												
MAPE	$d = 15$			$d = 30$			$d = 45$			$d = 60$		
ANN	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$
Model A (%)	7.29	7.55	6.35	7.23	5.65	5.83	6.47	6.83	6.67	5.83	6.55	5.83
Model B (%)	7.25	7.84	6.96	7.11	6.73	5.95	5.49	7.36	5.28	7.14	6.98	6.75
Model C (%)	9.77	9.12	8.39	7.20	7.01	7.15	6.10	5.90	6.21	4.33	7.97	7.03
Model D (%)	6.30	6.57	6.49	6.04	4.46	6.17	5.66	6.09	7.31	4.39	5.01	6.44
10 February 2006												
MAPE	$d = 15$			$d = 30$			$d = 45$			$d = 60$		
ANN	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$
Model A (%)	7.06	7.43	8.16	7.47	8.24	8.46	7.91	9.13	8.94	8.70	9.56	9.31
Model B (%)	6.93	7.38	7.68	7.51	8.68	8.52	7.91	8.53	7.79	8.56	9.77	9.91
Model C (%)	6.98	7.36	7.72	7.42	8.77	8.75	7.38	7.69	8.17	8.25	9.03	8.58
Model D (%)	5.96	7.25	8.21	6.72	7.41	9.89	6.13	7.14	9.42	7.47	8.88	7.62
5 March 2006												
MAPE	$d = 15$			$d = 30$			$d = 45$			$d = 60$		
ANN	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$
Model A (%)	6.97	7.10	7.48	7.42	7.54	8.23	8.38	7.61	9.91	7.82	8.23	8.32
Model B (%)	6.97	7.19	7.58	7.14	7.66	8.88	5.82	5.27	5.26	8.24	8.34	8.16
Model C (%)	8.21	8.87	9.45	8.40	8.89	9.93	4.99	4.79	5.52	8.99	9.19	8.23
Model D (%)	6.56	6.62	7.41	5.45	7.08	8.24	5.03	5.55	7.48	6.73	5.90	7.69
7 April 2006												
MAPE	$d = 15$			$d = 30$			$d = 45$			$d = 60$		
ANN	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$
Model A (%)	7.17	7.06	6.63	5.78	5.89	5.22	9.19	9.31	6.07	6.39	6.64	5.42
Model B (%)	6.81	7.22	7.10	6.25	5.87	4.93	9.19	6.14	5.74	6.53	5.91	5.02
Model C (%)	7.07	6.63	6.47	5.85	5.69	6.05	6.67	7.22	6.29	7.09	7.64	5.53
Model D (%)	7.13	7.05	6.95	5.27	6.19	5.27	6.15	6.35	6.40	7.35	6.84	5.28
13 May 2006												
MAPE	$d = 15$			$d = 30$			$d = 45$			$d = 60$		
ANN	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$
Model A (%)	11.50	13.65	14.11	10.91	12.87	11.63	5.08	7.67	5.43	10.35	9.78	9.73
Model B (%)	16.09	13.47	15.38	8.65	10.25	8.54	6.60	10.34	6.23	8.96	8.73	10.24
Model C (%)	9.05	8.05	8.87	6.46	6.22	5.27	5.85	7.44	6.88	9.27	6.71	7.28
Model D (%)	6.71	8.33	9.19	4.99	5.34	8.42	4.67	6.24	7.28	7.92	6.95	6.19

Table 5. Five-day average MAPE analysis (Artificial Neural networks).

MAPE	$d = 15$			$d = 30$			$d = 45$			$d = 60$		
	ANN	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$	$N = 10$	$N = 3$	$N = 5$
Model A (%)	8.00	8.56	8.55	7.76	8.04	7.87	7.41	8.11	7.40	7.82	8.15	7.72
Model B (%)	8.81	8.62	8.94	7.33	7.84	7.36	7.00	7.53	6.06	7.89	7.95	8.02
Model C (%)	8.22	8.01	8.18	7.07	7.32	7.43	6.20	6.61	6.61	7.59	8.11	7.33
Model D (%)	6.53	7.16	7.65	5.69	6.10	7.60	5.53	6.28	7.58	6.77	6.72	6.64

5.3. Comparison of Regression Forecasting and Neural Network Forecasting

Using model D (with the condition of $d = 45$ and $N = 3$), the best forecasting result parameters to proceed the prediction performance comparison by two methods: regression model and neural network (Forecast days are 20 January, 10 February, 5 March, 7 April and 13 May of 2006). The simulation results are illustrated as Figure 7a–e.

Obviously, from Figure 7, it can be seen that using the neural network to predict location marginal price is closer to actual LMP. The performance using the neural network is better than the regression model forecasting method. The reason is, before predicting, the neural network goes through repeat learning training and the process of error adjustment weight value, whereas regression model lack of the learning character. The MAPE and FMSE of regression model and neural network forecasting price are illustrated in Table 6.

Table 6. Regression model and neural network forecast price to MAPE and FMSE.

Method	RM		ANN		Deviation	
	MAPE (%)	FMSE (\$/MWh)	MAPE (%)	FMSE (\$/MWh)	MAPE (%)	FMSE (\$/MWh)
January 20	8.10	5.397	5.66	4.232	2.44	1.165
February 10	8.93	7.153	6.13	4.382	2.80	2.771
March 5	6.21	3.429	5.03	2.820	1.18	0.609
April 7	12.53	8.117	6.15	3.787	6.38	4.33
May 13	11.29	5.203	4.67	2.138	6.62	3.065
Average	9.41	5.860	5.53	3.472	3.88	2.388

In Table 6, the forecasting accuracy of neural network is better than that of regression model by calculating MAPE and FMSE. Some statistics regarding single forecast day result: For the accuracy of using neural network is minimum 1.18% improved, 6.62% maximum improved, and average forecasting improved is 3.88% in MAPE.

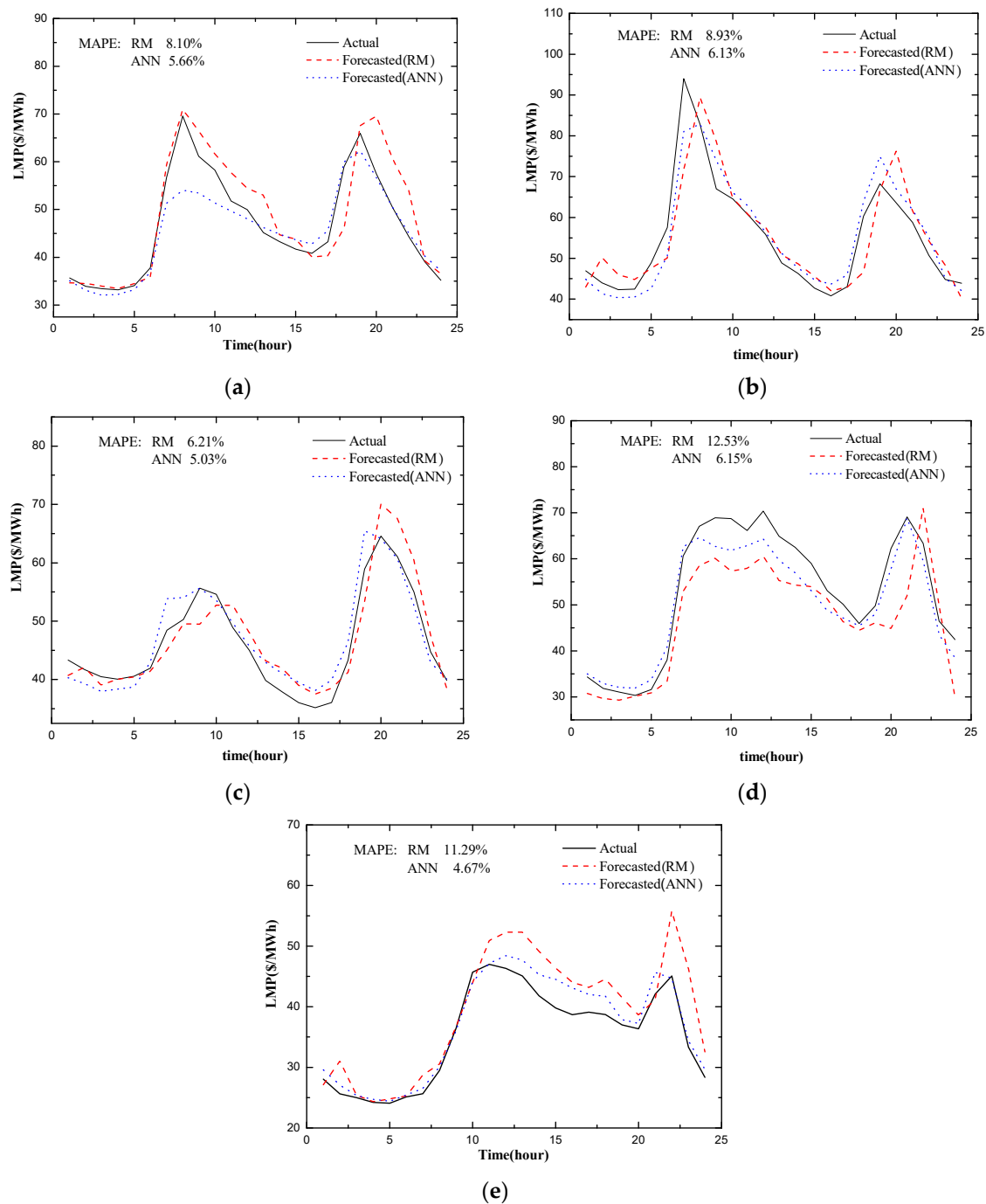


Figure 7. Results of regression model and neural network forecast price (2006). (a) 20 January, (b) 10 February, (c) 5 March, (d) 7 April, (e) 13 May.

6. Conclusions

This research analyzed and evaluated the four refined mathematical distance models to the application of forecasting short-term electricity prices. Apply the neural network to predict the LMP makes it closer to the actual LMP. Obviously, the performance using the neural network is better than the regression model forecasting method. Moreover, the simulation results show that using the Pearson correlation coefficient model can improve the accuracy of LMP obtained by the regression and neural network models when choosing the time interval of 45 days and similar days of three days.

Author Contributions: Conceptualization, C.-Y.L.; methodology, C.-Y.L.; software, C.-Y.L. and C.-E.W.; validation, C.-Y.L. and C.-E.W.; formal analysis, C.-Y.L. and C.-E.W.; investigation, C.-Y.L.; resources, C.-Y.L. and C.-E.W.; data curation, C.-Y.L.; writing—original draft preparation, C.-Y.L.; writing—review and editing, C.-Y.L.; visualization, C.-Y.L.; supervision, C.-Y.L.; project administration, C.-Y.L.; funding acquisition, C.-Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Ministry of Science and Technology of Republic of China under Contract MOST 108-2221-E-033-022-MY2.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Chitsaz, H.; Zamani, D.P.; Zareipour, H.; Parikh, P.P. Electricity price forecasting for operational scheduling of behind-the-meter storage systems. *IEEE Trans. Smart Grid* **2018**, *9*, 6612–6622. [\[CrossRef\]](#)
- Hubicka, K.; Marcjasz, G.; Weron, R. A note on averaging day-ahead electricity price forecasts across calibration windows. *IEEE Trans. Sustain. Energy* **2019**, *10*, 321–323. [\[CrossRef\]](#)
- Ranjbar, M.; Soleymani, S.; Sadati, N.; Ranjbar, A.M. Electricity price forecasting using artificial neural network. In Proceedings of the 2006 International Conference on Power Electronic, Drives and Energy Systems, New Delhi, India, 12 December 2006; pp. 1–5.
- Bissing, D.; Klein, M.T.; Chinnathambi, R.A.; Selvaraj, D.F.; Ranganathan, P. A hybrid regression model for day-ahead energy price forecasting. *IEEE Access* **2019**, *7*, 36833–36842. [\[CrossRef\]](#)
- Vu, D.H.; Muttaqi, K.M.; Agalgaonkar, A.P.; Bouzardoum, A. Short-term forecasting of electricity spot prices containing random spikes using a time-varying autoregressive model combined with kernel regression. *IEEE Trans. Indus. Inf.* **2019**, *15*, 5378–5388. [\[CrossRef\]](#)
- González, C.; McWilliams, M.J.; Juárez, I. Important variable assessment and electricity price forecasting based on regression tree models: Classification and regression trees, bagging and random forests. *IET Gener. Trans. Dist.* **2015**, *9*, 1120–1128. [\[CrossRef\]](#)
- Pourdaryaei, A.; Mokhlis, H.; Illias, H.A.; Kaboli, S.H.A.; Ahmad, S.; Ang, S.P. Hybrid ANN and artificial cooperative search algorithm to forecast short-term electricity price in de-regulated electricity market. *IEEE Access* **2019**, *7*, 125369–125386. [\[CrossRef\]](#)
- Alanis, A.Y. Electricity prices forecasting using artificial neural networks. *IEEE Lat. Am. Trans.* **2018**, *16*, 105–111. [\[CrossRef\]](#)
- Mosbah, H.; El-hawary, M. Hourly electricity price forecasting for the next month using multilayer neural network. *Can. J. Elect. Comp. Eng.* **2016**, *39*, 283–291. [\[CrossRef\]](#)
- Voronin, S.; Partanen, J.; Kauranne, T. A hybrid electricity price forecasting model for the Nordic electricity spot market. *Int. Trans. Elec. Energy Syst.* **2014**, *24*, 736–760. [\[CrossRef\]](#)
- Lee, C.Y.; Tuegeh, M. Optimal optimisation-based microgrid scheduling considering impacts of unexpected forecast errors due to the uncertainty of renewable generation and loads fluctuation. *IET Renew. Power Gen.* **2020**, *14*, 321–331. [\[CrossRef\]](#)
- Lee, C.Y.; Tuegeh, M. An optimal solution for smooth and non-smooth cost functions-based economic dispatch problem. *Energies* **2020**, *13*, 3721. [\[CrossRef\]](#)
- Shrivastava, N.A.; Khosravi, A.; Panigrahi, B.K. Prediction interval estimation of electricity prices using PSO-tuned support vector machines. *IEEE Trans. Ind. Inf.* **2015**, *11*, 322–331. [\[CrossRef\]](#)
- Peesapati, R.; Kumar, N. Electricity price forecasting and classification through wavelet–dynamic weighted PSO–FFNN approach. *IEEE Syst. J.* **2017**, *12*, 3075–3084.
- Pourdaryaei, A.; Mokhlis, H.; Illias, H.A.; Kaboli, S.H.A.; Ahmad, S. Short-term electricity price forecasting via hybrid backtracking search algorithm and ANFIS approach. *IEEE Access* **2019**, *7*, 77674–77691. [\[CrossRef\]](#)
- Darudi, A.; Bashari, M.; Javidi, M.H. Electricity price forecasting using a new data fusion algorithm. *IET Gener. Trans. Distr.* **2015**, *9*, 1382–1390. [\[CrossRef\]](#)
- González, J.P.; Muñoz San Roque, A.M.S.; Pérez, E.A. Forecasting functional time series with a new hilbertian armax model: Application to electricity price forecasting. *IEEE Trans. Power Syst.* **2018**, *33*, 545–556. [\[CrossRef\]](#)
- Bello, A.; Bunn, D.W.; Reneses, J.; Muñoz, A. Medium-term probabilistic forecasting of electricity prices: A hybrid approach. *IEEE Trans. Power Syst.* **2017**, *32*, 334–343. [\[CrossRef\]](#)

19. Khosravi, A.; Nahavandi, S.; Creighton, D. Quantifying uncertainties of neural network-based electricity price forecasts. *Appl. Energy* **2013**, *112*, 120–129. [[CrossRef](#)]
20. Chang, P.C.; Fan, C.Y.; Lin, J.J. Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network approach. *Int. J. Elect. Power Energy Syst.* **2011**, *33*, 17–27. [[CrossRef](#)]
21. Esmaili, M.; Shayanfar, H.A.; Moslemi, R. Locating series FACTS devices for multi-objective congestion management improving voltage and transient stability. *Eur. J. Oper. Res.* **2014**, *236*, 763–773. [[CrossRef](#)]
22. Bharatwaj, N.V.; Abhyankar, A.R.; Bijwe, P.R. Improved loss distribution and modeling in DCOPF. *Int. J. Elec. Power Energy Syst.* **2013**, *53*, 416–425. [[CrossRef](#)]
23. Sarkar, V.; Khaparde, S.A. Reactive power constrained OPF scheduling with 2-d locational marginal pricing. *IEEE Trans. Power Syst.* **2013**, *28*, 503–512. [[CrossRef](#)]
24. Lewis, G. Estimating the value of wind energy using electricity locational marginal price. *Energy Policy* **2010**, *38*, 3221–3231. [[CrossRef](#)]
25. Gautam, D.; Mithulananthan, N. Locating distributed generator in the LMP-based electricity market for social welfare maximization. *Electr. Power Compon. Syst.* **2007**, *35*, 489–503. [[CrossRef](#)]
26. Nozawa, T.; Konda, M.; Fujibayashi, M.; Imai, M.; Kotani, K.; Sugawa, S.; Ohmi, T. A parallel vector-quantization processor eliminating redundant calculations for real-time motion picture compression. *IEEE J. Solid State Circuits* **2000**, *35*, 1744–1751. [[CrossRef](#)]
27. Bastian, J.; Zhu, J.X.; Banunaryanan, V.; Mukerji, R. Forecasting energy prices in a competitive market. *IEEE Comp. Appl. Power* **1999**, *12*, 40–45. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).