IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received May 13, 2019, accepted June 3, 2019, date of publication June 12, 2019, date of current version June 27, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2922420

Short-Term Electricity Price Forecasting via Hybrid Backtracking Search Algorithm and ANFIS Approach

ALIREZA POURDARYAEI[1], HAZLIE MOKHLIS[1], (Senior Member, IEEE), HAZLEE AZIL ILLIAS[1], S. HR. AGHAY KABOLI[2], AND SHAMEEM AHMAD[1]

[1]Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur 50603, Malaysia
[2]Electrical Engineering Department, College of Engineering and Petroleum, Kuwait University, Safat 13060, Kuwait

Corresponding author: Hazlie Mokhlis (hazli@um.edu.my)

This work was supported by the University of Malaya (UM) under postgraduate research under Grant (PPP): PG323-2016A.

ABSTRACT In this paper, a hybrid electricity price forecasting method which is composed of two-stage feature selection method and optimized adaptive neuro-fuzzy inference system (ANFIS) technique as a forecasting engine is proposed to accurately forecast electricity price. A multi-objective feature selection approach comprises of multi-objective binary-valued backtracking search algorithm (MOBBSA) as an efficient evolutionary search algorithm and ANFIS method is developed in this paper to extract the most influential subsets of input variables with maximum relevancy and minimum redundancy. Through the combination of backtracking search algorithm (BSA) in learning process of ANFIS approach, a hybrid machine learning algorithm has been developed to forecast the electricity price more accurately. Real-world electricity demand and price dataset from Ontario power market; which is reported as among the most volatile market worldwide, has been used as case study to validate the performance of the proposed approach. From the simulation results, it has been seen that the proposed hybrid forecasting method was effective in accurately forecast the Ontario electricity price. In addition, to prove the superiority of the proposed hybrid forecasting method the simulation results obtained using ANN and ANFIS models optimized by other well-known optimization methods have been compared with that of proposed method.

INDEX TERMS Adaptive neuro-fuzzy inference system, backtracking search algorithm, electricity price forecasting, feature selection.

I. INTRODUCTION

When the reformation of electricity is introduced into the electric power industry, the electricity price has drawn all the attention in the power market. Electricity price forecast is key information for electricity market managers and participants. In addition, smart grid has already become a platform, which allows electricity providers to adjust their bidding strategies with respect to Demand-Side Management (DSM) models. Furthermore, reliability of system operation and capital cost investments can be improved by raising responsiveness of the electricity providers. Though various methodologies have been introduced for the forecasting of price, which mainly varies in the data preprocessing, model selection, calibration and testing phases, however, the literature is relatively limited in the field of price forecasting. The main reason for this limitation is the fact that most electricity providers were established based on monopoly systems, which has led to absence of or limited competitions in the market. However, the electricity market has now become more competitive. Hence, price forecasting is gaining its significance and many experts are beginning to take this factor into consideration. In fact, the energy market is now more open and continuously supporting competition, which makes price forecasting the center of attention of most electricity providers to invest more to introduce efficient and novel approaches. Moreover, electricity price forecasting (EPF) is a very valued tool for most electricity participants in this deregulated and competitive market. Electricity providers can apply prediction techniques to maximize their benefits and minimize their electricity cost by adjusting their production schedule and select the best bidding strategy.

The associate editor coordinating the review of this manuscript and approving it for publication was Kok Lim Alvin Yau.

Price forecasting guarantees strong, stable and well-organized operation of the power market. EPF plays a critical role when it comes to investment decisions and transmission expansion if the applied method comes with high accuracy. However, the forecasting with high accuracy is relatively a complicated task since several exclusive features are needed to be taken into consideration such as nonlinearity, high frequency, high volatility, multiple seasonality, mean reversion, price spikes and the calendar effect [1]. In addition, weather conditions and seasonal effects determine system load. On the other hand, prices are determined by a larger and diverse set of parameters [2]. Many of these parameters might not be accessible to researchers or engineers since they are classified and regulated due to the market competition. Therefore, various combinations of inputs should be examined, to develop an effective EPF price time series model.

Many methods have been proposed in past which are summarized in [2]–[5] to accurately forecast the electricity day-ahead price. For instance, in [6], a k-factor GIGARCH process has been applied to forecast price in the German electricity price market. ANN was used in [7] to forecast electricity demand and price. In [8], combinatorial neural network based forecasting engine was proposed to predict the future values of price data, while in [9] an iterative neural network based prediction technique has been used for load and price forecasting. Support Vector Machin (SVM) has been adopted in load forecasting area in [10]. An SVM model was adopted to estimate the electricity price prediction in [11], which measures the uncertainty by predicting the ranges of targeted quantities.

Apart from using the aforementioned data driven methods solely, they have been combined with other optimization algorithms to form machine learning methods to increase the forecasting methods accuracy. For example, in [12] the electricity price and ANN was combined with some data mining approach to forecast engine price. A combination of a feature selection technique and cascaded neuro-evolutionary algorithm (CNEA) was proposed for electricity price forecasting in [13]. Neural networks and stage selection feature system are combined in [14] for the purpose of price forecasting. Discrete cosine transforms based cascade-forward neural network approach was used to classify the electricity markets price forecasting of mainland Spain and New York in [15] and [16]. In [17] by combining a fuzzy system and ANN model (Neuro-Fuzzy) named ANFIS approach has been used to forecast electricity price. In [18], the accuracy of the forecasting and the non-linearity of the forecasted electricity price have been improved by integrating different techniques and forming a hybrid model. In [19], ANN, ANFIS, and ARIMA are adopted as the forecasters where Kalman filter is used to predict the price of electricity of a Spanish electricity system. In [20] three models are adopted but a modified ordered weighted average algorithm has been selected to do the prediction of price. In [21], a hybrid system is presented to predict the electricity prices in the Spanish market in which a mutual information method selects the inputs and feeds into the ANFIS. In [18], a hybrid approach was proposed for electricity price forecasting, which is based on a combination of ANFIS and evolutionary particle swarm optimization method.

Though many approaches have been applied to forecast electricity market price in the literature, however, all these approaches have possessed drawbacks. For instance, the major deficiencies of data driven approaches are that there are so many control parameters and they are very sensitive which make initialization of these parameters value very difficult. Although different machine learning methods have been applied for electricity price forecasting still new methodology is required to provide more accurate electricity price forecast. Furthermore, in most the aforementioned literature, long term price forecasting has been considered. However, the pattern of electricity demand depends on seasons, hence instead of long term price forecasting short term price forecasting will be more effective in deregulated electricity market for real time decision making. Another important factor that influences the efficiency and accuracy of the forecasting is feature selection. Proper selection of features helps to enhance the efficiency and accuracy of the forecasting. However, an observation has been made that it is very difficult to select the robust feature selection technique for the prediction of electricity price considering non linearity of price signal.

In this work, to resolve the above issues a hybrid electricity price forecasting methodology based on a combination of backtracking search algorithm (BSA) and ANFIS approach is proposed. In addition, to enhance the accuracy of the proposed hybrid price forecasting method a feature selection technique with the combination of multi-objective binary-valued backtracking search algorithm (MOBBSA) and ANFIS approach has been proposed. To examine the accuracy of the proposed methods, the Ontario electricity market has been considered as a case study. As, Ontario electricity market is recognized as one of the unstable markets due to its single settlement nature [22], [23]. Therefore, it leaves a massive challenge for those who are forecasting the electricity price.

The contributions of this paper are summarized as follows:

- One of the main contributions in this study is proposing a hybrid electricity price forecasting technique to provide more accurate forecasts through the use of an efficient BSA technique with ANFIS method. The BSA is applied to improve the forecasting accuracy of ANFIS by tuning the membership functions to achieve a lower error. BSA is applied in this study as its structure is simple but very effective in terms of solving multidimensional functions which makes it easy to adapt in different numerical problems. Moreover, BSA can be controlled by only one factor.
- To address the feature selection problem, a multi-objective feature selection technique has been proposed which is another contribution of this study. The proposed feature selection approach is composed of a multi-objective binary-valued backtracking search

algorithm (MOBBSA) and ANFIS approach. In the developed multi-objective feature selection method, MOBBSA is used to search within different combinations of input variables and to select the non-dominated feature subsets, while ANFIS is applied as evaluation metric to determine the performance of each feature subset.
- The proposed method can be examined carefully by comparing with data from ANN and ANFIS models optimized by other well-known optimization methods. Additionally, to assess the applicability of the proposed multi-objective feature selection technique for electricity price forecasting, its obtained results are compared with other well-known multi-objective optimization methods.

The reminder of this paper is structured as follows: The main principles of ANFIS are described in Section II, which is followed by the explanation of implementation of BSA in the learning process of ANFIS in section III. Then, section IV contains the explanation of mutual information technique that has been used to filter the redundant features. In Section V, the method of multi-objective BSA implementation in developing short-term EP forecasting models has been described. Simulation results along with the comparisons of proposed method performance with ANFIS-BSA and other artificial intelligence (AI) based approaches are presented in Section VI. In the same section, statistical analysis is also studied to achieve the robustness of the proposed method in forecasting the future electricity price. Finally, by consolidating the important features of the proposed work conclusion has been drawn.

II. ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM (ANFIS)

Jang [24] hybridized a fuzzy inference system (FIS) with ANN to form ANFIS to overcome the drawback of the fuzzy logic approach which lacks the ability to learn and adapts themselves to a new state due to predefined rules (if-then). In addition, this approach is capable to simulate complex nonlinear mappings using fuzzy system with ANN learning and it is considered as a universal estimator which predicts short, high and long term effect [25]–[27].

Generally, five different layers construct the ANFIS structure while each layer consists of node functions and the inputs of the nodes in the present layer are obtained from the previous layers [25]. The consecutive layers of ANFIS structure are: layer 1 is fuzzification (if-part), layer 2 is production part, layer 3 is normalization part, layer 4 is defuzzification (then-part), and eventually layer 5 is the total output generation part. Fig. 1 shows the structure of ANFIS with two independent variables (x and y) as input and one dependent variable f_{out} as an output.

For fuzzy inference systems, difference in the consequence of the set of fuzzy rules (if-then) and defuzzification procedures lead to two different types of fuzzy inference systems known as Mamdani type FIS and Sugeno type FIS.

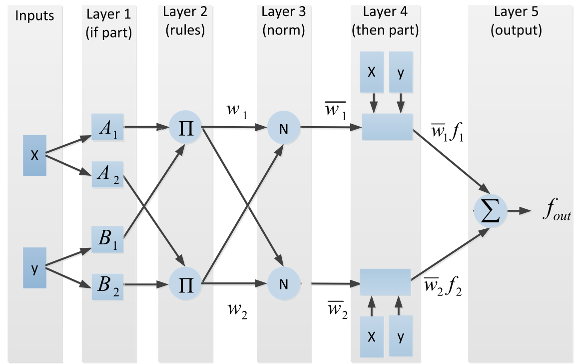

FIGURE 1. The general structure of ANFIS.

In this work, Sugeno type FIS has been chosen as Sugeno has more flexibility in system design which latter can be integrated with ANFIS tool to model the systems more precisely [28].

Considering ANFIS with Sugeno type FIS, the rule base of ANFIS contains fuzzy IF-THEN rules of a first order Sugeno type FIS and are stated as:

$$\begin{aligned}&\text{Rule 1}: \textit{ If } x \textit{ is } A_1 \textit{ and } y \textit{ is } B_1 \textit{ then } z \textit{ is } f_1\,(x, y; p_1,\\ &q_1, r_D) = x\,p_1 + y\,q_1 + r_1\\ &\text{Rule 2}: \textit{ If } x \textit{ is } A_2 \textit{ and } y \textit{ is } B_2 \textit{ then } z \textit{ is } f_2\,(x, y; p_2,\\ &q_2, r_2) = x\,p_2 + y\,q_2 + r_2 \end{aligned} \tag{1}$$

where, $f_i(x, y; p_i, q_i, r_i)$ is a first order polynomial function, which represents the outputs of the Sugeno type FIS, A_i and B_i are the fuzzy sets, and x and y are two different input and z is an output of ANFIS model.

In the ANFIS structure, different layers consist of different node function. As shown in Fig. 2, adaptive nodes which represent the adjustable parameter sets are denoted by squares whereas fixed nodes which represent the fixed parameter sets in the system are denoted by circles.

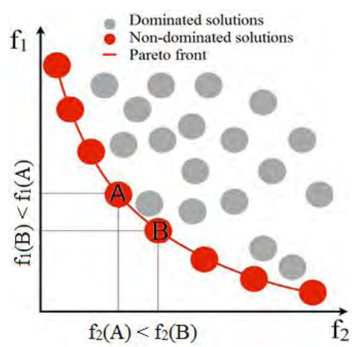

FIGURE 2. The Pareto optimal set for the two objective functions (A and B are two sample from non-dominated solutions).

A. LAYER 1

The node functions are adaptive ones and they are followed as:

$$Q_{1,i} = \mu_{Ai}(x), \quad i = 1, 2 \tag{2}$$

$$Q_{1,i} = \mu_{Bi-2}(y), \quad i = 3, 4 \tag{3}$$

where, x and y are the inputs to node i, A_i and B_i are linguistic labels, μ_{Ai} and μ_{Bi} are the membership functions for Ai and Bi fuzzy sets, respectively and $Q_{1,i}$ is the membership grade of a fuzzy set and considered as the output of node i in the first layer which specifies the degree to the given input (x or y) satisfies the quantifies.

Typically, in ANFIS, the MF (membership function) for a fuzzy set can be any parameterized membership function, such as generalized Bell shaped function, Gaussian, trapezoidal or triangular.

A generalized Bell shaped MF (bell MF) is specified as follows:

$$\mu_A = (x; a, b, c) = \frac{1}{1 + \left(\frac{x-c}{a}\right)^{2b}} \tag{4}$$

A Gaussian MF is specified as follows:

$$\mu_A = (x; c, \sigma) = e^{-0.5\left(\frac{x-c}{\sigma}\right)^2} \tag{5}$$

where, $\sigma$and c are the determined width and center of Guassian MF, respectively.

A trapezoidal MF is specified as follows:

$$\mu_A = (x; a, b, c, d) = \max(\min(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}), 0) \tag{6}$$

The parameters with a < b ≤ c < d specify the x coordinates of the four corners for the underlying trapezoidal MF.

A triangular MF is specified as follows:

$$\mu_A = (x; a, b, c) = \max(\min(\frac{x-a}{b-a}, \frac{c-x}{c-b}), 0) \tag{7}$$

The parameters with a < b < c specify the x coordinates of the three corners for the underlying triangular MF. Where in this layer, the parameters *a, b, c, d* and σ are the antecedent parameters.

B. LAYER 2

Every node in this layer is a fixed node whose output is the product of all the incoming signals. In this layer, through multiplication of input signals, the firing strength of each rule is determined.

$$Q_{2,i} = w_i = \mu_{Ai}(x)\mu_{Bi}(y) \quad i = 1, 2 \tag{8}$$

where, w_i is the output signal which represents the firing strength of a rule.

C. LAYER 3

Every node in this layer is a fixed node. In this layer, the firing strength provided in previous layer is normalized by computing the ratio of the i$^{\text{th}}$ rule's firing strength to the sum of all rules' firing strengths.

$$Q_{3,i} = \bar{w_l} = \frac{w_i}{w_1 + w_2} \quad i = 1, 2 \tag{9}$$

where, $\bar{w_l}$ is the output signal which represents the normalized firing strength of a rule.

D. LAYER 4

In this layer, every node i is adaptive with a node function.

$$Q_{4,i} = \bar{w_l} f_i \quad i = 1, 2 \tag{10}$$

where, f_1 and f_2 are the fuzzy IF-THEN rules as follows:

$$\begin{aligned} &Rule1 : If\ x\ is\ A_1\ and\ y\ is\ B_1\ then\ z = f_1(x, y; p_1, q_1, r_1) \\ &Rule2 : If\ x\ is\ A_2\ and\ y\ is\ B_2\ then\ z = f_2(x, y; p_2, q_2, r_2) \end{aligned} \tag{11}$$

where r_i , q_i and p_i are the parameter set, referred to as the linear consequent parameters.

E. LAYER 5

This layer has only one fixed node that computes the overall output of ANFIS by summation of all incoming signals.

$$Q_{5,i} = \sum_i \bar{w_l} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} = overaloutput \quad i = 1, 2 \tag{12}$$

The overall output is linear combination of the consequent parameters. Thus, the final output of ANFIS is expressed as:

$$\begin{aligned} f_{out} &= \bar{w}_1 f_1 + \bar{w}_2 f_2 = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 \\ &= \left(\bar{w}_1 x\right) p_1 + \left(\bar{w}_2 x\right) p_2 + \left(\bar{w}_1 y\right) q_1 + \left(\bar{w}_2 y\right) q_2 \\ &\quad + \left(\bar{w}_1\right) r_1 + \left(\bar{w}_2\right) r_2 \end{aligned} \tag{13}$$

Eventually, ANFIS applies a hybrid learning algorithm for parameters tuning. It utilizes the back propagation algorithm and the least squared method for updating the input MF parameters (antecedent parameters) in layer 1, and training the consequent parameters, respectively.

III. BACKTRACKING SEARCH ALGORITHM (BSA)

BSA is recently developed as evolutionary optimization algorithm with simple structure and it is highly effective in solving complex optimization problems. This capability allows BSA to be effortlessly adapted to a wide range of numerical optimization problems. A single control parameter and two advanced cross-over and mutation operators are normally used in this algorithm to balance the exploitation of better results and search space. This feature of BSA has confirmed it as an enhanced model to solve high multimodal optimization benchmarks compared to other evolutionary methods (DE, GA) [29]. Lacking the learning from the optimal individual (lack of elitist selection) can be considered as the only disadvantage of BSA. The general framework including the relevant equations of BSA model is described in Table 1.

IV. MULTI-OBJECTIVE BACKTRACKING SEARCH ALGORITHM (MOBSA)

In the context of multi-objective optimization, instead of unique solution, there is a Pareto optimal set corresponding to the optimal value of each objective. Considering the two solutions from Pareto optimal set as denoted by $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_N)$

TABLE 1. General framework of BSA.

Step 1: Initialization Scattering the population members in the solution space(**Eq. 14**)	$$P_{i,j;g=0} \sim U(low_j, up_j)\ ,\ y_i = f(P_i)\quad for\quad i=\{1,2,3,\ldots,nPop\},\ j=\{1,2,3,\ldots,nVar\} \tag{14}$$ where: *nPop* is population size. *nVar* signifies the optimization variable. Uniform distribution function is *U*. low_j and up_j are lower and upper search space limits of jth variable. y_i is productivity of ith individual. *g* is generation number.
Step 2: Selection-I 1) Initializing a historical population (**old P**) to determine the search-direction matrix (**Eq. 15**) 2) Redefining the historical population at each iteration based on (if-then) rule by comparing two random number ***a*** and ***b***. Subsequently population(**P**) pursues ***old P*** until it is changed to provide a memorization process and facilitate the exploration search space (**Eq.16**) 3) At the end of **step 2** , a hierarchical sequence has been permuted by shuffling random function (**Eq.17**)	$$oldP_{i,j} \sim U(low_j, up_j) \tag{15}$$ where, *OldP* is historical population $$if\ a<b\,\Big|_{a,b\sim U(0,1)}\ then\ oldP := P \tag{16}$$ where, := is the updated operation. *a* and *b* are randomly generated numbers $$oldP := permuting(oldP) \tag{17}$$ where, permuting *(oldP)* is a random shuffling function.
Step 3: Mutation The Wiener process (***F***) is implemented to control the amplitude of the search matrix according to (**Eq.18**)	$$Mutant = P + F.(oldP - P) \tag{18}$$ $$F = 3.rndn\,\Big|\ rndn \sim N(0,1)$$ where, *N* is standard normal distribution
Step 4: Crossover Determine the binary integer–valued matrix (***map***) and control parameter of individuals in **BSA** according (**Eq.19**)	$$\begin{aligned} & map_{i,j}=1, \\ & if\ a<b\,\Big|_{a,b\sim U(0,1)}\ then \\ & map_{i,u_{(1:\lceil mixrate.\ rnd\ .\ nVar\rceil\ )}}=0\ \Big|\ rnd\sim U(0,1),\ u=permuting(1,2,3,\ldots,nVar) \\ & else \\ & map_{i,randi(nVar)}=0 \\ & T := Mutant \\ & if\ \ map_{i,j}=1\ then\ T_{i,j} := P_{i,j} \\ & for \\ & i=\{1,2,3,\ldots,nPop\},\ j=\{1,2,3,\ldots,nVar\} \end{aligned} \tag{19}$$ where, *mixrate* is the control parameter of optimization algorithm.
Step 5: Boundary control At the end of **step 4**, if an individual in generated offspring (***T***) violates the boundary condition, the control mechanism developed in **step 5** is updated according to (**Eq.20**)	$$if\ \ (T_{i,j}<low_j)\ or\ (T_{i,j}>up_j)\ then\ T_{i,j}\sim U(low_j, up_j) \tag{20}$$ where, *T* is generated offspring
Step 6: Selection- II Calculating the fitness and the position(**Eq.21**)	$$\begin{aligned} & if\ \ f(T_i)<y_i\ then\ ,\ y_i := f(T_i),\ P_i := T_i \\ & y_g = min(\,f(P_g)) \\ & if\ \ y_g < y_{g-1}\ then\ \ global\,minimum := y_g\ ,\ global\,minimizer := P_g\ , \\ & g=g+1 \\ & for \\ & i=\{1,2,3,\ldots,nPop\}\ ,g=\{1,2,3,\ldots,gMax\} \end{aligned} \tag{21}$$

and $\partial_1 = (\partial_1, \ldots, \partial_N)$ and their corresponding objective functions represented by $f(\varepsilon) = (f_1(\varepsilon), \ldots, f_m(\varepsilon))$ and $f(\partial) = (f_1(\partial), \ldots, f_m(\partial))$. The solution ∂ is dominated by the solution ε, denoted by $f(\varepsilon) < f(\partial)$, if and only if the conditions described in Eq. (22) are satisfied. Therefore, the solution ε is considered as a non-dominated solution.

$$\begin{aligned} &\forall i \in \{1, \ldots, m\} : f_i(\varepsilon) \leq f_i(\partial) \\ &\exists j \in \{1, \ldots, m\} : f_j(\varepsilon) < f_j(\partial) \end{aligned} \tag{22}$$

For two objective functions denoted by f_1 and f_2, the Pareto optimal set is illustrated in Fig. 2, where the dominated solutions are represented by the grey circles and the Pareto optimal set of two objectives (f_1 and f_2) is represented by red dots.

Multi-objective optimization yields Pareto front as a set of optimal solutions rather than a single optimal. Any solution in the Pareto front is not inferior to another. Improvement to one objective cannot be achieved without sacrificing another. A trade-off between solutions should lead to the best compromise solution. As objectives of multi-objective optimization are on different scales, a normalization technique is utilized to provide a notionally common scale for all objectives. Then, all the normalized values of each Pareto set are added together and the lowest value of Pareto set is selected as the best compromise solution.

A. EXTERNAL ELITIST ARCHIVE

An elitist reservation mechanism is adopted in this study as an external elitist archive to update and retain the non-dominated solutions in each generation of BSA. Initially the external elitist archive is empty; within the optimization progresses, it stores the non-dominated solutions according to the following 'if-then' rules:

I. If the trial pattern (a new generated solution) dominates some of the archived elitist, then all dominated members of the external elitist archive are replaced by non-dominated trial pattern;
II. If the trial pattern is dominated by at least one member of the external elitist archive, then the trial pattern is disregarded for elitist archive;
III. If the archived members of the external elitist archive are not dominated by the trial pattern and trial pattern is not dominated by the archived members, then the external elitist archive retains the trial pattern as a new elitist member (non-dominated solution)

As the optimization progresses, the members of external elitist archive increases. Thus, to prevent overpopulation of the external elitist archive, the crowding distance of all members is measured and the extra members of the archive are removed according to their crowding distance value.

B. CROWDING DISTANCE

To keep the external elitist archive to its maximum capacity, the crowding distances (CD) of all solutions in Pareto-front (external elitist archive) are computed and the solution with the lowest CD value is subject to deletion when the archive is overloaded. The crowding distance is a factor to evaluate the distribution of the solutions in Pareto -front by measuring the density around a solution. The crowding distances compute the distance of two neighbor points around the solution. The crowding distances of i^{th} solution in the Pareto -front is calculated by

$$CD_i = \sum_{j=1}^{m} \frac{f_j(i+1) - f_j(i-1)}{f_j^{\max} - f_j^{\min}} \tag{23}$$

where, CD_i is crowding distance of solution i. f_j is the j^{th} objective function. f_j^{max} and f_j^{min} are maximum and minimum values of the j^{th} objective function respectively. m is number of objectives. For the boundaries solutions (f_j^{max} and f_j^{min}), the crowding distance is set to infinite as there is only one neighbor point for those solutions.

C. PROCEDURE OF MULTI-OBJECTIVE BSA (NON-DOMINATED APPROACH)

In the context of multi-objective optimization, instead of unique solution, there is a Pareto optimal set corresponding to the optimal value of each objective. Thus, to extend BSA to a multi-objective optimization approach, the replacement mechanism is adapted according to the concept of Pareto dominance [30]. Similar to BSA, in multi-objective BSA mutation and crossover operators are first applied to produce the offspring (T). Then, the comparison in the final step of BSA (export the global minimum) is modified according to the concept of Pareto dominance to replace i^{th} individual of population (P_i) by i^{th} individual in offspring (Ti) if the P_i is dominated by T_i. The consecutive steps in BSA algorithm are followed to form the multi-objective BSA, expect the last step (export the global minimum). Instead of exporting a global minimum, a Pareto optimal set is stored. The steps necessary are given below for multi-objective BSA to measure distance:

Step 1: The initial population (P) equals to the dimension of the optimization variables is randomly generated according to Eq. (14).

Step 2: Fitness function is determined and non-dominated solutions are stored for each initial population.

Step 3: The historical population is determined randomly according to Eq. (15).

Step 4: The historical population is updated at each iteration through the 'if-then' decision rule according to Eq. (16)

Step 5: Apply the mutation operator to generate an initial form of trial population (Mutant) according to Eq. (18).

Step 6: Apply the crossover operator over the initial form of trial population obtained in pervious step to generate the final form of offspring (T) according to Eq. (19).

Step 7: At the end of the crossover process, if an individual in the generated offspring (T) violates the boundary condition of the optimization problem, the related individual in offspring is updated according to the boundary control developed by Eq. (20).

Step 8: If the i$^{\text{th}}$ element of generated offspring (T_i) dominates the i$^{\text{th}}$ element of population (P_i), then T_i is replaced by P_i.

Step 9: Update the external elitist archive according to aforementioned 'if-then' rules.

Step 10: If the external elitist archive exceeds its maximum capacity, the crowing distance for all members of elitist archive are computed and the less crowded solutions are removed from the archive one after another.

Step 11: If the stopping criteria are not satisfied, set $g = g + 1$ and return to the *step 4*.

V. MUTUAL INFORMATION (MI) TECHNIQUE

MI technique has been widely implemented in electricity market price forecasting to observe the mutual relationship between two random variables, X and Y [31]. It measures the amount of information that is obtained by one variable with respect to the other random variable. As a result, all the information conveyed by one variable will be shared by the other variable. Thus, reduce the uncertainty about one random variable that knows the other random variable. Therefore, the mutual information will be the same as the uncertainty obtained by either one of the variables.

In [32], the fundamental concept of the information theory is explained in terms of entropy $H(X)$, which measures the uncertainty of random variable X. Higher entropy indicates more uncertainties about a random variable. If the random variable of X is in a discrete value with the probability of *P(X), H(X)* can be defined as follows:

$$H(X) = -\sum_{i=1}^{n} P(X_i)\log_2(P(X_i)) \tag{24}$$

For the continuous random variable, the $H(X)$ is defined as follows:

$$H(X) = -\int P(X)\log_2(P(X))dX \tag{25}$$

As the concept of mutual information is intricately linked by the entropy of the random variable, the mutual information between random variables X and Y, *MI(X, Y)* with the joint probability distribution of P_{XY} *(X, Y)* can be defined as follows:

$$MI(X,Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} P(X_i,Y_j)\log_2(P(X_i,Y_j)) \tag{26}$$

$$MI(X,Y) = -\iint P(X,Y)\log_2(P(X,Y))dXdY \tag{27}$$

Besides the entropy, there is a conditional entropy, which measures the average uncertainty of the first random variable after the second random variable is observed. For example, the conditional entropy, *CE* for random variable X after observing the random variable Y with the conditional probability distribution of *P(X|Y)* is expressed as

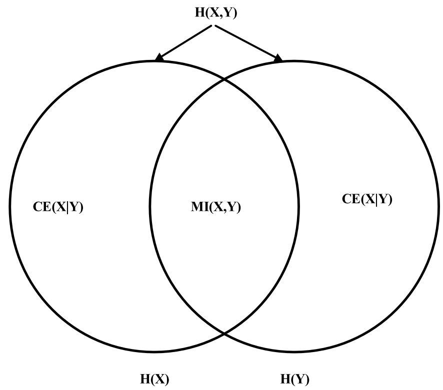

FIGURE 3. Graphical representation between mutual information and conditional entropy.

follows:

$$CE(X|Y) = \sum_{j=1}^{m} P(Y_j)\left[-\sum_{i=1}^{n} P(X_i|Y_j)\log_2(P(X_i|Y_j))\right] \tag{28}$$

$$CE(X|Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} P\left(X_i,Y_j\right)\log_2\left(P\left(X_i|Y_j\right)\right) \tag{29}$$

$$P\left(X_i|Y_j\right) = \frac{P\left(X_i|Y_j\right)}{P\left(Y_j\right)} \tag{30}$$

Fig. 3 shows the graphical representation of the relationship between the *MI* and *CE*. If *MI* is large enough, it indicates that X and Y are closely related and dependent on each other. If X and Y are independent variables, then *MI(X, Y)* will be equal to zero and thus $P(X, Y) = P(X)P(Y)$.

The relationship between *MI* and *CE* can be referred as the chain-rule and it can be calculated as follows:

$$MI(X,Y) = H(Y) - CE(Y|X) = H(X) - CE(Y|X) \tag{31}$$

$$H(X,Y) = H(X) + CE(Y|X) = H(Y) + CE(X|Y) \tag{32}$$

VI. ELECTRICITY PRICE FORECASTING (EPF)

In the deregulated electricity market, electricity price is a function of electricity demand. It can be considered that in electricity price forecasting process, electricity price at time t depends both on electricity demand and past values of electricity demand and price which is showed as follows (33), as shown at the top of the next page, where, *EP(t)* and *ED(t)* present the electricity price and demand at time t while they are assumed as a time series with t interval, NL_{EP} presents the number of lag order for the electricity price and similarly NL_{ED} denotes electricity demand lag order.

The hybrid ANFIS-BSA is applied in this study to forecast hourly Ontario electricity price (HOEP) according to two types of input historical data. The input historical data sets of HOEP and hourly Ontario electricity demand (HOED) in 2017 are adapted from [22]. Due to the wide range of historical data, Eq (34) is utilized to normalize both the dependent and independent variables. The main feature of data normalization is adjusting raw data observed on different

$$EP(t) = f\begin{pmatrix} EP(t-1), EP(t-2), EP(t-3), \ldots, EP(t-NL_{EP}), \\ ED(t), ED(t-1), ED(t-2), ED(t-3), \ldots, ED(t-NL_{ED}) \end{pmatrix} \tag{33}$$

scales to a notionally common scale, often prior to data processing.

$$\overline{Z}(t) = \frac{Z(t) - min(Z)}{max(Z) - min(Z)} + 1 \tag{34}$$

where, $\overline{Z}$ is the normalized value, Z is the value to be normalized, and t is hourly interval.

Assuming only one week with hourly lagged values of exogenous variables (NL_{EP}=NL_{ED} =168) are used to forecast electricity price, this constitutes 336 lagged values of exogenous variables. If all aforementioned exogenous variables are applied in the forecasting process, this may cause slow down the learning process, and also degrades the performances and overfit will happen in the training data. In other words, although these factors are of vital importance for electricity price forecasting process, only those features exhibiting significant influence on the output should be picked.

Though all subsets of variables have causal relation with electricity price, it is neither efficient nor feasible to employ all of them as inputs and assess their performances. Instead, an efficient feature selection is applied to select the most effective subset of variables in model construction. A search technique is combined with a learning algorithm (evaluation metric) for construction of feature selection methods which are classified into three categories, namely wrappers, filters, and embedded [33]. In this study, wrapper method has been adopted as search technique. As wrapper methods train a new predictive model for each candidate subsets, they often provide the best performing feature set for that particular type of model at the expense of computationally intensive tasks. Though wrapper methods often provide the most relevant features for particular type of model, but it requires a systematic searching algorithm in their evolutionary training process. Hence, the randomized search algorithm known as metaheuristic algorithm is proposed to explore the space of different feature subsets. Metaheuristic algorithms (GA, SA, ACO and PSO) incorporate randomness into search procedure of feature selection to escape local optima [33]. These metaheuristic algorithms deal with feature selection as a single objective optimization problem, so the number of relevant features should be predefined and always find the subset of features with fixed number of features. Generally, feature selection has two main conflicting objectives, which are minimizing simultaneously both the estimation error and the number of features. Therefore, feature selection problem can be expressed as a multi-objective problem that has two main objectives, maximizing the accuracy of model and minimizing the number of features and the decision is a tradeoff between these two objectives. Treating feature selection as a multi-objective problem leads to a set of non-dominated feature subsets to meet different requirements in real-world applications.

Existing multi-objective feature selection algorithms like multi-objective particle swarm optimization (MOPSO), NSGA-II suffer from the problems of high computational cost. In addition, the control patterns and parameters are many [35]–[37]. In contrast, statistical analyses in [30] confirm that multi-objective BSA (MOBSA) is a promising optimization method for solving high dimensional multi-objective problems over MOPSO and NSGA-II. Therefore, in this study binary-valued BSA (BBSA) based multi-objective feature selection algorithm is developed as a promising technique to generate a Pareto front of non-dominated feature subsets.

To assess the performance of each candidate feature subsets ANFIS is used in the evolutionary training process of MOBSA feature selection algorithm due to its fast learning capability to approximate nonlinear functions [25]; hence, it is adopted as an evaluation metric method in proposed multi-objective wrapper-based feature selection method. In particular, ANFIS employs an efficient hybrid learning method that combines the least squares method and gradient descent. The least square method is the main factor for quick training [38]. Thus, after only few epoch of training, ANFIS is able to construct the predictive model.

Before applying the feature selection to extract the most influential subsets of input variables with maximum relevancy and minimum redundancy for short-term EPF, both dependent and independent variables are randomly divided into two sets: 70% as the training set and 30% as the test set. The training set is used to construct the ANFIS models with different subsets of input variables, while the test set is used to access the strength and utility of generated models.

In the developed multi-objective feature selection method, MOBBSA is used to search within different combinations of input variables and selects the non-dominated feature subsets, while ANFIS is applied as evaluation metric to determine the performance of each feature subset. During the training process of applied learning method (ANFIS), each individual of MOBBSA represents one input variable. Feature selection has two main conflicting objectives of maximizing the forecasting accuracy and minimizing the number of features [35]. Since the goal of feature selection technique is extracting the most influential subsets of input variables with maximum relevancy and minimum redundancy, the number of input variables and root mean square error (RMSE) are selected as targets in MOBBSA feature selection technique. The developed feature selection strategy uses an elitist external archive to store non-dominated feature subsets, which simultaneously minimize both the root mean square error (RMSE) on the test set and the number of input variables as the optimal solution

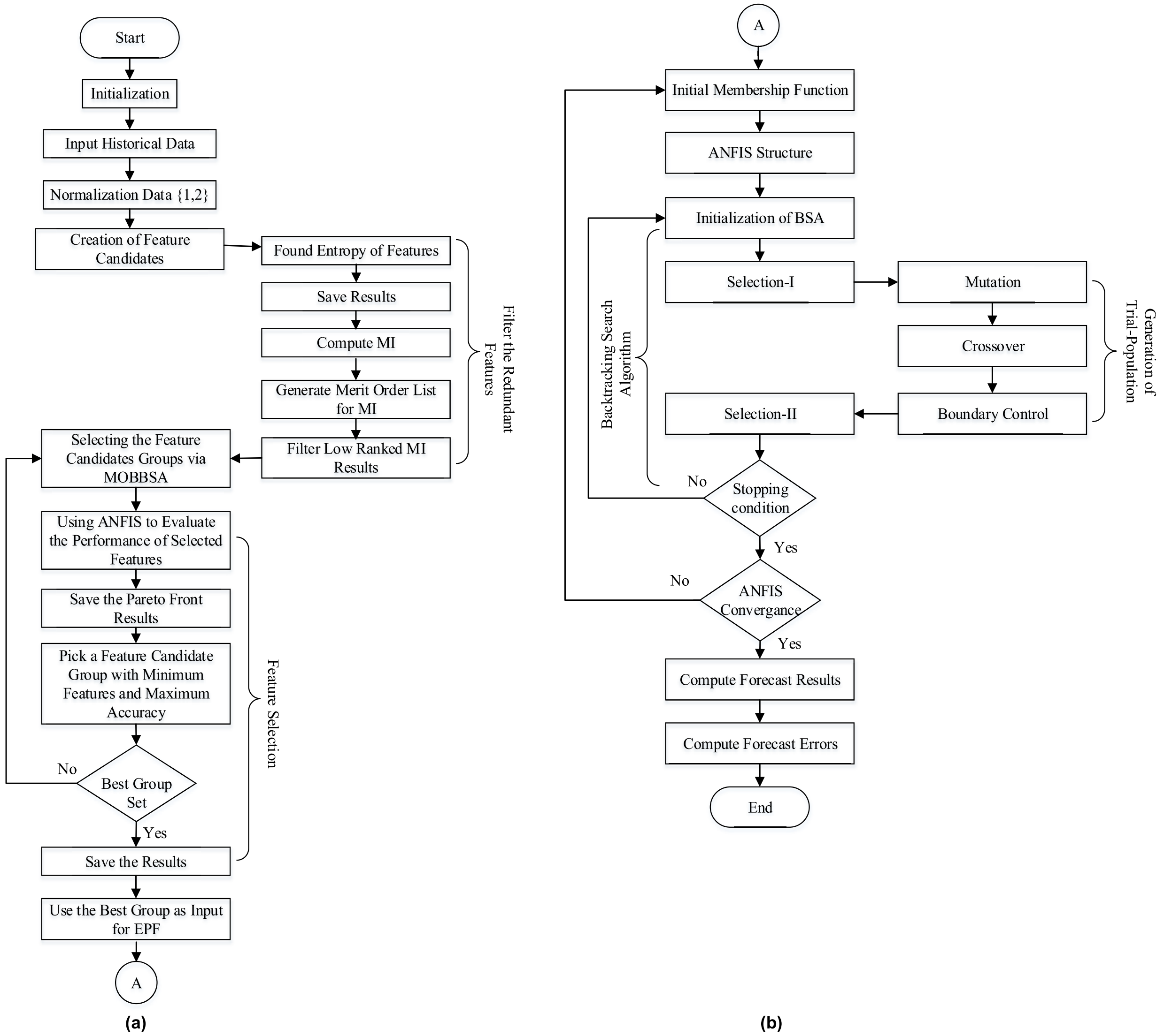

FIGURE 4. The flowchart of (a) Two-stage feature selection framework and (b) Electricity price forecasting process.

set according to the concept of Pareto dominance. MOBBSA as an extension of BSA has only one control parameter named "*mixrate*", which controls the number of individuals to be engaged in the crossover process. Value of 100 is set as population of individuals and highest value of *mixrate* (for example 100%) has been measured in the developed feature selection strategy to engage all the individuals in the crossover process.

To form the structure of ANFIS for feature selection, the Sugeno-type FIS is designed. Based on the work conducted in [39], the scatter partitioning can be used to facilitate the training process of ANFIS in feature selection. Since subtractive clustering is a realization of scatter partitioning, it is used to set up the ANFIS for feature selection.

In order to reduce running time of feature selection, two stages feature selection is proposed in this work. The stages are presented in Fig 4 (a). In the first stage of proposed feature selection technique, the mutual information between each individual input variable and output feature is computed according to (27). As the value goes high, it indicates that each input variable and output is very dependent on each other. To filter the redundant features, the relevancy threshold is considered as TH = 0.46. After the filtering process, the most relevant attributes are selected as 60 features. In the 2nd stage, MOBBSA is used to search non-dominated feature subsets, while ANFIS is applied as evaluation metric to determine the performance of each feature subset. In this stage, among 60 selected candidates, most relevance and dissimilar are 23 features which are selected by hybrid MOBBSA & ANFIS and later used as inputs for the next forecasting process.

Additionally, to calculate the efficiency of the suggested multi-objective feature selection method, comparison has been done with 5 different hybrid well known optimization

TABLE 2. The optimal subsets of input variables selected by different multi-objective feature selection techniques and their corresponding performances in terms of RMSE Value.

MOBBSA+ANFIS	MOPSO+ANFIS	NSGA-II+ANFIS	MOBBSA+ANN	MOPSO+ANN	NSGA-II+ANN
EP(t-1)	EP(t-1)	EP(t-1)	EP(t-1)	EP(t-1)	EP(t-1)
EP(t-2)	EP(t-24)	EP(t-2)	EP(t-2)	EP(t-2)	EP(t-2)
EP(t-23)	EP(t-25)	EP(t-3)	EP(t-24)	EP(t-3)	EP(t-3)
EP(t-24)	EP(t-47)	EP(t-24)	EP(t-25)	EP(t-24)	EP(t-24)
EP(t-47)	EP(t-72)	EP(t-48)	EP(t-48)	EP(t-25)	EP(t-25)
EP(t-71)	EP(t-95)	EP(t-72)	EP(t-49)	EP(t-48)	EP(t-48)
EP(t-95)	EP(t-120)	EP(t-95)	EP(t-72)	EP(t-49)	EP(t-49)
EP(t-119)	EP(t-144)	EP(t-119)	EP(t-73)	EP(t-72)	EP(t-72)
EP(t-143)	EP(t-168)	EP(t-121)	EP(t-96)	EP(t-73)	EP(t-73)
EP(t-167)	EP(t-169)	EP(t-143)	EP(t-97)	EP(t-96)	EP(t-96)
EP(t-191)	EP(t-191)	EP(t-144)	EP(t-120)	EP(t-97)	EP(t-97)
EP(t-335)	EP(t-335)	EP(t-167)	EP(t-121)	EP(t-120)	EP(t-120)
EP(t-504)	EP(t-336)	EP(t-168)	EP(t-144)	EP(t-121)	EP(t-121)
ED(t)	EP(t-503)	EP(t-192)	EP(t-145)	EP(t-144)	EP(t-144)
ED(t-1)	ED(t)	EP(t-193)	EP(t-168)	EP(t-145)	EP(t-145)
ED(t-23)	ED(t-1)	EP(t-335)	EP(t-169)	EP(t-168)	EP(t-168)
ED(t-71)	ED(t-23)	EP(t-336)	EP(t-192)	EP(t-169)	EP(t-169)
ED(t-96)	ED(t-24)	EP(t-503)	EP(t-193)	EP(t-192)	EP(t-192)
ED(t-120)	ED(t-96)	EP(t-504)	ED(t)	EP(t-193)	EP(t-193)
ED(t-144)	ED(t-120)	ED(t)	ED(t-1)	ED(t)	ED(t)
ED(t-168)	ED(t-144)	ED(t-1)	ED(t-24)	ED(t-1)	ED(t-1)
ED(t-192)	ED(t-168)	ED(t-2)	ED(t-72)	ED(t-3)	ED(t-24)
ED(t-335)	ED(t-191)	ED(t-24)	ED(t-96)	ED(t-24)	ED(t-25)
	ED(t-192)	ED(t-48)	ED(t-120)	ED(t-72)	ED(t-48)
	ED(t-336)	ED(t-72)	ED(t-144)	ED(t-96)	ED(t-49)
	EP(t-504)	ED(t-96)	ED(t-168)	ED(t-120)	ED(t-72)
		ED(t-120)	ED(t-336)	ED(t-144)	ED(t-96)
		ED(t-144)		ED(t-168)	ED(t-120)
		ED(t-336)		EP(t-192)	ED(t-144)
					ED(t-168)
					ED(t-504)
RMSE					
12.23	13.05	13.87	13.47	13.85	14.74

techniques namely MOPSO+ANFIS, NSGA-II+ANFIS, MOBSA+ANN, MOPSO+ANN and NSGA-II+ANN. The optimal subsets of input variables selected by studied multi-objective feature selection methods and their corresponding performances in terms of RMSE value are tabulated in Table 2. The *EP (t)* and *ED (t)* in Table 2 are EP and ED at t-time as optimal subsets of input variables selected by different multi-objective feature selection techniques. Thus, for electricity price forecasting only the value of these selected inputs should be predefined as they are input subset of forecasting engine. According to the obtained results, the suggested selection technique is very much efficient than other during the same test, as it provides less estimation error and number of features.

VII. SIMULATION RESULTS AND DISCUSSION

In this study, ANFIS-BSA is employed to enhance the accuracy of electricity price forecasting (EPF) of HOEP, as its electricity market is recognized as the one of the unstable market due to its single settlement nature. The most effective input variables are selected by proposed feature selection method. In addition, to demonstrate the effectiveness of short-term EPF using ANFIS-BSA, the obtained results are compared with the following

TABLE 3. Parameter setting of applied methods.

Methods		Parameters	Value
ANN	MLP	Hidden layer	2
		Transfer function	logarithmic sigmoid
		Learning algorithm	Levenberg-Marquardt PB
ANFIS	SC (Subtractive clustering)	Cluster radius	0.8
		FIS structure	Sugeno-type
		Membership function	Gaussian
Metaheuristic optimization	GA	Population	Size:100,Type: double vector, Creation function: uniform
		Selection function	Stochastic uniform
		Elite count	5.0
		Crossover	Fraction: 0.8, Function: scattered
		Mutation	Function: Gaussian, Scale:1.0, Shrink:1.0
		Migration	Direction: forward, Fraction:0.2, Interval:20
	PSO	Swarm population	100
		w	[0.4, 0.9]
		c1=c2	2
	BSA	Number of individuals	100
		Control parameter rate (P)	100%

techniques: ANN, ANFIS, ANN-GA, ANN-PSO, ANFIS-GA, ANN-BSA and ANFIS-PSO.

The framework of proposed methodology for short-term electricity price forecasting (EPF) is illustrated in Fig. 4 (b). Generally, in performing the short-term EPF based on optimal AI models, the following steps are implemented.

a) The electricity price is measured with hourly intervals which is a role of electricity price and demand in preceding hours. The particular data of Ontario electricity market in 2017 (HOEP and HOED) selected by proposed feature selection method are taken as variables which do not depend on any variable and HOEP is calculated as a dependent variable.

b) As the pattern of electricity demand seasonally changes, one month from each season (i.e., February, May, August and November) is considered to assess the effectiveness of applied methods for forecasting the electricity price in different seasons. There are two types of variables; such as dependent and independent variables. They are divided into two subdivisions, where the hourly data of first three weeks is used for training of design phase and that last week of each month is utilized for testing phase of models which are obtained from the analysis.

c) The learning process occurs in the training phase. The independent variables are connected to dependent variables by the computer programs which are derivative as a learning process. According to (34) it can be made faster with the help of normalizing input and output.

d) Although this phase of testing does not have any role in developing the models, it is utilized to evaluate AI-based methods performance. By caring out MAPE (mean absolute percentage error) and RMSE (root mean square error), performance of model in predicting the EPF is demonstrated. Thiel's inequality coefficient (U-statistic) is also such type of performance index. The mathematical equations of these indexes are as follows (35)–(37), as shown at the bottom of this page, U-statistic is a measure to show how well-fitted a time series of forecasted values are to a corresponding time series of actual data. It is always between 0 and 1. The closer to zero, the more accurate forecasting will be obtained with a perfect fit. Hence, the closer to 1, the less precision in the forecasting value or even a simple guess if the U-static is 1.

e) In order to ensure that the achieved models are sufficient to describe a given data series, Whiteness test (Durbin-Watson test) is implemented [40]. This test is

$$MAPE\% = \frac{1}{N}\sum_{t=1}^{N}\frac{|(EP(t)_{actual} - EP(t)_{estimated}|}{EP(t)_{actual}} \times 100 \tag{35}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(EP(t)_{\mathrm{actual}} - EP(t)_{estimated})^2} \tag{36}$$

$$U = \frac{RMSE}{\sqrt{\frac{1}{N}\sum_{t=1}^{N}(EP(t)_{actual})^2} + \sqrt{\frac{1}{N}\sum_{t=1}^{N}(EP(t)_{estimated})^2}} \tag{37}$$

TABLE 4. Performance indexes.

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
RACF	Training	0.0005	0.0021	0.0017	0.0002	0.0012	0.0145	0.0045	0.0055
	Testing	0.0007	0.0009	0.0014	0.0005	0.0023	0.0034	0.0103	0.0037
	Whole set	0.0003	0.0012	0.0001	0.0019	0.0029	0.0006	0.0025	0.0143
Absolute error	Training	20.0043	19.0045	18.1357	17.7653	19.5678	17.6546	16.9876	16.7089
	Testing	7.0543	6.0654	5.9876	5.5673	6.7864	5.1328	4.8987	4.7936
	Whole set	27.0586	25.0699	24.1233	23.3326	26.3542	22.7874	21.8863	21.5026
RMSE	Training	0.0875	0.0853	0.0823	0.0789	0.0815	0.7654	0.7001	0.0683
	Testing	0.0845	0.0812	0.0802	0.0775	0.0798	0.0734	0.0678	0.0628
	Whole set	0.0867	0.0834	0.0821	0.0781	0.0801	0.0745	0.0699	0.0670
U-statistic	Training	0.0449	0.0413	0.0404	0.0367	0.0409	0.0353	0.0321	0.0317
	Testing	0.0411	0.0396	0.0384	0.0346	0.0386	0.0332	0.0309	0.0297
	Whole set	0.0438	0.0409	0.0391	0.0355	0.0395	0.0345	0.0327	0.0312
MAPE (%)	Training	5.0345	4.8965	4.5615	4.0018	4.8763	3.5683	3.0054	2.8777
	Testing	4.8865	4.4445	4.2347	3.7654	4.7899	3.0643	2.7647	2.5237
	Whole set	4.9818	4.5679	4.3781	3.8967	4.8011	3.1268	2.8964	2.7892

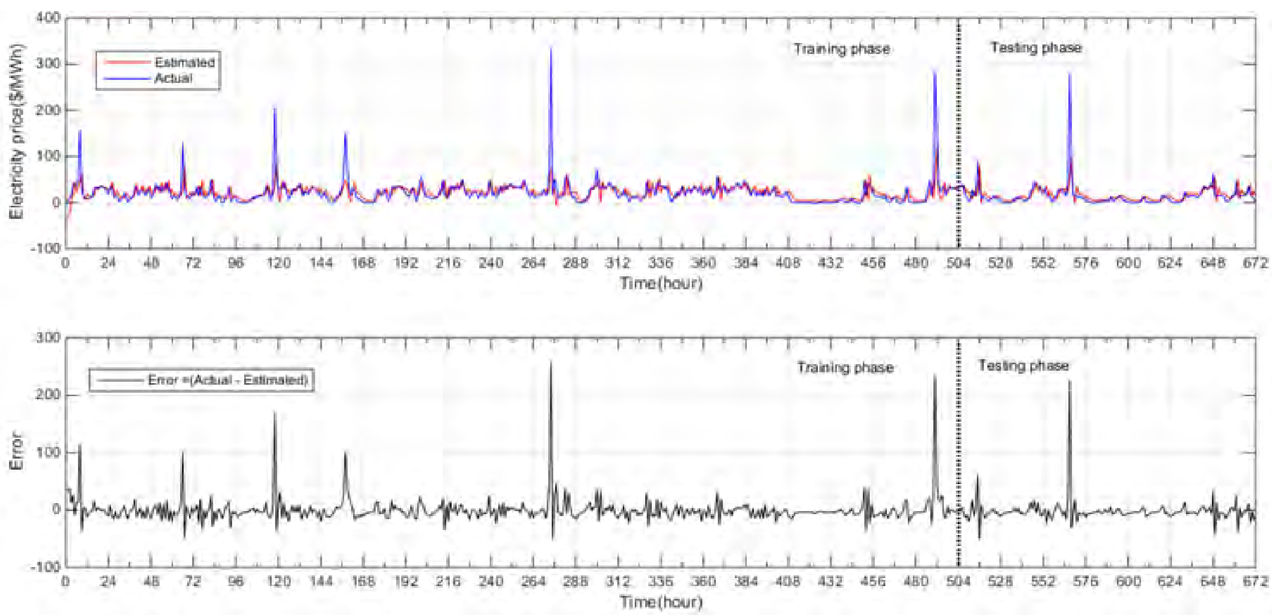

FIGURE 5. The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Ontario in February.

obtained through a confirmatory analysis. The purpose of confirmatory analysis is to ensure the whiteness of estimated residuals *(e(t))*. The whiteness of estimated residuals implies that they are uncorrelated. Residuals autocorrelation function (RACF) is utilized to analyze the correlation of whiteness of estimated residuals by the following equation:

$$RACF = \frac{\left|\sum_{i=2}^{N}(e(t)e(t-1))\right|}{\sum_{i=1}^{N}(e(t))^2} \tag{38}$$

The *RACF* values can be in the range of 0 to 1. When *RACF* value is far from zero, it falls far from the confidence level. It means that a vital independent variable has been neglected and the residuals are not white (correlated).

The control parameters of adopted methods are usually set based on the similar methodologies have been successfully applied in the literature for energy price or demand forecasting since there is no unanimity in the optimal values of the AI-based methods parameters setting. Table 3 summarizes all parameter settings of applied methods.

The machine learning method are summarized in Table 4 for Ontario EPF in February 2017. It is observed that the calculated *RACF* values validate the whiteness of probable

TABLE 5. Comparison between forecasting accuracy of studied methods for epf of ontario in may 2017.

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
RACF	Training	0.0002	0.0033	0.0021	0.0005	0.0030	0.0019	0.0014	0.0004
	Testing	0.0011	0.0016	0.0009	0.0003	0.0004	0.0008	0.0019	0.0010
	Whole set	0.0004	0.0005	0.0024	0.0015	0.0013	0.0026	0.0005	0.0016
Absolute error	Training	6.4578	6.2346	6.0045	5.8941	6.2354	5.7654	6.0007	5.5994
	Testing	2.5436	2.2549	2.1643	1.9745	2.4543	1.8435	2.1003	1.3796
	Whole set	9.0014	8.4895	8.1688	7.8686	8.6897	7.6089	8.1010	6.9790
RMSE	Training	0.0321	0.0301	0.0288	0.0267	0.0293	0.0251	0.0279	0.0247
	Testing	0.0245	0.0244	0.0229	0.0194	0.0251	0.0178	0.0201	0.0133
	Whole set	0.0311	0.0298	0.0262	0.0243	0.0271	0.0239	0.0256	0.0226
U-statistic	Training	0.0194	0.0184	0.0155	0.0145	0.0164	0.0121	0.0129	0.0119
	Testing	0.0107	0.0092	0.0082	0.0077	0.0089	0.0069	0.0071	0.0065
	Whole set	0.0191	0.0178	0.0147	0.0141	0.0155	0.0118	0.0124	0.0110
MAPE (%)	Training	1.3223	1.2132	1.2001	1.0059	1.2267	0.9278	0.9976	0.8987
	Testing	1.1397	1.1014	1.0009	0.9879	1.1165	0.8963	0.9181	0.7957
	Whole set	1.2466	1.1887	1.0679	0.9977	1.1999	0.9145	0.9459	0.8754

TABLE 6. Comparison between forecasting accuracy of studied methods for epf of ontario in august 2017.

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
RACF	Training	0.0011	0.0014	0.0012	0.0001	0.0010	0.0006	4.4E-5	0.0003
	Testing	0.0002	0.0018	0.0009	5.6E-6	0.0007	0.0003	0.0015	0.0005
	Whole set	0.0004	0.0003	0.0008	0.0013	0.0002	0.0004	0.0004	0.0005
Absolute error	Training	16.6789	16.5432	14.5431	15.3456	16.3458	13.0432	12.9807	12.4135
	Testing	3.9875	3.8976	3.3247	3.4568	3.4567	3.0543	2.7896	2.5635
	Whole set	15.7646	15.5432	15.2123	15.3002	15.3246	15.0004	14.9996	14.9770
RMSE	Training	0.0682	0.0646	0.0603	0.0612	0.0653	0.0601	0.0598	0.0581
	Testing	0.0357	0.0345	0.0312	0.0323	0.0337	0.0278	0.0254	0.0223
	Whole set	0.0643	0.0637	0.5987	0.0601	0.0632	0.5432	0.5362	0.0522
U-statistic	Training	0.0348	0.3246	0.3001	0.0314	0.0335	0.2798	0.2781	0.0264
	Testing	0.0198	0.0188	0.0178	0.0185	0.0189	0.0132	0.0115	0.0102
	Whole set	0.0289	0.0278	0.0261	0.0269	0.0267	0.0256	0.0241	0.0237
MAPE (%)	Training	2.7897	2.6543	2.4321	2.5638	2.4567	2.2271	1.9861	1.7909
	Testing	2.6543	2.5328	2.3456	2.4538	2.2456	1.9643	1.7632	1.4117
	Whole set	2.7065	2.5895	2.3953	2.5431	2.3689	2.0048	1.8732	1.7053

residuals of affirmed confidence range. Based on the data of Table 4, the predicting methods on Ontario electricity market according to different-criteria which made decisions are ordered as ANFIS-BSA > ANFIS-PSO > ANFIS-GA > MLP-BSA > MLP-PSO > MLP-GA > ANFIS > MLP. ANFIS method of estimation provides better result as shown in Table 4 than other studied methods.

Additionally, BSA is found the most efficient algorithm for controlling the ANFS, while ANFIS-BSA achieved to $MAPE = 2.79\%$, U-statistic $= 0.03$, $RMSE = 0.07$ and absolute error = 21.5, which is very less in value according to other algorithms. Fig. 5 presents the performance test of ANFIS-BSA when the training is executed for design phase and testing phase.

Machine learning methods performances for EPF of Ontario in May 2017 are tabulated in Table 5. From table 5 it is confirmed by the *RACF* values that all the estimated values are in the white confidence interval level. So, it can be said from the table that ANFIS-BSA is the most efficient model, where the values are of *MAPE* (0.87%),U-statistic (0.01), *RMSE* (0.02) and absolute error (6.98) belong to this model. The performances of ANFIS–BSA method for Ontario EPF in May is illustrated in Fig. 6.

Table 6 shows the performances of the applied methods for EPF of Ontario in August 2017. From this table it is confirmed by the *RACF* values that all the estimated values are in the white confidence interval level. Based on the data of Table 6, the predicting methods on Ontario electricity market according to different-criteria which made decisions are ordered as ANFIS-BSA > ANFIS-PSO > ANFIS-GA > ANFIS > MLP-BSA > MLP-GA > MLP-PSO > MLP. It can be deduced from the comparison that optimized ANFIS approaches are better than other methods. Table 6 shows the value of superior *MAPE* (1.7%), U-statistic (0.02), *RMSE* (0.05) and absolute error (14.98) which belong to ANFIS-BSA. Therefore, it is concluded that the

TABLE 7. Statistical factors of the anfis-bsa model for epf of ontario in november 2017.

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
RACF	Training	0.0012	0.0020	0.0018	0.0022	0.0011	0.0003	0.0012	0.0015
	Testing	0.0017	0.0023	0.0011	0.0013	0.0020	0.0008	0.0028	0.0019
	Whole set	0.0002	0.0016	0.0007	0.0004	0.0016	0.0010	0.0006	0.0009
Absolute error	Training	28.0031	27.9832	27.5421	26.3454	27.6067	27.0003	23.0045	22.6287
	Testing	8.2345	8.0046	7.8743	7.2451	7.9812	6.0321	5.0001	4.4944
	Whole set	36.2376	35.9878	35.4164	33.5905	35.5879	33.0324	28.0046	27.1231
RMSE	Training	0.0826	0.0819	0.0811	0.0802	0.0814	0.0804	0.0799	0.0798
	Testing	0.0794	0.0743	0.0720	0.0701	0.0785	0.0676	0.0503	0.0431
	Whole set	0.0803	0.0786	0.0754	0.0704	0.0793	0.0743	0.0737	0.0729
U-statistic	Training	0.0302	0.0298	0.0287	0.0278	0.0283	0.0276	0.0241	0.0202
	Testing	0.0497	0.0478	0.0463	0.0456	0.0482	0.0451	0.0389	0.0358
	Whole set	0.0392	0.0378	0.0364	0.0356	0.0375	0.0355	0.0341	0.0331
MAPE (%)	Training	4.6429	4.5329	4.4723	4.4011	4.4963	4.1212	3.7591	3.4031
	Testing	4.0428	3.9797	3.8346	3.7543	3.8203	3.3451	2.7521	2.4180
	Whole set	4.5465	4.3458	4.0768	3.9732	4.3103	3.7543	3.3541	3.1732

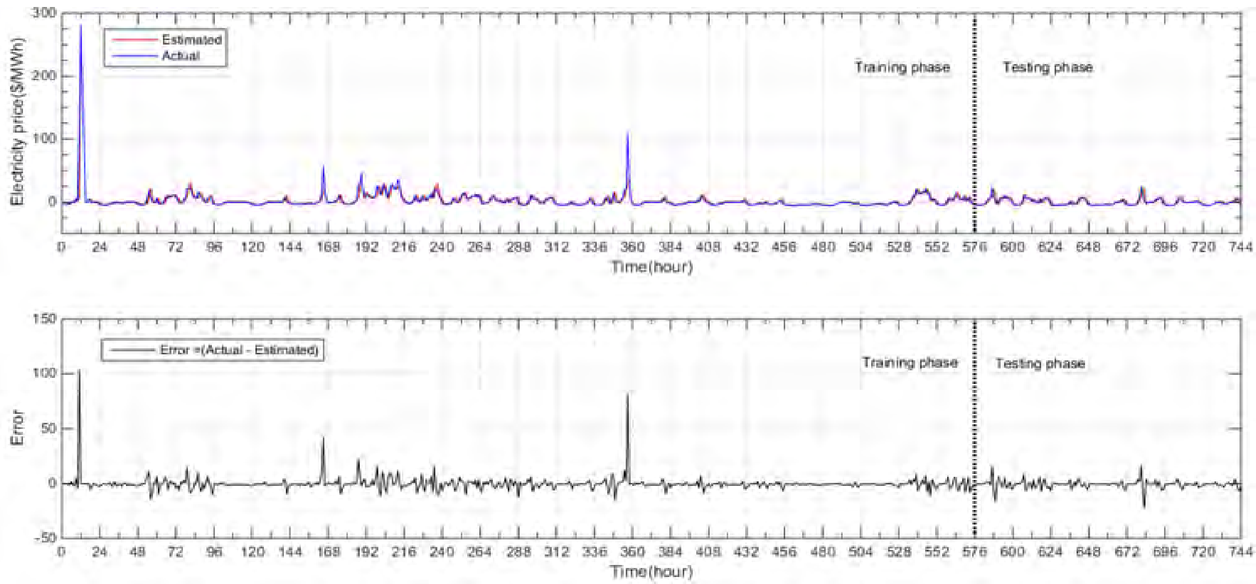

FIGURE 6. The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Ontario in May.

most efficient optimization algorithm for training ANFIS is BSA. Design phase and testing phase data are depicted in Fig. 7 which shows the ANFIS-BSA performance for forecasting Ontario electricity price in August.

Table 7 shows the comparison of soft computing approaches of performance of optimized ANFIS for EPF of Ontario in November 2017. From the table 7 *RACF* values, it is observed that all residual models are uncorrelated and this tables shows that it adequately describes data set. Based on the data of Table 7, the predicting methods on Ontario electricity market according to different-criteria which made decisions are ordered as ANFIS-BSA > ANFIS-PSO > ANFIS-GA > MLP-BSA > MLP-PSO > ANFIS > MLP-GA > MLP. It can be deduced from the comparison that optimized ANFIS approaches are better than other methods.

Table 7 also depicts the value of superior *MAPE* (3.17%), *U*-statistic (0.03), *RMSE* (0.07) and absolute error (27.12) which are belong to ANFIS-BSA. Fig. 8 demonstrates the performance of ANFIS-BSA method for forecasting Ontario's electricity price in November.

Different statistical methods are also applied as external validation to verify the validity of models developed by ANFIS-BSA. To assess the performance of the obtained model the following qualities were recommended [41]:

I. When a model shows $|R| > 0.8$, it is meant that a strong correlation is there between predicted and observed values.
II. When a model provides $0.2 < |R| < 0.8$, there is a correlation which is existing between the predicted and observed values.

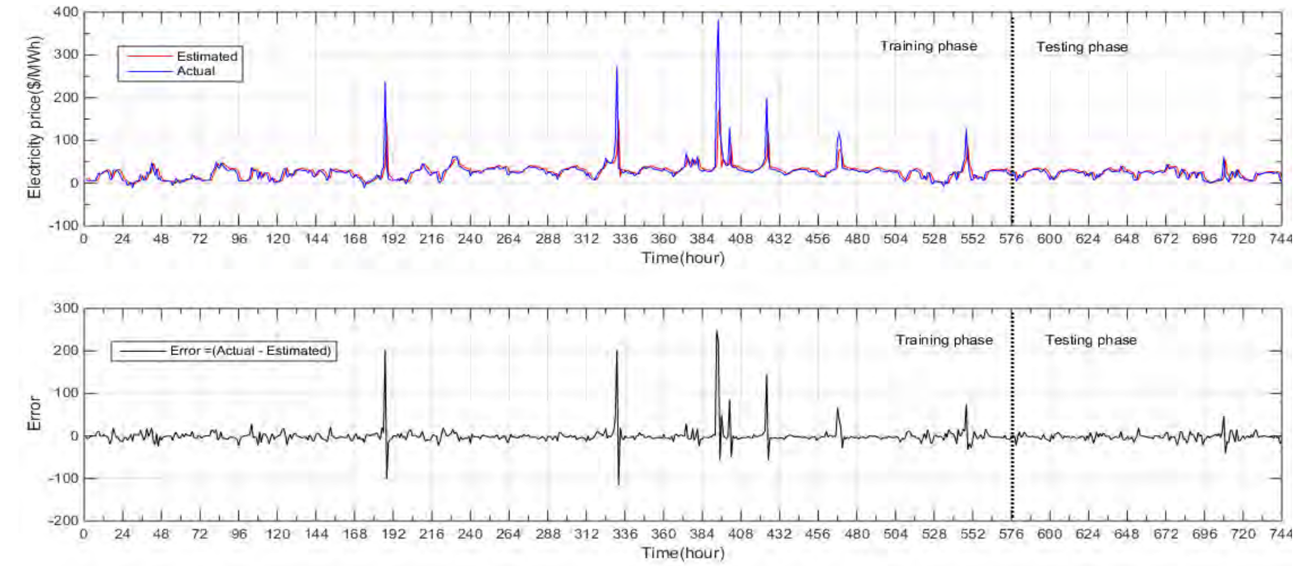

FIGURE 7. The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Ontario in August.

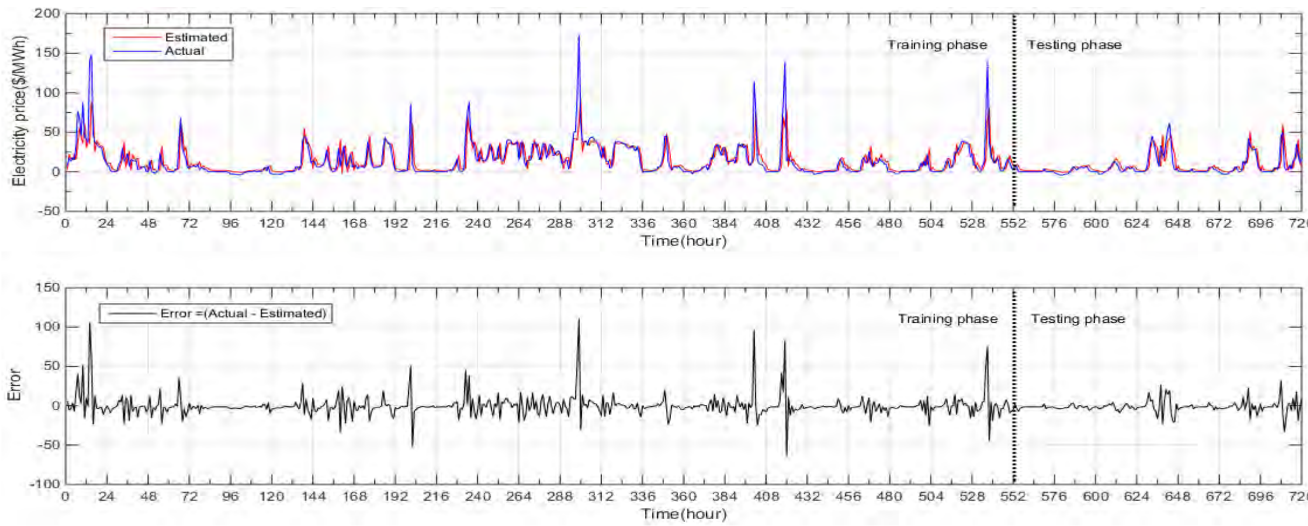

FIGURE 8. The performance of ANFIS-BSA during training of design phase and testing phase and its corresponding error for EPF of Ontario in November.

III. When a model shows $|R| < 0.2$, there is a weak correlation.

In addition, new factors suggested by [42] are checked for external validation of the obtained models on the testing phase. One slope of the regression lines (k or k') should be close to one through the origin. k or k' slopes are amid of the regressions of actual output (h_i) against predicted output (t_i) or vice versa and it goes through the origin. This can be written as $h_i \ =k\ t_i$ and $t_i \ =k'\ h_i$, correspondingly.

Moreover, the value of m and n should be less than 0.1. Here, m and n are two factors which are used for performance of variables. Recently, a confirmed indicator (R_m) is presented by Roy and Roy [43]. The condition suits when $R_m > 0.5$. From [44] it can be seen that the squared correlation coefficient (through the origin) (R_o^2) or the squared correlation coefficient between experimental and predicted values ($R_o'^2$) should be close to one. R should be greater than 0.8 in first item. In case of second one, value of k must be in between 0.85 and 1.15. For the item not three, the value ofk' must be in between 0.85 and 1.15. According to items four and five, the values of m and n values should be less than 0.1. Finally, R_m should be greater than 0.5.

Table 8 tabulates the statistical factors of the ANFI-BSA model for EPF of Ontario in February, May, August and November 2017. As shown in the table, the developed models satisfy all the requisite conditions. The validation phase ensures that ANFIS-BSA provides precise models, which is

TABLE 8. Statistical factors of the ANFIS-BSA model for EPF of Ontario in February, May, August and November 2017.

Item	Formula	Condition	ANFIS-BSA February	ANFIS-BSA May	ANFIS-BSA August	ANFIS-BSA November
1	R	$0.8<R_0$	0.9992	0.9977	0.9966	0.9992
2	$K=\frac{\sum_{i=1}^{n}(h_i\times t_i)}{\sum_{i=1}^{n}h_i^2}$	0.85<k<1.15	0.9989	0.9991	0.9959	0.9983
3	$K^{'}=\frac{\sum_{i=1}^{n}(h_i\times t_i)}{\sum_{i=1}^{n}t_i^2}$	0.85<k'<1.15	1.0010	1.0006	1.0038	1.0016
4	$m=\frac{R^2-R_o^2}{R^2}$	$\lvert m \rvert<0.1$	-0.0015	-0.0045	-0.0056	-0.0016
5	$n=\frac{R^2-R_{o^{'}}^2}{R^2}$	$\lvert n \rvert<0.1$	-0.0015	-0.0045	-0.0057	-0.0016
6	$R_m=R^2\times\left(1-\sqrt{\lvert R^2-R_o^2\rvert}\right)$	$0.5<R_m$	0.9970	0.9910	0.9881	0.9967
Where	$R_o^2=1-\frac{\sum_{i=1}^{n}(t_i-h_i^o)^2}{\sum_{i=1}^{n}\left(t_i-\bar{t}_i\right)^2}, h_i^o=k\times t_i$	$0.8<R_0^2<1$	1.0000	1.0000	0.9996	0.9999
	$R_{o^{'}}^2=1-\frac{\sum_{i=1}^{n}(h_i-t_i^o)^2}{\sum_{i=1}^{n}\left(h_i-\bar{h}_i\right)^2}, t_i^o=K^{'}\times h_i$	$0.8<R_{0^{'}}^2<1$	1.0000	1.0000	0.9997	0.9999

strongly applicable for short term electricity price forecasting of Ontario.

VIII. CONCLUSION

This research has proposed a hybrid approach for day ahead electricity price prediction in the Ontario electricity market containing a multi-objective feature selection technique and hybrid forecast engine.

- From the obtained results, it has been seen that the number of selected features is 23 and RMSE value is 12.23 for the proposed MOBBSA feature selection method which are less than the features number (26, 29, 26, 29 & 32) and RMSE values (13.05, 13.87, 13.47, 13.85 & 14.74) obtained by using other multi-objective approaches namely MOPSO + ANFIS, NSGA-II + ANFIS, MOBAS + AN, MOPSO + ANN and NSGA-II + ANN respectively.
- The proposed hybrid BSA and ANFIS forecasting approach provided a higher forecasting accuracy with the least complexity for electricity price forecasting in term of, $MAPE = 2.79\%$, 0.87%, 1.7% and 3.17% in February, May, August and November respectively compared to other artificial intelligence (AI) models.
- On the basis of the results obtained, it can be concluded that the ANFIS approach optimized by BSA could be considered as robust and useful forecast engine to the actual needs of electricity market participants, including the self-producers and traditional generation companies, the suppliers or retailers and aggregators. These contributions may help market participants to bid effectively, maintaining efficient daily operation and eventually increasing the company's profit.

REFERENCES

[1] Z. Yang, L. Ce, and L. Lian, "Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods," *Appl. Energy*, vol. 190, pp. 291–305, Mar. 2017.
[2] I. P. Panapakidis and A. S. Dagoumas, "Day-ahead electricity price forecasting via the application of artificial neural network based models," *Appl. Energy*, vol. 172, pp. 132–151, Jun. 2016.
[3] S. K. Aggarwal, L. M. Saini, and A. Kumar, "Electricity price forecasting in deregulated markets: A review and evaluation," *Int. J. Electr. Power Energy Syst.*, vol. 31, pp. 13–22, Jan. 2009.
[4] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *Int. J. Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014.
[5] M. Cerjan, I. Krželj, M. Vidak, and M. Delimar, "A literature review with statistical analysis of electricity price forecasting methods," in *Proc. IEEE EUROCON*, Jul. 2013, pp. 756–763.
[6] A. K. Diongue, D. Guégan, and B. Vignal, "Forecasting electricity spot market prices with a k-factor GIGARCH process," *Appl. Energy*, vol. 86, pp. 505–510, Apr. 2009.
[7] S. K. Aggarwal, L. M. Saini, and A. Kumar, "Electricity price forecasting in Ontario electricity market using wavelet transform in artificial neural network based model," *Int. J. Control, Autom., Syst.*, vol. 6, no. 5, pp. 639–650, 2008.
[8] O. Abedinia, N. Amjady, M. Shafie-Khah, and J. P. S. Catalão, "Electricity price forecast using combinatorial neural network trained by a new stochastic search method," *Energy Convers. Manage.*, vol. 105, pp. 642–654, Nov. 2015.
[9] N. Amjady and A. Daraeepour, "Mixed price and load forecasting of electricity markets by a new iterative prediction method," *Electric Power Syst. Res.*, vol. 79, pp. 1329–1336, Sep. 2009.

[10] H. Hahn, S. Meyer-Nieberg, and S. Pickl, "Electric load forecasting methods: Tools for decision making," *Eur. J. Oper. Res.*, vol. 199, pp. 902–907, Dec. 2009.
[11] A. Shiri, M. Afshar, A. Rahimi-Kian, and B. Maham, "Electricity price forecasting using support vector machines by considering oil and natural gas price impacts," in *Proc. IEEE Int. Conf. Smart Energy Grid Eng. (SEGE)*, Aug. 2015, pp. 1–5.
[12] N. Amjady and A. Daraeepour, "Design of input vector for day-ahead price forecasting of electricity markets," *Expert Syst. Appl.*, vol. 36, pp. 12281–12294, Dec. 2009.
[13] N. Amjady and F. Keynia, "Day-ahead price forecasting of electricity markets by mutual information technique and cascaded neuro-evolutionary algorithm," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 306–318, Feb. 2009.
[14] N. Amjady and F. Keynia, "Day-ahead price forecasting of electricity markets by a new feature selection algorithm and cascaded neural network technique," *Energy Convers. Manage.*, vol. 50, pp. 2976–2982, Dec. 2009.
[15] S. Anbazhagan and N. Kumarappan, "Classification of day-ahead deregulated electricity market prices using DCT-CFNN," in *Proc. Int. Conf. Swarm, Evol., Memetic Comput.*, 2013, pp. 499–510.
[16] S. Anbazhagan and N. Kumarappan, "Day-ahead deregulated electricity market price classification using neural network input featured by DCT," *Int. J. Elect. Power Energy Syst.*, vol. 37, pp. 103–109, May 2012.
[17] H. M. I. Pousinho, V. M. F. Mendes, and J. P. S. Catalão, "Short-term electricity prices forecasting in a competitive market by a hybrid PSO–ANFIS approach," *Int. J. Elect. Power Energy Syst.*, vol. 39, pp. 29–35, Jul. 2012.
[18] J. P. S. Catalão, H. M. I. Pousinho, and V. M. F. Mendes, "Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 137–144, Feb. 2011.
[19] J. P. S. Catalão, S. J. P. S. Mariano, V. M. F. Mendes, and L. A. F. M. Ferreira, "Short-term electricity prices forecasting in a competitive market: A neural network approach," *Electr. Power Syst. Res.*, vol. 77, pp. 1297–1304, Aug. 2007.
[20] J. P. S. Catalão, S. J. P. S. Mariano, V. M. F. Mendes, and L. A. F. M. Ferreira, "An artificial neural network approach for short-term electricity prices forecasting," in *Proc. Int. Conf. Intell. Syst. Appl. Power Syst. (ISAP)*, Nov. 2007, pp. 1–6.
[21] J. P. S. Catalão, H. M. I. Pousinho, and V. M. F. Mendes, "Neural networks and wavelet transform for short-term electricity prices forecasting," in *Proc. 15th Int. Conf. Intell. Syst. Appl. Power Syst. (ISAP)*, Nov. 2009, pp. 1–5.
[22] IESO. Accessed: Nov. 17, 2018. [Online]. Available: http://www.ieso.ca/
[23] H. Zareipour, K. Bhattacharya, and C. A. Cañizares, "Electricity market price volatility: The case of Ontario," *Energy Policy*, vol. 35, pp. 4739–4748, Sep. 2007.
[24] J.-S. R. Jang, "ANFIS: Adaptive-network-based fuzzy inference system," *IEEE Trans. Syst., Man, Cybern.*, vol. 23, no. 3, pp. 665–685, May/Jun. 1993.
[25] M. Tavana, A. Fallahpour, D. Di Caprio, and F. J. Santos-Arteaga, "A hybrid intelligent fuzzy predictive model with simulation for supplier evaluation and selection," *Expert Syst. Appl.*, vol. 61, pp. 129–144, Nov. 2016.
[26] W. Gao, H. Moayedi, and A. Shahsavar, "The feasibility of genetic programming and ANFIS in prediction energetic performance of a building integrated photovoltaic thermal (BIPVT) system," *Solar Energy*, vol. 183, pp. 293–305, May 2019.
[27] W. Gao, V. Sarlak, M. R. Parsaei, and M. Ferdosi, "Combination of fuzzy based on a meta-heuristic algorithm to predict electricity price in an electricity markets," *Chem. Eng. Res. Des.*, vol. 131, pp. 333–345, Mar. 2018.
[28] I. Svalina, V. Galzina, R. Lujić, and G. Šimunović, "An adaptive network-based fuzzy inference system (ANFIS) for the forecasting: The case of close price indices," *Expert Syst. Appl.*, vol. 40, pp. 6055–6063, Nov. 2013.
[29] P. Civicioglu, "Backtracking search optimization algorithm for numerical optimization problems," *Appl. Math. Comput.*, vol. 219, no. 15, pp. 8121–8144, Apr. 2013.
[30] M. Modiri-Delshad and N. A. Rahim, "Multi-objective backtracking search algorithm for economic emission dispatch problem," *Appl. Soft Comput.*, vol. 40, pp. 479–494, Mar. 2016.
[31] N. Amjady and M. Hemmati, "Energy price forecasting-problems and proposals for such predictions," *IEEE Power Energy Mag.*, vol. 4, no. 2, pp. 20–29, Mar./Apr. 2006.
[32] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.
[33] E. T. Renani, M. F. M. Elias, and N. A. Rahim, "Using data-driven approach for wind power prediction: A comparative study," *Energy Convers. Manage.*, vol. 118, pp. 193–203, Jun. 2016.
[34] A. Unler and A. Murat, "A discrete particle swarm optimization method for feature selection in binary classification problems," *Eur. J. Oper. Res.*, vol. 206, no. 3, pp. 528–539, 2010.
[35] B. Xue, M. Zhang, and W. N. Browne, "Particle swarm optimization for feature selection in classification: A multi-objective approach," *IEEE Trans. Cybern.*, vol. 43, no. 6, pp. 1656–1671, Dec. 2013.
[36] B. Xue, M. Zhang, and W. N. Browne, "Particle swarm optimisation for feature selection in classification: Novel initialisation and updating mechanisms," *Appl. Soft Comput.*, vol. 18, pp. 261–276, May 2014.
[37] C. Zhang, J. Zhou, C. Li, W. Fu, and T. Peng, "A compound structure of ELM based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting," *Energy Convers. Manage.*, vol. 143, pp. 360–376, Jul. 2017.
[38] M. Alizadeh, F. Jolai, M. Aminnayeri, and R. Rada, "Comparison of different input selection algorithms in neuro-fuzzy modeling," *Expert Syst. Appl.*, vol. 39, pp. 1536–1544, Jan. 2012.
[39] J.-S. R. Jang, "Input selection for ANFIS learning," in *Proc. 5th IEEE Int. Conf. Fuzzy Syst.*, Sep. 1996, pp. 1493–1499.
[40] M. R. AlRashidi and K. El-Naggar, "Long term electric load forecasting based on particle swarm optimization," *Appl. Energy*, vol. 87, pp. 320–326, Jan. 2010.
[41] E. S. Mostafavi, S. M. Mousavi, and F. Hosseinpour, "Gene expression programming as a basis for new generation of electricity demand prediction models," *Comput. Ind. Eng.*, vol. 74, pp. 120–128, Aug. 2014.
[42] A. Golbraikh and A. Tropsha, "Beware of q^2!," *J. Mol. Graph. Model.*, vol. 20, pp. 269–276, Jan. 2002.
[43] P. P. Roy and K. Roy, "On some aspects of variable selection for partial least squares regression models," *QSAR Combinat. Sci.*, vol. 27, pp. 302–313, Mar. 2008.
[44] A. H. Alavi, P. Aminian, A. H. Gandomi, and M. A. Esmaeili, "Genetic-based modeling of uplift capacity of suction caissons," *Expert Syst. Appl.*, vol. 38, pp. 12608–12618, Sep. 2011.

ALIREZA POURDARYAEI received the bachelor's degree in electrical and electronic engineering from Hormozgan University, Iran, in 2010, the M.Eng.Sc degree in industrial electronic and control from the University of Malaysia (UM), in 2014, where he is currently pursuing the Ph.D. degree in power system. His current research interests include distribution automation, artificial intelligence (AI), and energy management and forecasting.

HAZLIE MOKHLIS (M'01–SM'18) received the B.Eng. and M.Eng.Sc. degrees in electrical engineering from the University of Malaya (UM), Malaysia, in 1999 and 2002, respectively, and the Ph.D. degree from The University of Manchester, Manchester, U.K., in 2009.

He is currently a Professor with the Department of Electrical Engineering, UM and the Head of UM Power and Energy System (UMPES) Research. His research interest includes fault location, distribution automation, power system protection, and renewable energy. He is a Chartered Engineer in U.K. and a Professional Engineer in Malaysia.

HAZLEE AZIL ILLIAS received the bachelor's degree in electrical engineering from the University of Malaya, Malaysia, in 2006, and the Ph.D. degree in electrical engineering from the University of Southampton, U.K., in 2011. Then, he immediately joined Freescale Semiconductors (M) Sdn. Bhd., as a Product Engineer in 2008. Since 2017, he has been an Associate Professor with the Department of Electrical Engineering, University of Malaya. He is an Active Reviewer for numerous high quality journals, which include the IEEE Transactions on Dielectrics and Electrical Insulation, the IEEE Transactions on Power Delivery and Measurement. Since 2011, he has been the Head of the University of Malaya High Voltage Laboratory and Research Group. He is currently a member of Institution of Engineering and Technology (IET). His research interests include modeling and measurement of partial discharge phenomena in solid dielectric insulation, lightning overvoltage, transmission line modeling, optimization methods, and artificial intelligence techniques.

S. HR. AGHAY KABOLI received the B.Eng. degree in electrical engineering and the M.Eng. and Ph.D. degree in power system engineering from the University of Malaya (UM), Kuala Lumpur, Malaysia, in 2009, 2012, and 2018, respectively. From 2011 to 2018, he was a High Impact Research Assistant for Campus Network Smart Grid for Energy Security Project with the UM, in the UM Power Energy Dedicated Advanced Center (UMPEDAC). He is currently with the Electrical Engineering Department, College of Engineering and Petroleum, Kuwait University, Safat, Kuwait. His research interests include artificial intelligence (AI), smart grid, and power electronic.

SHAMEEM AHMAD was born in Chittagong, Bangladesh, in 1985. He received the B.E. degree in electrical and electronics engineering from Visvesvaraya Technological University (VTU), Belgaum, India, in 2009, and the master's degree in electrical engineering from the University of Malaya, Malaysia, in 2014, where he is currently pursuing the Ph.D. degree. His research interests include FACTS controllers, power system stability, load flow analysis, and artificial intelligence.