Recent Development in Electricity Price Forecasting Based on Computational Intelligence Techniques in Deregulated Power Market

Alireza Pourdaryaei [1,2], **Mohammad Mohammadi** [1], **Mazaher Karimi** [3,*], **Hazlie Mokhlis** [4], **Hazlee A. Illias** [4], **Seyed Hamidreza Aghay Kaboli** [5,6] and **Shameem Ahmad** [4]

1   Department of Power and Control, School of Electrical and Computer Engineering, Shiraz University, Shiraz 7194684334, Iran; a.pourdaryaei@gmail.com (A.P.); m.mohammadi@shirazu.ac.ir (M.M.)
2   Department of Electrical and Computer Engineering, University of Hormozgan, Bandar Abbas 7916193145, Iran
3   School of Technology and Innovations, University of Vaasa, Wolffintie 34, 65200 Vaasa, Finland
4   Department of Electrical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur 50603, Malaysia; hazli@um.edu.my (H.M.); h.illias@um.edu.my (H.A.I.); ahmad05shameem@gmail.com (S.A.)
5   Power Systems & Markets Research Group, Electrical Power Engineering Unit, University of Mons, 7000 Mons, Belgium; kaboli0004@gmail.com
6   Electrical Engineering Department, Engineering Faculty, Razi University, Kermanshah 6714414971, Iran
*   Correspondence: mazaher.karimi@uwasa.fi

Abstract: The development of artificial intelligence (AI) based techniques for electricity price forecasting (EPF) provides essential information to electricity market participants and managers because of its greater handling capability of complex input and output relationships. Therefore, this research investigates and analyzes the performance of different optimization methods in the training phase of artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) for the accuracy enhancement of EPF. In this work, a multi-objective optimization-based feature selection technique with the capability of eliminating non-linear and interacting features is implemented to create an efficient day-ahead price forecasting. In the beginning, the multi-objective binary backtracking search algorithm (MOBBSA)-based feature selection technique is used to examine various combinations of input variables to choose the suitable feature subsets, which minimizes, simultaneously, both the number of features and the estimation error. In the later phase, the selected features are transferred into the machine learning-based techniques to map the input variables to the output in order to forecast the electricity price. Furthermore, to increase the forecasting accuracy, a backtracking search algorithm (BSA) is applied as an efficient evolutionary search algorithm in the learning procedure of the ANFIS approach. The performance of the forecasting methods for the Queensland power market in the year 2018, which is well-known as the most competitive market in the world, is investigated and compared to show the superiority of the proposed methods over other selected methods.

Keywords: electricity price forecasting; artificial intelligence; adaptive neuro-fuzzy inference; system feature selection; backtracking search algorithm; competitive market

1. Introduction

Modern power system planning encompasses diverse resources to incorporate the increasing demand subject to numerous techno-economic and environmental constraints. Price forecasting is of paramount importance to all aspects pertaining to power system operation, which is emulated by the abundance of researchers working on operation-related issues [1]. Several methodologies have been put forth by researchers that differ in data processing, model selection, calibration and testing phases. Profuse literature is available on load forecasting due to years of extensive research and application, while the

availability of the literature related to price forecasting is relatively limited. The reason being is that the wholesale competition was either limited or absent because the electricity markets were structured as monopolies. Electricity markets have been evolving recently due to the research in the field of price forecasting gaining attraction, resulting in an evolved competitive market [2]. Countless markets are structured as oligopolies, where the controlling elements are the market's dominant influencing level of prices. The necessity of extensive research on price forecasting approaches is validated by the fact that the market continues to grow more competitive as the importance of price forecasters keeps on increasing [3].

Forecasting of electricity price's role is vital during decision making on investments and the transmission system extension. The method of forecasting should be precise to guarantee that the decision is effective. However, high-precision forecasting is relatively complex, as a number of exclusive features must be considered. These comprise non-linearity, high volatility, multiple seasonality, average reversion and price increase [4]. Furthermore, the non-friendly properties of the price duration line make it even harder to be forecasted [5]. Market clearing price (MCP) is described as volatility, as its hourly determination remains in a shifting and contentious circumstance [6]. A set of multiple parameters influences the evolution of the MCP. Influential parameters are demand, fuel costing (e.g., coal, gas), the capacity of hydropower, the merit order of manufacturing, market players' strategies and grid congestion. In addition, weather patterns and seasonal factors may affect the load level. Contrarily, the price rates of electricity are dependent on a significant and distinctive set of parameters [7]. Some of these criteria may not be accessible to scholars because they are listed and controlled by market competitors. Thus, to achieve a robustness forecasting solution, tests should be conducted at the series of input combinations. The inputs, model parameters, calibration and practical executions must be selected meticulously. This research concentrates on price forecasting of the day-ahead, mainly based on atypical data. Therefore, based on the recently published papers, the study shows that feature extraction from the data set (typical data) given is executed by the improved feature selection algorithm, which generates the maximum impact towards the price forecasting efficiency [8]. Regardless of the utilization of metaheuristic methods to generate major improvement in prediction techniques, a certain number of limitations remain. These include the need for a predefined comprehension of the composition of the current connection within the variables of the fitness function in the algorithm that should be gained priority, high sensitivity of the selected parameter's value from the initial stage and abundance of parameter control. Even though the combination of optimization techniques provides encouraging results, the resolution of metaheuristic methods at a particular point of integration remains a significant issue. Moreover, due to the inherent complexities of a combinatorial optimization approach, there is a continual attempt to tune control parameters in a proper manner. Thus, in the training stage of the artificial intelligence process, various optimization algorithms can be explored to enhance the precision of forecasting [9].

ANFIS, ANN (artificial neural network), SVR (Support Vector Regression) and a combined hybrid of these methods has accomplished a tremendous breakthrough in providing precision in the forecasting of electricity. Nevertheless, additional efficient methods to facilitate electricity price precision are still needed [10]. In addition, for any deregulated power market, the method applied should be able to provide rational forecast prices for a better bidding strategy for generating companies. To ensure it is capable of producing highly reliable results, the accuracy of the electricity price forecasting system needs to be checked. In contemplation of performing this test, the most volatile system must be chosen and validated by means of a detailed statistical analysis. Until now, a rigorous statistical study in the field of electric price forecasting has not yet been performed. Earlier, only a few indices were utilized in electrical price forecasting, namely RMSE, MAE and MAPE. Therefore, the precision of the proposed methods is not completely assessed [11]. Compendious reviews of the latest price prediction approaches and techniques are available

in [12–14]. Meanwhile, stochastic models, artificial intelligence models and models of regression are most of the approaches published through extensive reviews that can be found in [3]. In [15], MLPs are barely implemented to predict, whereas in [16], they are merged with some other time series models. The literature indicates that electricity demand and price prediction can be carried out by utilizing ANN to deal with some approaches. Moreover, ANN is also combined with several data mining methods to identify specific training days [11]. The quality feature of the MLP is that it could be implemented within the hour-ahead timeframe [17], whereas most of the other existing studies are only related to the forecasting of day-ahead. The key role of the MLP merging with the hybrid model intends to enhance the forecast, which is through the classical time series model (i.e., ARIMA) derivations. An investigation on the ANN model in comparison with other models, for instance, AR and ARIMA, is established in [18].

A non-linear mapping of the original data into high dimensional space is provided by SVMs [19], in which a linear function is utilized to determine the boundaries of the new space. Unlike MLPs that have the potential to operate within local minima of their objective function, SVMs can provide a global solution to a problem. During load forecasting, this feature has been applied in [20]. In addition, the estimation on the predictions interval quantifies the uncertainty corresponding to the prediction on targeted numbers with the approximation ranges based on the SVM model [21]. Furthermore, another model for EPF has been proposed by combining the operation of ANNs and Fuzzy Logic (FL), such as adaptive neuro-fuzzy inference system (ANFIS) [22]. However, a predefined rule is the benchmark used by the ANFIS method for forecasting. For the new circumstances, it becomes incapable in the process of learning and self-adaption. By integrating a fuzzy system and an ANN model, this limitation was somewhat tackled in [23]. The main characteristic of a typical neuro-fuzzy model is the capability to adapt, which has made the ANFIS model a widespread predictor to forecast different ranges from the short-term to the long-term price of electricity in any deregulated market. The implementation of the latest AI development for EPF in the literature is depicted in Table 1.

Table 1. Review of the developed artificial smart techniques implemented in literature for the prediction of EPF.

Model	Year	Applied for	Characteristics	MAPE
SVM + GA [9]	2016	Ontario	A real-time cost of Hourly Ontario Electricity Price (HOEP) in Ontario mainland (HOEP and demand are considered as selected features for forecasting analysis.) is predicted hour by hour for 6 test weeks	9.22%
Environmentally adapted generalized neuron model [24]	2017	NSW	The test cycle on the Australian electricity market is a one week test on the basis of historical price and demand data	2.28%
Hybrid ANN-artificial cooperative search algorithm [25]	2019	Ontario	The test cycle on the Ontario electricity market is about 4 months by taking the consideration of different season conditions on the basis of historical price and demand data	1.1%
ANFIS-BSA [26]	2019	Ontario	The test cycle on the Ontario electricity market is about 1 week by taking the consideration of different season conditions on the basis of historical price and demand data for the year 2017	0.79%
Optimized heterogeneous LSTM [27]	2019	PJM	The autocorrelation analysis is conducted to determine price data based on LSTM and EEMD and is used to decompose the electricity price sequence	2.51%
Cuckoo search, singular spectrum analysis and SVM [28]	2019	New South Wales	This work considering the dynamic behavior of price series and to find optimum features by defining one threshold in different seasons	4%
Extreme learning machine [29]	2020	New York City	Real-time data of the market in 2017 in different seasons has been simulated in order to predict the electricity market. Data were converted to four previous hours and were taken into account to forecast the current price	1.44%

Table 1. *Cont.*

Model	Year	Applied for	Characteristics	MAPE
Similar day-based neural network [30]	2020	PJM	The was work conducted based on the availability of the selection of 45 framework days of 2006 of PJM, and three similar days based on the Pearson correlation coefficient model	4.67%
Spatiotemporal deep learning [31]	2020	PJM	The zone prices of all 21 zones in the PJM market and the nodal price of the target location were used as inputs to the CNN model	12.37%
Deep learning hybrid framework [32]	2020	PJM	The point prediction module combines different deep learning techniques, such as DBN, LSTM and CNN, to extract non-linear features	0.06%
Combined SSA, ANN and LSSVM [33]	2020	Australian market	Combined model using 48 data items from the previous Monday to the following Monday	7.58%
Dimension reduction strategy and rough ANN [34]	2020	Ontario and Canada	Grey correlation analysis is applied to select efficient EPF and deep neural networks with stacked denoising auto-encoders to denoise data from different sources	5.64%
Gan-based deep reinforcement learning [35]	2021	New England Electricity market	The proposed approach uses generative adversarial networks to collect synthetic data and increase the training set, and to enhance the forecasting system considering more features, such as temperature and load data, as inputs	Not reported
Hybrid deep neural network [36]	2021	New York City	Hourly electricity price data from 2015 to 2018 are tested using VMD, CNN and GRU for four seasons	0.73%
Multi-head self attention and EEMD framework [37]	2021	New England	LMP, hourly system load with temperature and dew point information, has been used as the input variable to the hybrid model	Not reported
LSTM-NN [38]	2021	PJM	Features are selected by a combination of entropy and mutual information, and wavelet transform is used to eliminate the fluctuation behavior of electricity load and price time series	0.63%
Combined integration via Stacked Pruning Sparse Denoising Auto Encoder [39]	2021	Australian electricity market	The proposed method has been used to decrease the noise of the data set and Tensor Canonical Correlation Analysis to select features with low correlation ranks	5%
Ensemble approach [40]	2021	Austria	A bootstrap aggregated-stack generalized architecture has been implemented to facilitate participants with renewable energy resources in real-time	5.13%

The contributions of this paper are as follows:

- In this work, investigation over the day-ahead price forecasting is based on the price of electricity (PoE) data and demand of electricity (DoE) data in different time intervals. Therefore, within the available data set, this work has developed a feature selection technique to extract the features with the highest impact on price prediction accuracy. The feature selection method is a combination of two techniques, namely ANFIS and MOBBSA. To select non-dominated features, subsets from different combinations of input variable MOBBSA are used, while the performance of every selected feature subset is determined by ANFIS. Additionally, various well-known feature selection methods based on multi-objective methods, such as MOPSO, NSGAIII and NSGAII, are simulated for EPF as a benchmark.
- Since electricity price and demand are inherently correlated, prediction of these parameters in a deregulated market, such as the Australian electricity market, is a challenging task in a smart grid environment. Therefore, to enhance the forecasting accuracy, a novel prediction approach is required, and to meet this requirement in this work, dif-

ferent optimization techniques are implemented to improve the forecasting accuracy in the training phase of an AI-based approach. Based on the studied research in price forecasting, to improve the forecasting accuracy of ANFIS, BSA is applied to tune the membership functions of ANFIS so that the error can be minimized. Based on the forecasting accuracy, the improvement rate of different robust metaheuristic algorithms as comparisons is verified in the training phase of ANFIS and ANN methods.

- Finally, the proposed method is evaluated through the comprehensive statistical analysis. It is observed that the developed model via a data-driven approach complies with all the necessary constraints. As such, it is suitable to be implemented for future EPF.

To evaluate the accuracy of electricity price forecasting, the Queensland electricity market has been considered as one the most unstable markets in the world. In Queensland, the Australian Energy Market Operator (AEMO) provides opportunities for a wide range of forecasting and planning trends to power suppliers and consumers in order to submit their offers for sale and bids to buy the electrical energy. Then, an Independent System Operator (ISO) is a utility or option in the market that arranges the price offers of the generators and the bid price of the consumers. In this open market, a single wholesale market for electricity is well known as a National Electricity Market (NEM). This infrastructure is responsible for purchasing and selling electrical energy between interconnected regions, generating units and also the retailers. The Queensland competitive market is considered as an electric grid that can deliver electricity in a controlled smart way from points of generation to active consumers. It is done by promoting the interaction and responsiveness of the customer as well as offers a broad range of potential benefits for system operation and expansion and for market efficiency. Therefore, it becomes a massive challenge for EPF. Since the pattern of electricity demand is changing based on seasonality; short-term EPF would be more useful for real-time decision making in the deregulated electricity market for the purpose of assessing price forecasting. Hence, to execute this work, feature selection and a forecasting method are adopted to cater to short-term EPF, and data from different seasons of this market are utilized for the verification of AI application for price prediction.

The remainder of the paper is organized as follows: section II summarizes the price forecasting classification. In the same section, the process of the development of the method is briefly explained. In section III, the typical price forecasting approaches, such as ANN and ANFIS, are explained separately. The process of selecting the most influential features for price forecasting is discussed in section V. After investigating the most popular AI techniques, the recent techniques for price forecasting are simulated for the QLD competitive market for different seasons. Finally, the last section concludes the paper and recommends the potential future development of methodologies for accurate electricity price forecasting.

2. General Framework for the Development of Price Forecasting Method

Forecasting of electricity prices is generally divided into short, medium and long term [13] categories. Nonetheless, there is no particular boundary line in the literature to distinguish them. Generally,

- Short-term: This significant subcategory is the most relevant for daily market operations, and the forecast time varies from a few or several minutes to several days upwards.
- Medium-term: The establishment of balancing sheet estimation includes medium electricity price forecasting, such as the derivatives pricing (change structure), and strategies of the risk management process, and the forecast duration starts from a few days to a few months after. The development in the forecast of electricity prices is generally based on the factor of price distribution on the future horizon rather than on the actual point predictions.
- Long-term: The forecasting implementation of this scenario concentrates mostly on the preparation of profitability investment analysis and planning; the prediction duration

is carried out for a month, quartile or even years in the future. This type of forecasting has generated useful information, which is appropriate to evaluate the potential site or generating facility-based fuel sources.

Precise forecasting is a prerequisite for key players and decision-makers in the electricity market to develop an optimal strategy that includes the improvement of socio-economic benefits and risk reduction. Short-term forecasts attract substantial attention and are extensively utilized for economic dispatch and power system control in electricity markets. Therefore, removing impediments in the short-term forecasting of electricity prices will play an instrumental role in managing power systems to meet the growing demand, keeping in line with economic growth that is imperative for sustainable development of different competitive electricity markets.

The proposed electricity price forecasting strategy is presented in Figure 1. After collecting data on historical prices and demand, it is required to prepare constrained data through significant feature selection. Therefore, in the first step, an enhanced feature selection is utilized via hybrid filtering and embedded techniques to assess the quality of features for the forecasting process. In the first stage, MI is applied to reduce the time of training due to the availability of data with a high dimension. In the next stage, MOBBSA is applied to select the features that represent the most important information of the original set. In step 2, the robust forecasting technique, based on combined ANFIS-BSA, is designed for day-ahead price forecasting in the highly volatile Queensland market. In this study, two types of ANFIS are developed. During the feature selection, in order to evaluate the selected input variable for forecasting purposes, as well as during forecasting, the electricity price is used to improve the forecasting accuracy. In the second step, BSA and other well-known optimization techniques are utilized to tune the membership function parameters to improve the price forecasting accuracy.
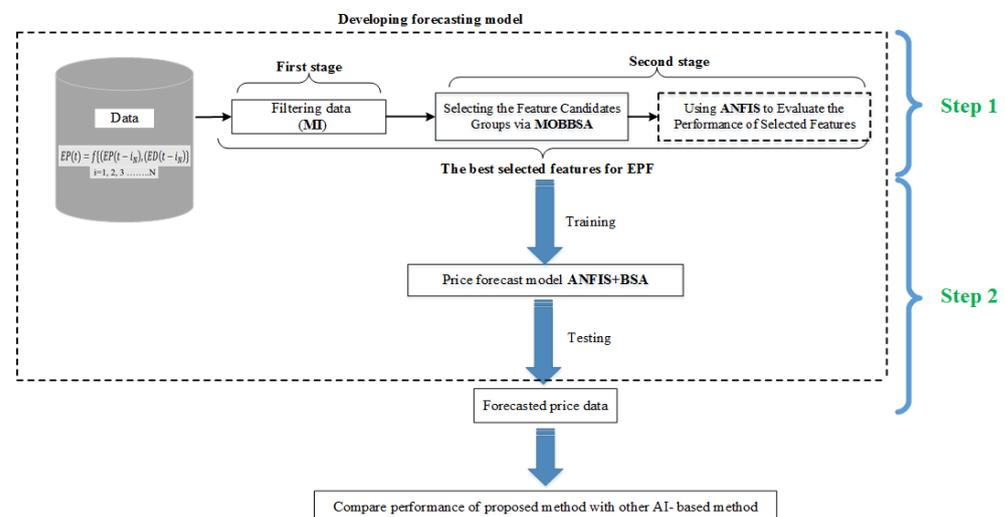
Figure 1. The recent development process for EPF integrated with feature selection.

The electricity price is represented as a function of the demand for electricity in a deregulated electricity market. During the process of electricity price prediction, it can be determined that the electrical energy prices in time, t, depend not only on the demand of electricity, but the previous values also affect the prices. The usual relationship between electricity demand and price from the previous values is stated in the following equation:

$$PoE(t) = f \left(\begin{array}{c} PoE(t-1), PoE(t-2), PoE(t-3), \ldots, PoE(t-NL_{PoE}), \\ DoE(t), DoE(t-1), DoE(t-2), DoE(t-3), \ldots, DoE(t-NL_{DoE}) \end{array} \right) \quad (1)$$

where the demand and price of electricity at time t are represented by $DoE(t)$ and $PoE(t)$, respectively, whereby the assumption has been made for them as a *t* interval time series.

The number of lag sequence for the electricity price is represented by NL_{PoE}, and similarly, the electricity demand lag sequence is represented by NL_{DoE}.

In this research, by considering dual sets of historical input data, the hybrid ANFIS-BSA can be carried out to forecast Queensland electricity price (PoE) hourly. In 2018, historical data sets of input consisting of PoE and hourly Queensland DoE were found in [41]. In order to implement the dependent and independent variables normalization, Equation (2) has been applied based on the wide-ranging historical data. Data normalizations main feature is the adaptation of unrefined data encountered on various scales to a conceptually familiar scale, relatively early to the processing of data by

$$\overline{Z}(t) = \frac{Z(t) - \min(Z)}{\max(Z) - \min(Z)} \tag{2}$$

where Z and t are described as the normalized value, the value that should be normalized and hourly interval, respectively.

Assume that only one week with exogenous variables hourly lagged values ($NL_{EP} = NL_{ED} = 168$) are applied for the prediction of electricity price, which represents 336 exogenous variables lagged. If any of the above exogenous variables are used in the forecasting phase, the learning process may be slowed down, performance deteriorated and training data will cause overfitting. Since these aspects are of utmost significance in the process of electrical energy price prediction, it should pick only those features that have a major impact on performance.

The primary goal selection of the featured method is purposely to determine the importance of input features with the consistency aspect for the best subset selection in the original feature set of the suitable information. Due to the principle of dimensionality, numerous predictors may lead to a lower performance of extracted models. The core principle for using a feature selection technique has to be the removal of redundant or irrelevant features from a data set with various features, with no significant loss of predictive precision. A search strategy with a measurement metric is used for seeking candidates and for rating the performance of these candidates in a feature selection algorithm. The simplest function selection algorithm is to evaluate every potential feature subset to find one that reduces the error rate that provides the feature space with a thorough search, but it is computationally intractable [42]. Therefore, to discover all potentials, integration is implemented by a comprehensive assessment measure. Meanwhile, to determine the feature quality with a significant impact on the effective algorithm for the feature selection integration, a search strategy is utilized.

Based on the combination of searching techniques with well-known learning algorithms (assessment metric) to construct a model, three classes of feature selection methods are formed, namely wrappers, filters and embedded devices [43]. A predictive model based on wrapper (the search driven by accuracy) methods is implemented for feature subset evaluation. Every candidate function subset is applied to wrapper methods to train a model that is assessed on a holdout set. The error rate evaluation of the model for a testing set has generated the score for each subset of participants. As the wrapper for each candidate subset technique provides training to a trendy predictive model intended for the expenses of computationally intensive tasks for that particular type of model, they will regularly offer the best-performing feature set.

In filter (information gain) methods, a proxy measure is used instead of an error rate to score a candidate subset of the feature. The selection of proxy measures comprises the pointwise information, the product-moment correlation coefficient and mutual information for faster assessment of efficiency feature set. The MI technique is already extensively used in electricity market price prediction [44]. Nevertheless, this technique encounters challenges because of the lagged values of the candidate inputs given by the electricity market, which consist of load demand, price and other variables. Therefore, it is hard to acquire both individual and joint probability distributions of the candidate inputs [11]. In addition, it should be pointed out that the price of electricity is also recognized as a

time-variant signal. Hence, it is not necessary to use a long history of candidate inputs, as market circumstances fluctuate most of the time. As such, due to the shortage of information values, it may deceive or give a less accurate price forecast process [11].

The wrapper methods are much more computationally intensive compared to filter methods, but wrapper methods have used a specific type of learning algorithm to obtain a subset of features for performance evaluation. Due to the learning algorithm lacking in filters, this has caused less prediction performance generated in comparison with the wrapper method set performed by the filter methods with a feature set, which is normally a common method. Commonly, the filter type method is utilized to discover variables' relationships, and instead of including an explicit best feature subset, the rank of features is more preferred. Hence, the hybrid feature selection method can be created by utilizing a filter to do a wrapper pre-processing step. A selection in this hybrid method for the most suitable features of bigger sets of data, based on the dimensionality reduction method by filter method type, allows the wrapper to do the proper selection.

Another subset of features, known as embedded methods, has the best contribution to precision during the process of designing the model. Embedded methods must not distinguish between the feature selection component and the training process, as the selection of the features and model construction steps are accomplished concurrently. Although computation is less comprehensive using the built-in methods compared to aggregation methods, this method has major limitations with certain features towards the base model, being sensitive to its structure. Hence, the approach is normally accurate to their learning algorithms. Dissimilar categories of the embedded method are classification trees, known as random forests and regularization approaches. The most generic version of the embedded feature selection method is the regularization approach, which is often called the penalization approach. The penalization approach inserts more constraints to the model development, which simplifies the model by penalizing the model for higher intricacy.

3. Electricity Price Forecasting (EPF) Techniques

In assessing electricity price forecasting, there are two general techniques; hard and soft computing techniques. There are various studies on hard computing approaches with various objectives, such as transfer function model, autoregressive integrated moving average (ARIMA), wavelet-ARIMA and mixed model. This approach needs an accurate model of the system to utilize the algorithm in finding the optimal solution considering physical phenomena. Although the accuracy of this approach is found to be high, it needs a large number of information and is computationally outrageous. There are also several studies on soft computing approaches in electricity price forecasting. Some recent approaches include artificial neural networks (ANN) and adaptive neuro-fuzzy inference system (ANFIS). Generally, this approach does not need any system modeling since it develops an input-output mapping based on historical data. Therefore, these approaches are computationally more efficient, and it has higher accuracy as well as higher resolution subject to correct inputs [13]. Therefore, the main focus of this paper is to review the methods and techniques that have been developed and introduced adopting soft computing models, namely AI techniques.

3.1. Artificial Neural Network

Artificial neural networks (ANN) are one of the promising technologies found in the last few decades that are used extensively in numerous functions in various fields. In the 1980s, the ANN approach, basically a mathematical model, was introduced for the very first time. An artificial neural network, simply known as a neural network, is developed based on the architecture and activity of biological neural networks in the brain. Numerous numbers of artificial neurons are gathered, and together, they construct an artificial neural network. Each neuron is connected to other neurons through synaptic weights (or directly weights). A simple biological neuron has four main parts—a cell body (soma), axons, dendrites and synapses. The dendrites help to take input signals

into the cell body. Axons' responsibility is to transfer the signals from one neuron to the others. On the other hand, the dendrites of one cell and the axon of another cell meet at a point called a synapse. An artificial neural network consists mainly of; weights, bias and activation functions. Generally, an artificial neural network can be divided into two main parts; neurons and connections between network layers where the neurons are located. A typical ANN consists of three main layers, such as input layer, hidden layer and output layer. The ANN uses the concept of multilayer perception (MLP) that is the most popular ANN method among researchers. However, the outputs (Y_n) of ANN are determined from Equation (3) as follows:

$$Y_n = f_n(\sum_{i=1}^{m} W_{ni}X_n + b_n) \tag{3}$$

where X_n presents the input values, W_{ni} stands for the connection weight values among the input, hidden and output layers. b_n and f_n are the bias and transfer functions, respectively. From the above model and equation of ANN, the main challenge is found in handling the unknown variable-transfer function. The responsibility of the transfer function is to determine the characteristics of an artificial neuron.

3.2. Adaptive Neuro-Fuzzy Inference System (ANFIS)

Based on the early defined rules, fuzzy logic methods cannot learn as well as adapt to a new condition by themselves. To succeed, the authors of [45] have mixed two different methods with each other and made a hybrid method named ANFIS, which is nothing but a mixture of Fuzzy Inference System (FIS) and ANN. The methodology named ANFIS can be acknowledged as a characterized system that can be matched with ANN. In ANN, the output parameters of the fuzzy system can be adaptive to train the system parameters of the fuzzy membership function. The benefit of FIS and ANN is processed by ANFIS. Different types of drawbacks are being analyzed through different difficult procedures of neural networks where all of the networks are being bypassed by linguistic variables of the FIS system. On the other hand, the neural inference system solves the problem by creating the ability to learn, as well as adapting themselves to a new condition. Hence, complicated non-linear mappings can be assumed by the competency of this method by applying the fuzzy system with ANN learning. Furthermore, it is acknowledged as a comprehensive estimator of long-lasting, medium and short forecasting [46].

The main reason to develop a system such as ANFIS is to adopt a system with a tunable membership function (MF), as well as a set of fuzzy rules during a phase of training. Two individual parameters can be optimized to implement the learning steps:

- Parameters of antecedent (the MF parameters)
- Parameters of consequent (the fuzzy system output function)

Here, the characteristic is linear in the following parameters. The linear least-square is applied to optimize the predecessor parameters that look very similar to neural networks' backpropagation algorithm in conjunction, where gradient descent is applied for optimization.

Usually, the ANFIS is constructed by five individual stages. Among those, each of the stages has a node function. From the earlier layers, the next layer gets an input node [47]. The sequential layers of ANFIS can be arranged as fuzzification (if-part) in layer 1, production part in layer 2, normalization part in layer 3, defuzzification (then-part) in layer 4 and lastly, total output generation part in layer 5. There are dual inputs that are independent variables (x and y) and a single output, which is a dependent variable (f_{out}) included in the composition of ANFIS.

Dual diverse kinds of fuzzy inference systems are generated by alternating the fuzzy rules (if-then) of the consequences set with the procedure of defuzzification. This system is called Mamdani-based FIS and Sugeno-based FIS.

In numerous regards, the Mamdani-based FIS approach is similar to the Sugeno approach. A comparative fluffy deduction is prepared for both sorts by the implementation

of the fuzzifying process upon the input information and the fluffy administrators. The foremost distinction between Sugeno-based FIS and Mamdani-based FIS is that the manner of the fluffy inputs has changed over to a fresh yield. In Mamdani-based FIS, for computing the fresh yield, the fluffy yield is employed in the defuzzification strategy, whereas in Sugeno-based FIS, the weighted normal strategy is utilized. The idea of disposing of Mamdani interpretability and expressive control is the aim of the strategies due to the reason that the standard consequence of the Sugeno strategy is not fluffy. Sugeno has a quicker interim time compared to Mamdani-based FIS; rather than the time-devouring defuzzification, it prepares the connection to the weighted normal strategy. Due to the instinctive nature of and operation of this view, it has led to the complex strategy of Mamdani, with the choice-back application thought to be linked. Additionally, Sugeno and Mamdani-based FIS's show more contrast between them due to the fact that Sugeno has no yield participation capacities compared to Mamdani FIS yield participation, so the Sugeno strategy gives a yield that is either a direct (weighted) numerical expression or is steady. The Mamdani strategy provides a yield that is a fluffy set. Sugeno has more adaptability in the framework plan than Mamdani-based FIS, as demonstrated by the more efficient frameworks that can be achieved if the ANFIS device is coordinated with [48].

Conceding ANFIS is linked with Sugeno-based FIS, the composition of the fuzzy IF-THEN rules of Sugeno-based FIS of the first-order are known as the ANFIS rules, and are indicated as:

Rule 1: *If x is A1 and y is B1 then z is f1(x, y; p1, q1, r1) = x p1 + y q1 + r1*
Rule 2: *If x is A2 and y is B2 then z is f2(x, y; p2, q2, r2) = x p2 + y q2 + r2*

where Ai and Bi are sets of the fuzzy, fi (x, y; pi, qi, ri) is known as the first-order polynomial function that defines the Sugeno-based FIS outputs, x and y are two separate facts, and z is an ANFIS model output.

Inside the ANFIS mainframe, the layers of the distinctive comprise of distinctive hubs work. The nodes (hubs) within the same layer of this network perform functions of the same type. The layers are described more in detail as follows:

Layer 1: In this layer, the inputs are x and y to hub i, etymological names are Ai and Bi, enrollment capacities for Ai and Bi fluffy sets are μ_{Ai} and μ_{Bi}, separately, and the enrollment review of a fluffy set is known as $q1$. i is regarded as the yield of hub i within a layer that indicates the degree to which the specified input (x or y) fulfills the evaluation. In ANFIS, the MF (enrollment work) for a fluffy set can regularly be any parameterized participation work, such as universal Chime molded work, Gaussian, trapezoidal or triangular.

Layer 2: Each hub in this layer may be a settled hub that yields the item of all the approaching signals. In this layer, through an increase of input signals, the terminating quality of each run is decided.

Layer 3: Each hub in this layer may be a settled hub. Throughout this layer, the terminating quality given in the past layer is normalized by computing the proportion of the ith rule's terminating quality to the entirety of all rules' terminating qualities.

Layer 4: Each hub is adaptive with a node feature in this layer.

Layer 5: This layer has one settled hub that computes the large yield of ANFIS by summation of all approaching signals.

Lastly, the hybrid learning algorithm has been utilized by ANFIS to tune the parameter. At the same time, for updating the input MF parameters (antecedent parameters) and training the consequent parameters in layer 1, respectively, the backpropagation algorithm and the least-squared method have been used.

Based on recent research in AI techniques, it was concluded that the simulation results of the Queensland market in 2018 in different seasons are based on the hybrid ANFIS-BSA. Therefore, BSA is explained as the most recently developed optimization technique in the training phase of ANFIS and is compared with well-known optimization techniques to prove the proposed method can be applied for any deregulated electricity market.

4. Multi-Objective Backtracking Search Algorithm (MOBSA)

In multi-objective optimization problems, the Pareto (French economist) optimality method is applied to generate a set of solutions to the objectives instead of looking for a single solution. The set of optimal solutions is needed because a single point may not optimize all the objective functions at the same time due to a conflict sometimes arising among the objectives. There are two feasible optimal solutions determined by the Petro optimality technique, which are designated by $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_N)$ and $\partial = (\partial_1, \ldots, \partial_N)$, respectively. In order to accomplish the solutions, two sets of objective functions are used; $f(\varepsilon) = (f_1(\varepsilon), \ldots, f_m(\varepsilon))$ and $f(\partial) = (f_1(\partial), \ldots, f_m(\partial))$, developed as shown in Equation (4). Solution (∂) is accepted only as an optimal solution over solution (ε) when a mathematical condition ($f(\varepsilon) < f(\partial)$) and Equation (4) are satisfied simultaneously. Hence, solution (ε) is called the non-dominant solution, corresponding to the solutions incorporated in the Pareto optimal set. The depiction of the Pareto optimal set (containing the objecting functions and decision variables) is designated as Pareto front [49–56].

$$\forall i \in \{1, \ldots, m\} : f_i(\varepsilon) \leq f_i(\partial)$$
$$\exists j \in \{1, \ldots, m\} : f_j(\varepsilon) < f_j(\partial) \tag{4}$$

The Pareto optimal set of a multi-objective BSA scheme is shown in Figure 2, in which the first and second objective functions are named f_1 and f_2, respectively. The figure consists of the gray colored dominated and red circled non-dominated solutions of the Pareto optimality. The two samples (A and B) are collected from the non-dominated solutions, as represented in Figure 2.
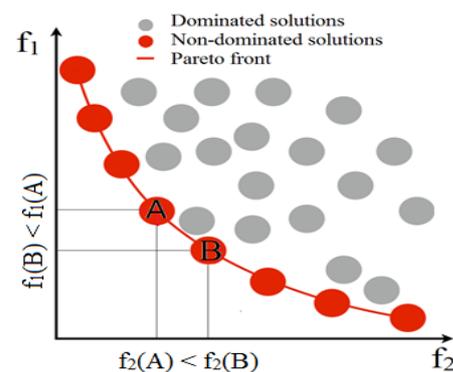
Figure 2. Dual fitness functions based on the optimal set of the Pareto front method (two examples of non-dominated solutions; A and B).

The Pareto front produces a group of optimal outcomes regardless of generating a unique optimal solution in the BSA optimization analysis. In fact, no single solution should be neglected over other solutions in the Pareto front as they are all considered crucial parts of the optimization technique. It may be an impossible task to attain greater development in the determined objectives if any one of them is eliminated from the optimization process. Consequently, trade-offs of the solutions are expected and satisfied with a most convincing solution while manipulating the multi-objective optimization problems.

A standardization mechanism is designed to provide an estimated common scale in the objectives that are originally designed for distinct scale factors in the multi-objective algorithm. To evaluate a precise solution (least valued solution) from the Pareto set, other normalized outcomes have to be accumulated and measured with the common scale.

The backtracking search algorithm (BSA) is the latest evolutionary algorithm with a simple structure. It has the capability to solve multimodal functions and different numerical optimization problems. In BSA, two advanced crossover and mutation operators are proposed to generate the trial population. These operators are unique and different from other evolutionary algorithms (e.g., GA and DE) in terms of their structure. It only has one control parameter, and it is insensitive to the initial parameter value. As such, it

overcomes the drawbacks of metaheuristic methods that have a lot of control parameters, are sensitive to the initial value of these parameters, premature convergence and time-consuming computation. The overall optimization process of BSA in the selection of features and price prediction is obtainable, which has six stages: initialization, selection-I, mutation, crossover, boundary control and selection-II [57].

In the case of multi-objective BSA schemes, several numbers of optimal solutions are generated as a Pareto optimal set, rather than dealing with a unique solution. To compute well the generated dominant and non-dominant solutions and to bring improvements in the existing algorithm, a sophisticated mechanism is included based on the superiority idea of the Pareto technique [58]. In an initial stage, there should be a generation of a considerable number of offspring (T) by using the parameters of the crossover and mutation processes in the multi-objective BSA algorithm. Hence, a comparative study is taken place depending upon the notion of Pareto dominance between the individual members (i^{th}) of the offspring and the population (P_i). In this comparison analysis, the individuals of the population are replaced by the offspring ones due to (T_i) led by (P_i) in the optimization process. In the next stage, it is important to transform the BSA algorithm into the multi-objective functioned method to reach the global minimum optimization. A Pareto optimal mechanism set is developed to store many dominant and non-dominant solutions as an alternative to exporting a global minimum approach. To establish the concepts of the external elitist archive and crowding procedure, there are many steps that need to be accomplished in the multi-functioned BSA algorithm, which are shown in Figure 3 and described in the following order.

Step 1: The optimization parameters are the same size as the main arbitrarily developed population (P), as referred to in (5).

Step 2: Now, there is a need to figure out every member of the primary population with their fitness strengths. After identifying the categories of different populations, only non-dominant solutions should be stored in the external archive.

Step 3: Equation (6) is used to compute the historical archive of the parent participants in this BSA optimization algorithm.

Step 4: The archived members of the non-dominant solutions are updated in every consecutive repetition of the optimization procedure by following the "if-then" rules in Equation (7).

Step 5: A mutation technique is implemented to determine a single offspring from only one population stored in the historical archive, which is manipulated by Equation (9).

Step 6: To Equation (10), eventually, a unique solution of the offspring (T) can be achieved using a crossover strategy in every consecutive iterative operation from the trial population previously stored in the archive.

Step 7: After the crossover analysis is completed, an already produced member of the offspring population should be replaced by an alternative one if it breaches the threshold condition (stated in Equation (11)) of the non-dominant set size in the external elitist archive.

Step 8: The entries (i^{th}) of the produced offspring (T_i) have to be changed by the members (i^{th}) of the parent (P_i), when T_i exceeds Pi in number.

Step 9: Then, reorder the solutions in the elitist archive based on the commands explained in Step 4.

Step 10: The area of objective functions is split after measuring the crowding inter-spaces of the solutions in the external elitist archive. Then every single solution is stationed in a specific destination based on the parameters of their objectives. When there is no more space to store a newcomer of the non-dominant solutions into the external archive, an arbitrarily chosen solution from the densely populated area should be eliminated to give access to the new incoming solution.

Step 11: If the redistribution process does not meet the optimization requirements, apply the formula g = g + 1 and start repeating the optimization process from Step 4.

$$P_{i,j;g=0} \sim U(low_j, up_j) \quad , y_i = f(P_i)$$
$$for$$
$$i = \{1,2,3,...,nPop\}, \quad j = \{1,2,3,...,nVar\} \tag{5}$$

where:

nPop is population size. nVar signifies the optimization variable. uniform distribution function is defined as U. low$_j$ and up$_j$ are lower and upper

$$oldP_{i,j} \sim U(low_j, up_j) \tag{6}$$

where,

$$if \ a < b \Big|_{a,b \sim U(0,1)} \ then \quad oldP := P \tag{7}$$

where:

:= is the updated operation. a and b are randomly generated numbers.

$$oldP := permuting(oldP) \tag{8}$$

where:

$$Mutant = P + F.(oldP - P)$$
$$F = 3.rndn \Big|_{rndn \sim N(0,1)} \tag{9}$$

$$map_{i,j} = 1,$$

$$if \ a < b \Big|_{a,b \sim U(0,1)} \quad then$$

$$map_{i,u_{(1:\lceil mixrate.rnd.nVar \rceil)}} = 0 \Big|_{rnd \sim U(0,1), \ u = permuting(1,2,3,...,nVar)}$$
$$else$$
$$map_{i,randi(nVar)} = 0$$
$$T := Mutant$$
$$if \ map_{i,j} = 1 \ then \ T_{i,j} := P_{i,j}$$
$$for$$
$$i = \{1,2,3,...,nPop\}, \quad j = \{1,2,3,...,nVar\} \tag{10}$$

where, mixrate is the control parameter of optimization algorithm.

$$if \ (T_{i,j} < low_j) \ or \ (T_{i,j} > up_j) \ then \ T_{i,j} \sim U(low_j, up_j) \tag{11}$$

where,

$$if \ f(T_i) < y_i \ then \ , \ y_i := f(T_i), \ P_i := T_i$$
$$y_g = min(f(P_g))$$
$$if \ y_g < y_{g-1} \ then \ global \ minimum := y_g \ , \ global \ minimizer := P_g \ ,$$
$$g = g+1$$
$$for$$
$$i = \{1,2,3,...,nPop\}, \ g = \{1,2,3,...,gMax\} \tag{12}$$

Step 1 Initialization

Step 2 Selection-I

Step 3 Mutation

Step 4 Crossover

Step 5 Boundary Control

Step 6 Selection-II

Check maximum number of Iteration reached for a predetermined value
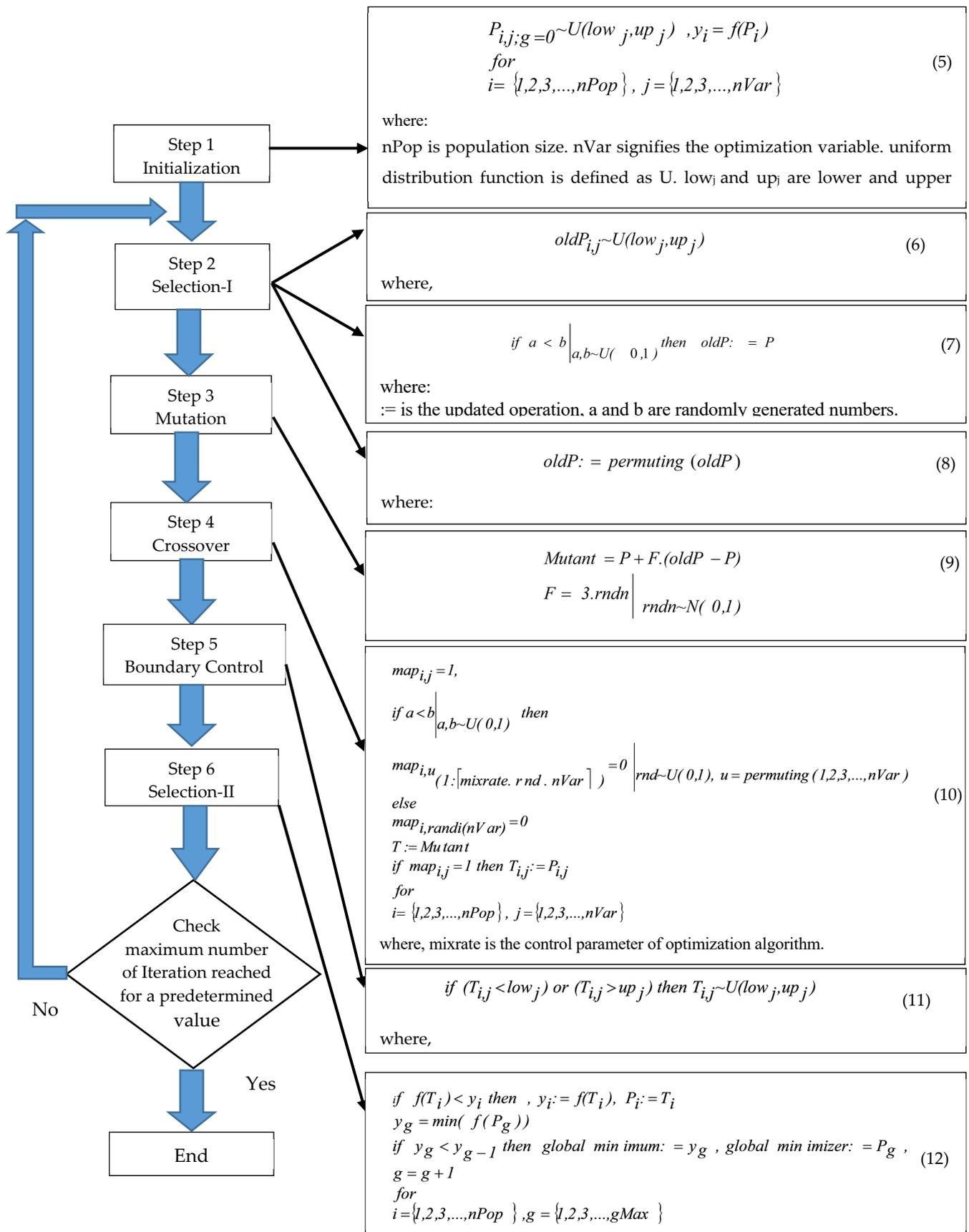
No

Yes

End

Figure 3. Steps need to be accomplished in the multi-functioned BSA algorithm.

5. Selection of Forecast Model Inputs via Multi-Objectives

In this section, feature selection is developed based on the filtering and wrapper techniques to determine the most effective input for the forecasting of the Queensland deregulated electricity market. The wrapper method for a particular model type always provides the most important features, but their efficient research algorithm in the evolving training phase is required. The sequential search-based metaheuristic technique for feature selection is recommended for the purpose of adding or eliminating the features before the efficiency of the model is improved, as the exhaustive search is normally impractical. During the local optima point, the sequential search has a probability of turning out to be stagnated. Thereby, to search for a different feature subsets space, the metaheuristic algorithm, also known as the randomized search algorithm, is suggested. The application of metaheuristic algorithms for feature selection has included random searching procedures to avoid becoming trapped in a local optima point. The multiple metaheuristic algorithms, which are GA, SA, ACO and PSO, have been used in the context of feature selection [42].

The formulation of the objective optimization problem of the feature selection is done in metaheuristic algorithms; therefore, the number of appropriate features must be predefined and continually locate the features subset along with a number of static features. Typically, the selection of features has dual main-diverging purposes; simultaneously minimizing both measurement error and the number of features. Consequently, multi-objective problems formulated from the selection of features consist of two key goals, which are optimizing the model effectiveness and considering the number of features to be minimized, and a decision is a trade-off of these two objectives. The multi-objective topic is a formulation of feature selection that corresponds to a non-dominant set of feature subsets in order to fulfill multiple requirements.

The investigation through NSGA-III, NSGA-II and multi-objective particle swarm optimization (MOPSO) is carried out to attain the Pareto front of feature subsets [58,59], yet a still more effective search approach is required to enhance the solution to feature selection issues [60]. There are some issues experienced by the existing multi-objective feature selection algorithms where the obstacles are referring to the computational cost, which is too high, more parameter control, and are also very sensitive to the initial value of the BSA parameter, whereby it is considered to be less costly to computationally implement instead of other metaheuristic techniques with just a single control parameter [54]. A Binary-Valued BSA (BBSA) in [61] is proposed for solving the discrete form parameter optimization.

Nevertheless, BBSA is being used as an effective search algorithm for the selection of features in [60]; it handles the role of a single objective issue and is not specifically applied to tackle multi-objective problems of feature selection. Numerous variations of Multi-Objective BSA (MOBSA) were established in [54,62]. Statistical analyses in [54] indicate that MOBSA is becoming a promising strategy of optimization to solve high-dimensional multi-objective problems among multiple established multi-objective evolutionary algorithms (e.g., MOPSO, NSGAIII and NSGA-II). Hence, this work proposes an algorithm for solving multi-objective problems of feature selection based on BBSA that can be a potential algorithm for obtaining a non-dominant subset Pareto front. The evolutionary training procedure can utilize any learning algorithm (e.g., ANN, SVM, ANFIS) for solving BSA-based multi-objective problem algorithms to assess the quality of individual candidate feature subsets. Due to its rapid learning ability to estimate non-linear functions, ANFIS has been known as a universal estimator [60]; thus, it is incorporated in the proposed wrapper-based multi-objective feature selection method as an evaluation metric method. In particular, ANFIS deploys an effective hybrid learning technique, which integrates the least-square method with the gradient descent. The least-square method contributed to the speed of the training [63]. Thus, ANFIS has the capabilities to develop the predictive model after only a few epochs of training. The models have been designed for different combinations of features selected by Multi-Objective BBSA (MOBBSA), as the least-square approach is computationally effective, a single/few runs of the least-square technique is used to train them, and then to establish the model, a subset with non-dominated features

with the best performance quality is preferable. Prior to implementing a feature selection to gain the most prominent subsets with maximum validity variables of input and minimal of EPF short-term redundancy, dependent variables, as well as the independent variables, are put into two sets randomly: 70% as training sets and 30% as test sets. ANFIS models are constructed by the training set with various input variables subsets, whilst the test set provides access to the vigor and utility generated models.

MOBBSA is utilized to appear within a diverse mix of variable inputs as the technique of multi-objective feature selection progresses and the selection of the non-dominant subset of features advances, while ANFIS is used as a metric of evaluation to observe the work accomplishment of each subset of each feature. Each MOBBSA individual represents one input variable during the applied learning method (ANFIS) training process. The establishment of a feature selection method retains non-dominant feature subsets in an external elitist archive, which concurrently minimizes the root mean square error (RMSE) by using the principle of Pareto dominance on the test set and a variable number of the input to achieve the solution in global optima points. The MOBBSA upgrade from BSA currently only has a single control parameter, called "mixrate", responsible for restricting the number of individuals' involved in the crossover phase. Therefore, about 100% of the population has been determined with a maximum mixrate value (i.e., 100% of the population size) has been considered for the implementation of the feature selection approach by including each and every individual in the crossover stage. The alternative potential to the Mamdani-based FIS is the Sugeno-based FIS to build the framework of ANFIS for the purpose of feature selection since this type of model is ideally suitable for modeling non-linear functions by interpolating with numerous linear functions. Based on the research done [64], the author provided scatter partitioning into the training process to enhance the feature selection process. The main part of scatter partitioning, which is known as subtractive clustering, becomes a key point to create the ANFIS for feature-based selection.

The reciprocal information of the input characteristics is evaluated, and insignificant and redundant traits are filtered to create a lower input subset. Then the technique of feature selection is implemented into the reduced subset to identify a smaller set of features with high predictive precision. The feature selection technique has suggested that during the initial stage, the reciprocal details of the input and output features amongst every individual variable are tabulated based on MI formulation in [25]. The indication of mutual information with much higher values leads to a higher reliance on each output and input variable. The sorting process of input features is in descending order, which is based on specifically computing the mutual information.

In order to reduce the running time of feature selection, two-stage feature selection is proposed in this work. Input features exhibiting a lesser amount regarding the significant influence toward the output and allowing lesser value due to the reciprocal information than the relevancy threshold (TH) is eliminated. Due to filtering the purposes of the redundant features, the significance threshold must be considered as TH = 0.46. Therefore, after filtering procedures, the most important qualities of 69 features are selected. Out of the 69 candidates chosen, 27 features with the most significance and the highest dissimilarity are identified by MOBBSA as inputs for the next predictive procedures. Moreover, for the effectiveness of the evaluation phase in the proposed multi-objective feature selection method, the comparison with the MOPSO, NSGAIII and NSGAII has been carried out. A thorough analysis to find the optimal solution for subsets of input variables along with their computational time is achieved using the multi-objective feature selection methods and tabulated RMSE values, which represent their subsequent performances in Table 2. According to the obtained results, for the similar results to the test, the suggested feature selection technique is better than other techniques since it generates less error in estimation and feature numbers.

Table 2. The multi-objective feature selection methods select the optimal subsets of input variables by studied, RMSE and computational time value, which represent their subsequent performances.

MOBBSA + ANFIS	MOPSO + ANFIS	NSGAIII + ANFIS	NSGAII + ANFIS	MOBBSA + ANN	NSGAIII + ANN	MOPSO + ANN	NSGAII + ANN
PoE(t-1)	PoE(t-1)	PoE(t-1)	PoE(t-1)	PoE(t-1)	PoE(t-1)	PoE(t-1)	PoE(t-1)
PoE(t-2)	PoE(t-2)	PoE(t-2)	PoE(t-2)	PoE(t-2)	PoE(t-2)	PoE(t-2)	PoE(t-2)
PoE(t-3)	PoE(t-3)	PoE(t-3)	PoE(t-3)	PoE(t-3)	PoE(t-4)	PoE(t-3)	PoE(t-6)
PoE(t-23)	PoE(t-23)	PoE(t-23)	PoE(t-23)	PoE(t-23)	PoE(t-23)	PoE(t-71)	PoE(t-23)
PoE(t-24)	PoE(t-25)	PoE(t-47)	PoE(t-24)	PoE(t-24)	PoE(t-47)	PoE(t-73)	PoE(t-24)
PoE(t-25)	PoE(t-47)	PoE(t-48)	PoE(t-48)	PoE(t-25)	PoE(t-96)	PoE(t-96)	PoE(t-48)
PoE(t-47)	PoE(t-48)	PoE(t-49)	PoE(t-49)	PoE(t-48)	PoE(t-120)	PoE(t-97)	PoE(t-120)
PoE(t-48)	PoE(t-49)	PoE(t-71)	PoE(t-72)	PoE(t-49)	PoE(t-121)	PoE(t-120)	PoE(t-144)
PoE(t-72)	PoE(t-71)	PoE(t-72)	PoE(t-73)	PoE(t-73)	PoE(t-144)	PoE(t-121)	PoE(t-168)
PoE(t-95)	PoE(t-72)	PoE(t-96)	PoE(t-96)	PoE(t-94)	PoE(t-167)	PoE(t-144)	PoE(t-169)
PoE(t-120)	PoE(t-96)	PoE(t-97)	PoE(t-97)	PoE(t-120)	PoE(t-168)	PoE(t-145)	PoE(t-192)
PoE(t-167)	PoE (t-119)	PoE(t-120)	PoE(t-120)	PoE(t-167)	PoE(t-169)	PoE(t-167)	PoE(t-193)
PoE(t-168)	PoE(t-168)	PoE(t-123)	PoE(t-121)	PoE(t-168)	PoE(t-192)	PoE(t-168)	PoE(t-335)
PoE(t-169)	PoE(t-169)	PoE(t-145)	PoE(t-144)	PoE(t-169)	PoE(t-336)	PoE(t-169)	PoE(t-337)
PoE(t-191)	PoE(t-192)	PoE(t-168)	PoE(t-145)	PoE(t-191)	PoE(t-337)	PoE(t-335)	PoE(t-503)
PoE(t-336)	PoE(t-334)	PoE(t-192)	PoE(t-168)	PoE(t-192)	PoE(t-504)	PoE(t-336)	PoE(t-504)
PoE(t-504)	PoE(t-335)	PoE(t-334)	PoE(t-169)	PoE(t-336)	PoE(t-505)	PoE(t-337)	DoE(t)
DoE(t)	PoE(t-336)	PoE(t-335)	PoE(t-192)	PoE(t-504)	DoE(t)	PoE(t-504)	DoE(t-1)
DoE(t-1)	DoE(t)	DoE(t)	PoE(t-193)	DoE(t)	DoE(t-1)	PoE(t-505)	DoE(t-4)
DoE(t-2)	DoE(t-1)	DoE(t-1)	DoE(t)	DoE(t-1)	DoE(t-3)	DoE(t)	DoE(t-12)
DoE(t-23)	DoE(t-2)	DoE(t-2)	DoE(t-1)	DoE(t-2)	DoE(t-24)	DoE(t-1)	DoE(t-24)
DoE(t-24)	DoE(t-24)	DoE(t-24)	DoE(t-2)	DoE(t-24)	DoE(t-25)	DoE(t-3)	DoE(t-25)
DoE(t-25)	DoE(t-25)	DoE(t-25)	DoE(t-24)	DoE(t-25)	DoE(t-71)	DoE(t-24)	DoE(t-48)
DoE(t-167)	DoE(t-167)	DoE(t-72)	DoE(t-25)	DoE(t-72)	DoE(t-72)	DoE(t-72)	DoE(t-72)
DoE(t-168)	DoE(t-168)	DoE(t-96)	DoE(t-72)	DoE(t-96)	DoE(t-95)	DoE(t-96)	DoE(t-96)
DoE(t-169)	DoE(t-169)	DoE(t-120)	DoE(t-96)	DoE(t-120)	DoE(t-96)	DoE(t-120)	DoE(t-120)
DoE(t-335)	DoE(t-335)	DoE(t-169)	DoE(t-120)	DoE(t-144)	DoE(t-144)	DoE(t-144)	DoE(t-144)
	DoE(t-336)	DoE(t-335)	DoE(t-144)	DoE(t-168)	DoE(t-168)	DoE(t-168)	DoE(t-168)
			DoE(t-168)	DoE(t-336)	DoE(t-192)	DoE(t-192)	DoE(t-335)
			DoE(t-336)			DoE(t-335)	DoE(t-504)
RMSE							
17.35	17.59	17.61	17.96	18.70	18.75	18.87	18.94
Computational Time							
98.6523	112.4573	110.6785	119.3236	128.0335	142.4235	140.3657	157.2518

6. Sequential Steps to Obtain AI-Based Models for Short-Term EPF

Figure 4 represents the sequential steps to obtain AI-based models for short-term EPF are carried out for all models, which are as follows:
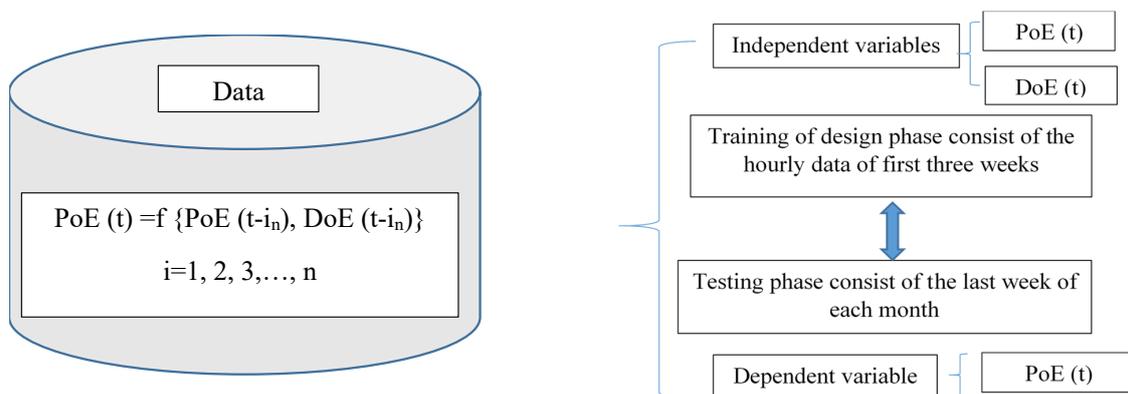

Figure 4. Process of obtain AI-Based Models for Short-Term EPF.

Step 1: To evaluate the efficacy of the applied methods for EPF in different seasons, one month in each season was considered (i.e., February (summer), May (autumn), August (winter) and November (spring)) due to seasonality effects.

Step 2: Designing the training phase entails the derivation of algorithms responsible for connecting the input variables to the output variables, and Equation (2) is used to normalize the input and output variables to make the learning process swift.

Once the selected features (PoE and DoE) are finalized, the most influential features will transfer as the input variables for the next forecasting process. In order to evaluate the forecasted price, a data set beyond the training data set, known as the testing data set, is employed. The reliability of the generated model can be determined utilizing the considered scores in the testing process.

Step 3: To predict the electricity price (PoE) precisely, ANFIS-BSA is implemented by minimizing the cost function, illustrated as follows:

$$F = \sum_{t=1}^{N} \left| PoE(t)_{observed} - PoE(t)_{forecasted} \right| \tag{13}$$

where the actual and predicted electricity prices are $PoE\ (t)_{observed}$ and $PoE(t)_{forecasted}$, respectively, and N represents the number of observations.

Step 4: The purpose of designing a testing phase is to evaluate model performance on results of AI approaches applied to datasets having no function in building models.

Once the selected features (PoE and DoE) are finalized, the most influential features will transfer as input variable for next forecasting process. In order to evaluate the forecasted price, a data set beyond the training data set known as the testing data is employed. Reliability of the generated model can be determined utilizing the considered scores in testing process.

$$MAPE\% = \frac{1}{N} \sum_{t=1}^{N} \frac{\left| PoE(t)_{observed} - PoE(t)_{forecasted} \right|}{PoE(t)_{observed}} \times 100 \tag{14}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} \left(PoE(t)_{observed} - PoE(t)_{forecasted} \right)^2} \tag{15}$$

$$U = \frac{RMSE}{\sqrt{\frac{1}{N} \sum_{t=1}^{N} \left(PoE(t)_{observed} \right)^2} + \sqrt{\frac{1}{N} \sum_{t=1}^{N} \left(PoE(t)_{forecasted} \right)^2}} \tag{16}$$

U-statistic always generates binary results [0, 1], where zero represents higher forecasting precision, and one represents the estimation is as inaccurate as a naïve guess.

The appropriateness description of a given data series obtained through models is ensured through the whiteness test, also known as the Durbin–Watson test [65], acquired after a confirmatory analysis. The objective of the confirmatory analysis is to confirm the whiteness of estimated residuals ($e(t)$), which confirms the non-correlation between them.

$$RACF = \frac{\sum_{t=2}^{N} (e(t)e(t-1))}{\sum_{t=1}^{N} (e(t))^2} \tag{17}$$

To prove the effectiveness of the proposed model, the Akaike Information Criterion (AIC) for different months is calculated [66] as (18). AIC deals with the trade-off between the goodness of fit of the model and the simplicity of the model. In other words, AIC deals with both the risk of overfitting and the risk of underfitting.

$$AIC = n \times \log(\sqrt{RMSE}) + 2k \tag{18}$$

where n represents the number of observations, k represents the number of coefficients optimized by a model and RMSE is the root mean square error.

7. Simulation Results and Discussion

The electricity price forecasting process and development of the whole feature selection techniques were coded in MATLAB (R2019a) and run on a personal computer, with a core-2 quad processor of 2.6 GHz clock speed and 4GB RAM. Regarded as one of the most volatile electricity markets, the implementation of ANFIS-BSA in this study is to improve the precision of Queensland mainland Electricity Price Forecasting (EPF). In this section, substantial features are established by emerging multi-objective BSA and ANFIS as inputs for prediction investigation. Additionally, an assessment of the efficacy of ANFIS-BSA for short-term EPF accurateness purposes is performed by comparison with the well-known AI techniques: ANN, ANFIS, ANN-GA, ANN-PSO, ANFIS-GA, ANN-BSA and ANFIS-PSO. There are two types of variables, dependent and independent variables. They are divided into two subdivisions, where the hourly data of the first three weeks is used for the training of design phase and that last week of each month is utilized for the testing phase of the models, which are obtained from the analysis. The adopted methods with the control parameters are typically based on identical methodologies that have been effectively utilized for energy price or demand forecasting in the literature, as there is no unanimity in the optimal values of the AI-based method settings of the parameters.

A backpropagation MLP as an ANN model, which adopts a feed-forward structure trained with a backpropagation (BP) learning model, is used. This network is usually referred to as a universal estimator because of its simple solution network and faster computational steps, which enables it to have conducted training over a large input of data sets. Several criteria, for example, architecture network, learning algorithm and transfer function impact the effectiveness of developed models by ANN. As recommended in [67], two hidden layers are adopted for short-term EPF to form an MLP model. The layers include Levenberg–Marquardt PB learning and the logarithmic sigmoid transfer function. The formation of the ANFIS structure is based on the selected Sugeno-type FIS instead of the Mamdani-type FIS (Fuzzy Inference System). The reason for selecting the Sugeno-type is that it is able to model non-linear systems based on the interpolation of multiple linear models. As suggested in [68], the selected membership function here is Gaussian, and ANFIS is set up by subtractive clustering (radii = 0.8). Based on the recent studies, different ANFIS models are computed in this work. The results in Table 3 show that subtractive clustering (SC) outperformed the other models of ANFIS.

Table 3. Comparison of different ANFIS models for different months.

Performance Indexes (ANFIS)		November			August			May			February		
		GP	FCM	SC	GP	FCM	SC	GP	FCM	SC	GP	FCM	SC
RACF	Training	0.0007	0.0005	**0.0003**	0.0008	0.0006	**0.0005**	0.0049	0.0048	**0.0047**	0.0031	0.0029	**0.0028**
	Testing	0.0009	0.0008	**0.0006**	0.0018	0.0016	**0.0012**	0.0032	0.0028	**0.0027**	0.0018	0.0016	**0.0015**
Absolute error	Training	31.5827	29.5825	**26.5822**	30.6788	28.6787	**25.6785**	28.3229	24.3228	**21.3226**	22.8533	20.8531	**18.8529**
	Testing	19.5481	16.5479	**12.5477**	21.2569	19.2567	**17.2565**	18.1251	15.1249	**12.1247**	10.3838	8.3836	**6.3834**
RMSE	Training	4.1336	2.1332	**0.1327**	7.2493	5.2481	**0.2478**	6.1169	3.1166	**0.1164**	5.0887	2.0886	**0.0883**
	Testing	3.1219	1.1216	**0.1205**	4.1888	3.1886	**0.1882**	4.0725	2.0723	**0.0721**	2.0817	1.0816	**0.0811**
***U*-statistic**	Training	0.0658	0.0657	**0.0654**	0.0922	0.0918	**0.0916**	0.0518	0.0517	**0.0514**	0.0469	0.0468	**0.0467**
	Testing	0.0558	0.0555	**0.0548**	0.0879	0.0878	**0.0876**	0.0393	0.0392	**0.0391**	0.0419	0.0418	**0.0417**
MAPE (%)	Training	9.1889	8.1887	**5.1884**	8.3792	7.3788	**6.3778**	9.5249	8.5246	**4.5245**	8.6251	6.6248	**4.6237**
	Testing	7.5241	6.5231	**4.5217**	7.6222	6.6216	**5.6212**	7.4342	6.4331	**4.4328**	7.4897	5.2878	**4.0798**

The method recommended in [69] is GA, which was implemented by the genetic algorithm optimization toolbox (GAOT) with the involvement of scattered crossover and mutation of Gaussian. All the control parameters, such as crossover rate and mutation rate, are set to be default as suggested in [70], and the population of the GA is set to 100.

In this study, a standard form of PSO algorithm is also implemented, in which 100 is set for the population number. Value 2.0 is selected for both acceleration factors (c_1 and c_2). A decaying inertia weight ω with a range between 0.9 and 0.4 is selected, as stipulated in [71] with the increase of the running time.

For a cross-over process participation, the control parameter of BSA sets the number of individuals within the population varies in the range of 0% and 100%. Nevertheless, the variation of this parameter is not influenced by the performance of BSA. Thus, to achieve the best solutions, the individuals with the maximum value were selected. The adjustable parameters of the whole technique are presented in Table 4.

Table 4. User-defined parameters for applied techniques.

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

For February, May, August and November 2018, the performance of machine learning methods applications for the EPF of QLD with their computational time are summarized in Tables 5–8, respectively. The accomplished values of RACF validate the estimated residual whiteness in the asserted confidence range for different months for all established models. Regarding Table 5, the precision of the QLD electricity market forecasting method can be achieved by adopting the mean rank approach, which was intended for the indicator based on multi-criteria evaluations. ANFIS-BSA > ANFIS-PSO > ANFIS-GA > MLP-BSA > MLP-PSO > MLP-GA > ANFIS > MLP is the ranked order (absolute error, RMSE, U-statistic and MAPE) for the applied methods of EPF for forecasting effectiveness based on multi-criteria decisions for each indicator, considering the mean rank of the methods.

Table 5. Comparison performance indexes of multiple metaheuristic techniques corresponding to the forecasting accuracy for EPF in February 2018 in QLD.

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
RACF	Training Testing	0.0013 0.0015	0.0006 0.0018	0.0026 0.0007	0.0021 0.0006	0.0028 0.0015	0.0007 0.0019	0.0009 0.0023	0.0003 0.0005
Absolute error	Training Testing	19.2059 7.1035	18.9638 6.5630	18.9977 6.6027	18.6238 6.2471	18.8529 6.3834	18.5327 6.2285	18.1481 5.4873	**17.7113** **5.2846**
RMSE	Training Testing	0.1246 0.0973	0.0926 0.0842	0.0974 0.0843	0.0862 0.0814	0.0883 0.0811	0.0847 0.0777	0.0787 0.0738	**0.0673** **0.0621**
***U*-statistic**	Training Testing	0.0624 0.0478	0.0578 0.0472	0.0516 0.0454	0.0453 0.0422	0.0467 0.0417	0.0432 0.0408	0.0388 0.0378	**0.0256** **0.0331**
MAPE (%)	Training Testing	5.0563 4.5624	4.8523 4.3411	4.9322 4.4503	4.5721 3.9846	4.6237 4.0798	4.4172 3.7671	3.9823 3.2207	**3.8542** **3.0785**
Computational Time		77.4532	76.9853	75.5236	74.3264	63.2352	62.9211	61.2014	**60.3252**

Table 6. Comparison performance indexes of multiple metaheuristic techniques corresponding to the forecasting accuracy for EPF in May 2018 in QLD.

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
RACF	Training	0.0012	0.0014	0.0029	0.0147	0.0047	0.0113	0.0046	0.0015
	Testing	0.0143	0.0003	0.0017	0.0019	0.0027	0.0036	0.0004	0.0039
Absolute error	Training	22.8950	21.7276	21.6552	21.4531	21.3226	20.8263	21.1213	**20.4314**
	Testing	12.2535	12.1624	12.1833	11.7235	12.1247	11.5578	12.1156	**11.3342**
RMSE	Training	0.1411	0.1322	0.1312	0.1226	0.1164	0.0817	0.0812	**0.0763**
	Testing	0.0865	0.0847	0.0835	0.0785	0.0721	0.0714	0.0648	**0.0547**
U-statistic	Training	0.0652	0.0566	0.0546	0.0525	0.0514	0.0457	0.0374	**0.0352**
	Testing	0.0577	0.0512	0.0402	0.0366	0.0391	0.0365	0.0344	**0.0327**
MAPE (%)	Training	4.8784	4.7504	4.6557	4.6263	4.5245	4.4381	4.3341	**3.4621**
	Testing	4.4424	4.3512	4.2769	4.2634	4.4328	4.0567	4.2620	**2.7756**
Computational Time		76.5342	75.3824	75.3343	74.3972	64.3846	63.4102	62.3753	**60.3655**

Table 7. Comparison performance indexes of multiple metaheuristic techniques corresponding to the forecasting accuracy for EPF on August 2018 in QLD.

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
RACF	Training	0.0016	0.0025	0.0021	0.0008	0.0005	0.0011	0.0028	0.0002
	Testing	0.0017	0.0006	0.0015	0.0026	0.0012	0.0014	0.0028	0.0015
Absolute error	Training	25.6465	25.3267	25.5615	25.2446	25.6785	24.2214	23.2596	**21.0278**
	Testing	17.2563	17.2123	18.2275	17.1147	17.2565	17.0014	16.6772	**14.5342**
RMSE	Training	0.3661	0.2537	0.2147	0.2041	0.2478	0.1852	0.1837	**0.1676**
	Testing	0.2048	0.1734	0.1922	0.1653	0.1882	0.1635	0.1613	**0.1453**
U-statistic	Training	0.0975	0.0776	0.0803	0.0738	0.0916	0.0734	0.0728	**0.0685**
	Testing	0.0932	0.0811	0.0873	0.0684	0.0876	0.0667	0.0623	**0.0601**
MAPE (%)	Training	6.3276	5.8356	6.0013	5.7127	6.3778	5.7055	5.6645	**4.4642**
	Testing	6.7728	5.6524	5.8897	5.5524	5.6212	4.5047	4.3357	**4.2225**
Computational Time		78.6574	78.2345	77.3886	77.3672	66.3756	65.3775	65.3473	**62.3585**

The rank (absolute error, RMSE, U-statistic and MAPE) for the entire set in May of 2018 is ANFIS-BSA > ANFIS-GA > MLP-BSA > ANFIS-PSO > MLP-PSO > MLP-GA > ANFIS > MLP. In August, the accurateness of the forecasting methods based on multi-criteria decisions that have been studied within the entire set are ANFIS-BSA > ANFIS-PSO > ANFIS-GA > ANFIS > MLP-BSA > MLP-GA > MLP-PSO > MLP, ranked for every indicator (absolute error, RMSE, U-statistic and MAPE) with the mean rank method application. Meanwhile, in November 2018, the accurateness of the forecasting methods based on multi-criteria decisions that have been studied within the entire set are ANFIS-BSA > ANFIS-PSO > ANFIS-GA > MLP-BSA > MLP-PSO > ANFIS > MLP-GA > MLP, ranked for every indicator (absolute error, RMSE, U-statistic and MAPE) with the mean rank method application. Table 9 reports the AIC index with the better-fit model. For different months of different seasons, the enhanced results indicate the better-fit estimation generated by the optimized ANFIS methodologies as compared to previous researched techniques. Additionally, for training the ANFIS, it is decided that BSA is the most effective optimization algorithm, whereas ANFIS-BSA attained the better MAPE (%), U-statistic, RMSE and absolute error, which is less than the values attained using other optimization algorithms.

Figures 5–8 present the design phase and testing phase of performance tests that use ANFIS-BSA during February, May, August and November, respectively, of each season's execution.

Table 8. Comparison performance indexes of multiple metaheuristic techniques corresponding to the forecasting accuracy for EPF in November 2018 in QLD.

Performance Indexes		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
RACF	Training	0.0008	0.0009	0.0015	2.60E-04	0.0003	0.0016	0.0008	0.0017
	Testing	0.0005	3.50E-05	0.0013	0.0006	0.0006	0.0015	0.0012	0.0006
Absolute error	Training	27.6740	25.0021	25.2140	24.3145	26.5822	24.2364	24.2012	**21.1121**
	Testing	12.6613	12.1456	12.3215	12.0138	12.5477	11.8823	10.5674	**10.0022**
RMSE	Training	0.1411	0.1022	0.1144	0.0934	0.1327	0.0814	0.0811	**0.0776**
	Testing	0.1218	0.0885	0.1106	0.0918	0.1205	0.0876	0.0847	**0.0725**
U-statistic	Training	0.0668	0.0622	0.0644	0.0557	0.0654	0.0524	0.0538	**0.0376**
	Testing	0.5824	0.0504	0.0526	0.0473	0.0548	0.0478	0.0459	**0.0326**
MAPE (%)	Training	5.2457	5.0031	5.1314	4.7810	5.1884	4.6734	4.7126	**3.5762**
	Testing	4.8724	4.4875	4.6810	4.0026	4.5217	3.7320	3.6255	**2.7123**
Computational Time		78.5462	77.3824	76.7943	75.4262	66.5736	66.5462	66.3873	**60.3465**

Table 9. Comparison of different models based on the AIC index for different months.

Akaike Information Criterion Index (AIC)		Methods							
		MLP	MLP-GA	MLP-PSO	MLP-BSA	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS-BSA
February	Training	94.50	93.15	93.38	92.82	64.93	64.74	60.41	63.69
	Testing	100.46	100.25	100.24	100.19	72.18	72.12	72.04	71.78
MAY	Training	93.79	93.45	93.42	93.06	64.79	62.95	62.91	62.59
	Testing	100.28	100.25	100.23	100.13	72.00	71.99	71.84	71.58
August	Training	98.76	96.85	95.98	95.72	68.73	67.21	67.17	66.69
	Testing	101.59	101.34	101.49	101.26	73.46	73.25	73.23	73.07
November	Training	94.22	92.61	93.17	92.16	65.91	63.47	63.45	63.23
	Testing	100.80	100.31	100.65	100.37	72.78	72.29	72.25	72.01

The performance of the models developed by ANFIS-BSA is verified through external validation using different statistical methods for the QLD competitive electricity market in the preceding year of 2018. To assess the performance of the obtained model, the following qualities were recommended [72–75]:

i.   For |R| > 0.8 given by a model, the prediction and observation values have a strong correlation.

ii. For 0.2 < |R| < 0.8 given by a model, the prediction and observation values have a mild correlation.

iii. For gives |R| < 0.2 given by a model, the prediction and observation values have a poor correlation.

The statistical factors of the ANFI-BSA model for EPF of Ontario in different months of 2018 are tabulated in Table 10. The table demonstrates that the developed models meet the entire requisite specifications. The precise models must be provided by ANFIS-BSA during the validation phase because it is strongly relevant to QLD and the implementation in the future of any competitive markets for short-term electricity price forecasting.
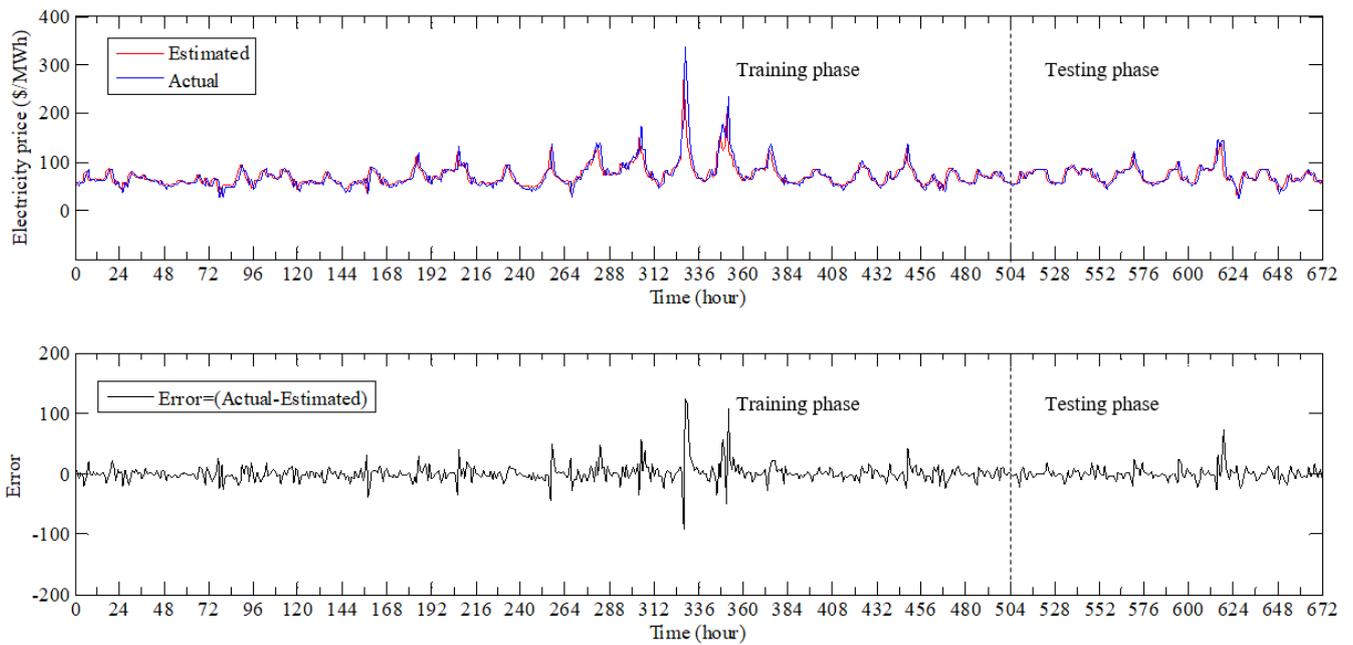
Figure 5. The evaluation of the performance throughout training at the designing and testing phases of ANFIS-BSA and its corresponding error of EPF in February 2018 for QLD.
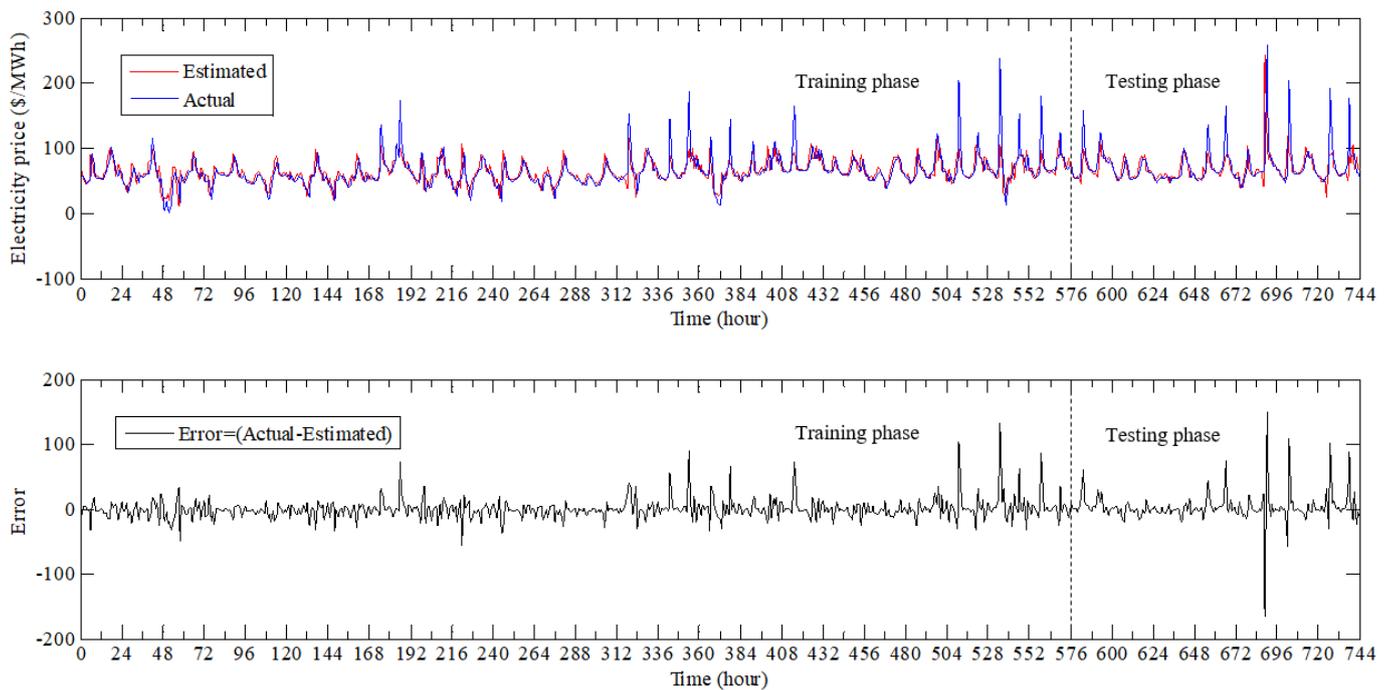

Figure 6. The evaluation of the performance throughout training at the designing and testing phases of ANFIS-BSA and its corresponding error of EPF in May 2018 for QLD.
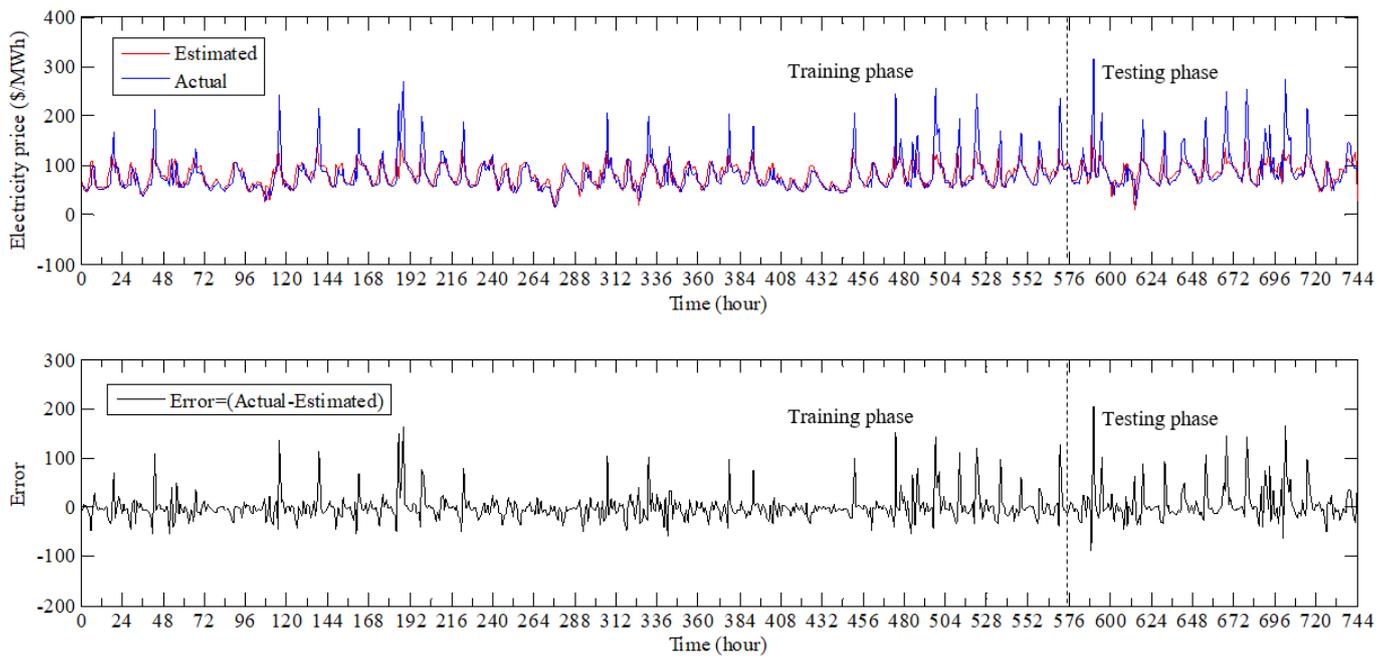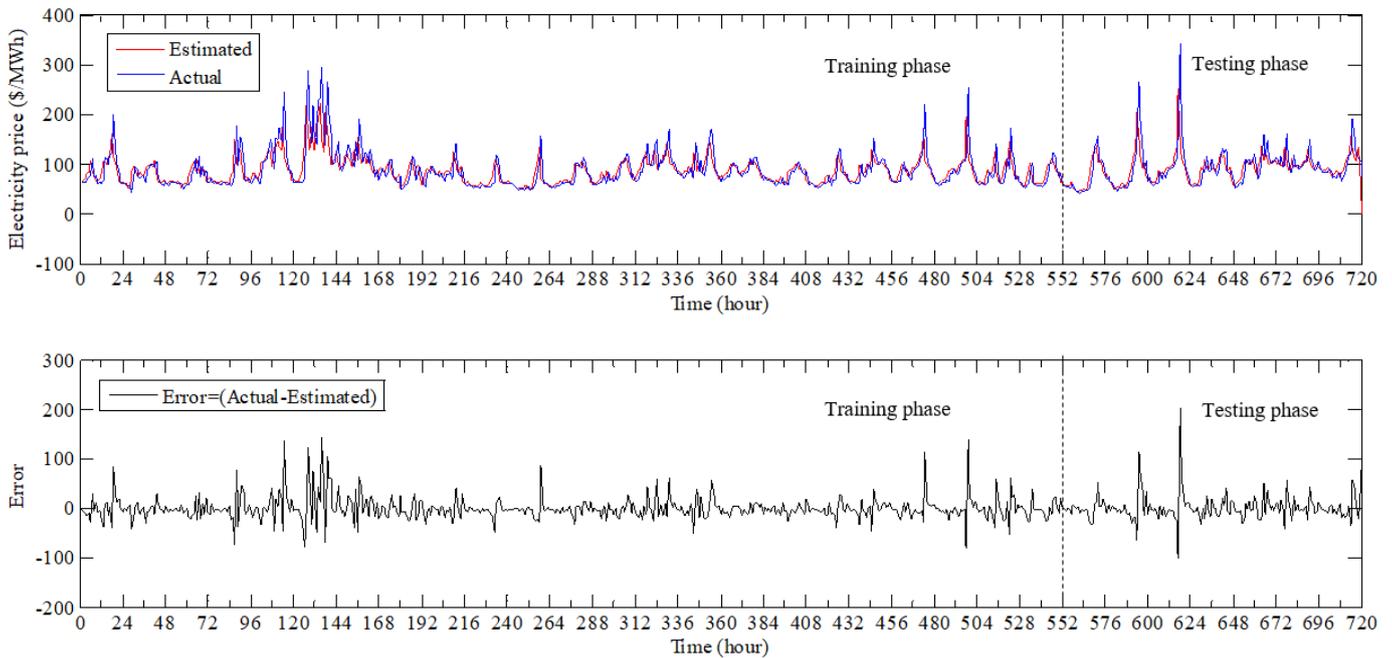
Figure 7. The evaluation of the performance throughout training at the designing and testing phases of ANFIS-BSA and its corresponding error of EPF in August 2018 for QLD.

Figure 8. The evaluation of the performance throughout training at the stage of designing and testing phase of ANFIS-BSA and its corresponding error of EPF in November 2018 for QLD.

Table 10. ANFIS-BSA model's statistical factors for EPF in one month of winter, spring, summer and autumn 2018 for QLD.

Item	Formula	Condition	ANFIS-BSA Approach			
			Winter	**Spring**	**Summer**	**Autumn**
1	R	$0.8 < R_0$	0.9987	0.9933	0.9850	0.9998
2	$K = \frac{\sum_{i=1}^{n}(h_i \times t_i)}{\sum_{i=1}^{n} h_i^2}$	$0.85 < k < 1.15$	0.9976	0.9985	0.9962	0.9970
3	$K' = \frac{\sum_{i=1}^{n}(h_i \times t_i)}{\sum_{i=1}^{n} t_i^2}$	$0.85 < k' < 1.15$	1.0030	1.0011	1.0023	1.0008
4	$m = \frac{R^2 - R_o^2}{R^2}$	$|m| < 0.1$	−0.0033	−0.0039	−0.0042	−0.0024
5	$n = \frac{R^2 - R_{O'}^2}{R^2}$	$|n| < 0.1$	−0.0033	−0.0039	−0.0041	−0.0024
6	$R_m = R^2 \times \left(1 - \sqrt{|R^2 - R_o^2|}\right)$	$0.5 < R_m$	0.9974	0.9923	0.9865	0.9980
where	$R_o^2 = 1 - \frac{\sum_{i=1}^{n}(t_i - h_i^o)^2}{\sum_{i=1}^{n}(t_i - \bar{t}_i)^2}, h_i^o = k \times t_i$	$0.8 < R_0^2 < 1$	0.9998	1.0000	1.0000	0.9996
	$R_{O'}^2 = 1 - \frac{\sum_{i=1}^{n}(h_i - t_i^o)^2}{\sum_{i=1}^{n}(h_i - \bar{h}_i)^2}, t_i^o = K' \times h_i$	$0.8 < R_0^2 < 1$	0.9998	1.0000	1.0000	0.9996

8. Conclusions and Recommendations

An investigation has been carried out over price forecasting methodologies used in recent studies related to the deregulated environment. As a result of the rapid transition in the power market's structure, prediction with a high degree of precision for upcoming prices, which has the capability to maximize the profit margin, has become an irresistible practice for market participants. It is a challenging task to consider the most influential electricity input variables caused by the reliance on the multiple parameters of electricity demand and electricity price. In order to overcome the challenges encountered in the forecasting of electricity prices, various techniques have been proposed to attain a robust model with high precision. The statistical techniques, artificial intelligence-based techniques and hybrid approach have been clearly deployed for the short-term forecasting of electricity prices. While most of the alternatives to electricity price forecasts in the literature were implemented, there are limitations for such approaches. For example, the main deficiencies in data-driven approaches are so many control parameters, and they are extremely sensitive, making it very complicated to initialize these parameters. Although multiple machine-learning techniques are often considered for electricity price forecasting, in order to give a more precise electricity price forecast, the newest methodology is still preferred. Moreover, long-term price forecasting was taken into account in most of the aforementioned literature. The electricity demand trend relies on the seasons. Thus, short-term price forecasting would become more efficient for real-time decision-making on a deregulated electricity market. Feature selection is another important factor affecting the efficacy and precision of the forecast. The selection of features significantly helps in enhancing forecasting abilities and accuracy. Nevertheless, a determination has been made that it is challenging to pick the most robust feature selection approach for the prediction of electricity price regarding the non-linearity of the price signal. In the most recent reference, the authors recommended the testing of a hybrid ANFIS-BSA method instead of Ontario to test its robustness and efficacy. Therefore, the efficiency of ANFIS-BSA on the Queensland electricity market is tested by the proposed method.

The suggested methodology for the selection of features evaluated not only whether the price signal of non-linear dependencies on its variable of input is better compared to the established multi-objective optimization methods, such as MOPSO, NSGAIII and NSGAII, but also eliminating features redundancy commonly to select the different related features. The classification of fewer input variables by the proposed combination of a MOBBSA-ANFIS feature selection approach, on the other hand, allows better accuracy in forecasting. The theoretical feature selection technique with the information of the non-redundant features between a big set input of candidates also has the potential to choose

the most appropriate minimum subset of the proposed strategies. Hence, the candidate inputs are eventually chosen with minimal redundancy, and the output is highly relevant.

Finally, the backtracking search algorithm (BSA) is used as an efficient optimization algorithm in the learning phase of the ANFIS approach to provide a more precise prediction engine for forecasting the price of electrical energy. The predicted results are compared with ANN and ANFIS models' results, which are optimized by particle swarm optimization (PSO) and genetic algorithms (GA) to determine the effectiveness and applicability of the proposed approach for electricity price forecasting. The Hybrid ANFIS-BSA approach provides a good forecasting accuracy, with an average MAPE of 3.07, 2.77, 4.22 and 2.71 in February, May, August and November, respectively, taking into account the results of previous publications for different case studies as benchmark described in Table 1. The verifications of the results of the simulations used the real sets of data based on the electricity market of Queensland, which is amongst the most volatile electricity markets globally. The BSA-optimized ANFIS strategy is worth being considered as a robust and useful forecast mechanism in order to fulfill the actual needs of electricity market participants, including the self-producers and traditional generation companies, the suppliers/retailers and aggregators. These contributions will help market players efficiently bid, keep their daily business productive and ultimately increase the revenue of the companies.

Author Contributions: Conceptualization, A.P., M.M. (Mohammad Mohammadi) and M.M. (Mazaher Karimi); methodology, A.P. and S.H.A.K.; software, A.P.; validation, A.P. and S.A.; formal analysis, A.P. and S.A.; Investigation, A.P. and M.M. (Mohammad Mohammadi); resources, M.M. (Mazaher Karimi); data curation, A.P. and H.M.; Writing—Original Draft, A.P.; Writing—Review and Editing, all authors; visualization, A.P. and S.H.A.K.; supervision, M.M. (Mohammad Mohammadi), H.M. and H.A.I.; Project administration, A.P., M.M. (Mohammad Mohammadi) and H.M.; funding acquisition, M.M. (Mazaher Karimi). All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: This research does not involve humans and animals related research.

Informed Consent Statement: This research does not involve humans.

Data Availability Statement: The data used in this study obtained from Australian Energy Market Operator and are available online: https://www.aemo.com.au/energy-systems/electricity/national-electricity-market-nem/data-nem (accessed on 20 September 2021).

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

Nomenclature

$\bar{Z}$	Normalized data	$f_i(x, y; p_i, q_i, r_i)$	Output of the Sugeno type FIS
Z	Un normalized data	y_i	Minimum value of the jth objective function
:=	Update operation	P_g	Productivity of ith individual
ARIMA	Auto regressive integrated moving average	NL_{PoE}	Number of lag order for electricity price
(A_i, B_i)	Fuzzy sets	g	Transfer function
(x, y)	Inputs to node i	$e(t)$	Estimated residuals whiteness
$\overline{w}$	Normalized firing strength of a rule	$H(X)$	Random variable X's entropy
μ_{A_i}	Membership function for A_i fuzzy sets	$P(X)$	Probability function of X
μ_{B_i}	Membership function for B_i fuzzy sets	mixrate	BSA Control parameter
∂ and ε	Solutions in Pareto optimal set	$H(X,Y)$	Variables X and Y joint probability distribution
PoE $(t)_{observed}$	During time t electricity price observed	g + 1	Next generation
PoE $(t)_{forecasted}$	During time t electricity price predicted	low_j	Lower search space limits of jth variable
map	Binary integer-valued matrix	nPop	Population size of host nests
Mutant	Initial form of trial population	nVar	Number of respective optimization variable
N	Standard normal distribution	P_{best}	Previous best position
oldP	Historical population	permuting	Random shuffling function

y_g	Global minimum	$Q_{1,i}$	Membership grade of a fuzzy set
rand	Distributed random numbers	T	Generated offspring at the end of crossover process
randi	Random selection function	U	Uniform distribution function
up_j	Upper search space limits of jth variable	σ	Width of Guassian MF
w_i	Output signal and represent the firing strength of a rule	F	Wiener process
a and b in BSA	Randomly generated numbers	DoE	Electricity demand at time t
f_{out}	Final output of ANFIS model	RACF	Residuals autocorrelation function
c	Center of Guassian MF	RMSE	Root mean square error
CE	Conditional entropy	MI	Mutual Information
AI	Artificial intelligence	ANFIS	Adaptive neuro-fuzzy inference system
PoE	Electricity price at time t	ANN	Artificial neural network
BSA	Backtracking search algorithm	$/MW.h	dollar per Megawatts hour
EPF	Electricity Price Forecasting	SC	Subtractive clustering
NL_{DoE}	Lag order number	QLD	Queensland electricity market
MAPE	Mean absolute percentage error	MOPSO	Multi-objective particle swarm optimization
NSGA	Non-dominated sorting genetic algorithm	MOBBSA	Multi-objective binary-valued BSA

References

1. Kurbatsky, V.G.; Sidorov, D.N.; Spiryaev, V.A.; Tomin, N.V. Forecasting nonstationary time series based on Hilbert-Huang transform and machine learning. *Autom. Remote Control* **2014**, *75*, 922–934. [CrossRef]
2. Gao, W.; Sarlak, V.; Parsaei, M.R.; Ferdosi, M. Combination of fuzzy based on a meta-heuristic algorithm to predict electricity price in an electricity markets. *Chem. Eng. Res. Des.* **2018**, *131*, 333–345. [CrossRef]
3. Panapakidis, I.P.; Dagoumas, A.S. Day-ahead electricity price forecasting via the application of artificial neural network based models. *Appl. Energy* **2016**, *172*, 132–151. [CrossRef]
4. Yang, Z.; Ce, L.; Lian, L. Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods. *Appl. Energy* **2017**, *190*, 291–305. [CrossRef]
5. Bask, M.; Widerberg, A. Market structure and the stability and volatility of electricity prices. *Energy Econ.* **2009**, *31*, 278–288. [CrossRef]
6. Ziel, F.; Steinert, R.; Husmann, S. Efficient modeling and forecasting of electricity spot prices. *Energy Econ.* **2015**, *47*, 98–111. [CrossRef]
7. Singhal, D.; Swarup, K. Electricity price forecasting using artificial neural networks. *Int. J. Electr. Power Energy Syst.* **2011**, *33*, 550–555. [CrossRef]
8. The Independent Electricity System Operator (IESO). Available online: http://www.ieso.ca/ (accessed on 17 November 2018).
9. Razak, I.A.W.A.; Abidin, I.Z.; Yap, K.S.; Abidin, A.A.Z.; Rahman, T.K.A.; Nasir, M.N.M. A novel hybrid method of LSSVM-GA with multiple stage optimization for electricity price forecasting. In Proceedings of the Power and Energy (PECon), 2016 IEEE International Conference, Melaka City, Malaysia, 28–29 November 2016; pp. 390–395.
10. Zareipour, H.; Bhattacharya, K.; Canizares, C.A. Electricity market price volatility: The case of Ontario. *Energy Policy* **2007**, *35*, 4739–4748. [CrossRef]
11. Amjady, N.; Daraeepour, A. Design of input vector for day-ahead price forecasting of electricity markets. *Expert Syst. Appl.* **2009**, *36*, 12281–12294. [CrossRef]
12. Aggarwal, S.K.; Saini, L.M.; Kumar, A. Electricity price forecasting in deregulated markets: A review and evaluation. *Int. J. Electr. Power Energy Syst.* **2009**, *31*, 13–22. [CrossRef]
13. Weron, R. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *Int. J. Forecast.* **2014**, *30*, 1030–1081. [CrossRef]
14. Cerjan, M.; Krželj, I.; Vidak, M.; Delimar, M. A literature review with statistical analysis of electricity price forecasting methods. In Proceedings of the EUROCON, 2013 IEEE, Tool, Zagreb, Croatia, 1–4 July 2013; pp. 756–763.
15. Abedinia, O.; Amjady, N.; Shafie-Khah, M.; Catalão, J. Electricity price forecast using combinatorial neural network trained by a new stochastic search method. *Energy Convers. Manag.* **2015**, *105*, 642–654. [CrossRef]
16. Amjady, N.; Daraeepour, A. Mixed price and load forecasting of electricity markets by a new iterative prediction method. *Electr. Power Syst. Res.* **2009**, *79*, 1329–1336. [CrossRef]
17. Amjady, N.; Hemmati, M. Day-ahead price forecasting of electricity markets by a hybrid intelligent system. *Int. Trans. Electr. Energy Syst.* **2009**, *19*, 89–102. [CrossRef]
18. Amjady, N.; Keynia, F. Day-ahead price forecasting of electricity markets by a new feature selection algorithm and cascaded neural network technique. *Energy Convers. Manag.* **2009**, *50*, 2976–2982. [CrossRef]
19. Anbazhagan, S.; Kumarappan, N. Day-ahead deregulated electricity market price forecasting using neural network input featured by DCT. *Energy Convers. Manag.* **2014**, *78*, 711–719. [CrossRef]

20. Hahn, H.; Meyer-Nieberg, S.; Pickl, S. Electric load forecasting methods: Tools for decision making. *Eur. J. Oper. Res.* **2009**, *199*, 902–907. [CrossRef]

21. Shiri, A.; Afshar, M.; Rahimi-Kian, A.; Maham, B. Electricity price forecasting using Support Vector Machines by considering oil and natural gas price impacts. In Proceedings of the Smart Energy Grid Engineering (SEGE), 2015 IEEE International Conference, Oshawa, ON, Canada, 17 August 2015; pp. 1–5.

22. Bigdeli, N.; Afshar, K.; Amjady, N. Market data analysis and short-term price forecasting in the Iran electricity market with pay-as-bid payment mechanism. *Electr. Power Syst. Res.* **2009**, *79*, 888–898. [CrossRef]

23. Pousinho, H.M.I.; Mendes, V.M.F.; Catalão, J.P.D.S. Short-term electricity prices forecasting in a competitive market by a hybrid PSO–ANFIS approach. *Int. J. Electr. Power Energy Syst.* **2012**, *39*, 29–35. [CrossRef]

24. Singh, N.; Mohanty, S.R.; Shukla, R.D. Short term electricity price forecast based on environmentally adapted generalized neuron. *Energy* **2017**, *125*, 127–139. [CrossRef]

25. 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]

26. 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]

27. Zhou, S.; Zhou, L.; Mao, M.; Tai, H.-M.; Wan, Y. An optimized heterogeneous structure LSTM network for electricity price forecasting. *IEEE Access* **2019**, *7*, 108161–108173. [CrossRef]

28. Zhang, X.; Wang, J.; Gao, Y. A hybrid short-term electricity price forecasting framework: Cuckoo search-based feature selection with singular spectrum analysis and SVM. *Energy Econ.* **2019**, *81*, 899–913. [CrossRef]

29. Parhizkari, L.; Najafi, A.; Golshan, M. Medium term electricity price forecasting using extreme learning machine. *J. Energy Manag. Technol.* **2020**, *4*, 20–27.

30. Lee, C.-Y.; Wu, C.-E. Short-Term Electricity Price Forecasting Based on Similar Day-Based Neural Network. *Energies* **2020**, *13*, 4408. [CrossRef]

31. Hong, Y.-Y.; Taylar, J.V.; Fajardo, A.C. Locational marginal price forecasting in a day-ahead power market using spatiotemporal deep learning network. *Sustain. Energy Grids Netw.* **2020**, *24*, 100406. [CrossRef]

32. Huang, C.J.; Shen, Y.; Chen, Y.H.; Chen, H.C. A novel hybrid deep neural network model for short-term electricity price forecasting. *Int. J. Energy Res.* **2021**, *45*, 2511–2532. [CrossRef]

33. Zhang, H.; Yang, Y.; Zhang, Y.; He, Z.; Yuan, W.; Yang, Y.; Qiu, W.; Li, L. A combined model based on SSA, neural networks, and LSSVM for short-term electric load and price forecasting. *Neural Comput. Appl.* **2021**, *33*, 773–788. [CrossRef]

34. Jahangir, H.; Tayarani, H.; Baghali, S.; Ahmadian, A.; Elkamel, A.; Golkar, M.A.; Castilla, M. A novel electricity price forecasting approach based on dimension reduction strategy and rough artificial neural networks. *IEEE Trans. Ind. Inform.* **2019**, *16*, 2369–2381. [CrossRef]

35. El Assadi, A. Multistep Electricity Price Forecasting for Deregulated Energy Markets: GAN-Based Reinforcement Learning. Ph.D. Dissertation, Texas A & M University, Texas, USA, 2021.

36. Zhang, R.; Li, G.; Ma, Z. A deep learning based hybrid framework for day-ahead electricity price forecasting. *IEEE Access* **2020**, *8*, 143423–143436. [CrossRef]

37. Azam, M.F.; Younis, S. Multi-Horizon Electricity Load and Price Forecasting using an Interpretable Multi-Head Self-Attention and EEMD-Based Framework. *IEEE Access* **2021**, *9*, 85918–85923. [CrossRef]

38. Memarzadeh, G.; Keynia, F. Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm. *Electr. Power Syst. Res.* **2021**, *192*, 106995. [CrossRef]

39. Sun, L.; Li, L.; Liu, B.; Saeedi, S. A Novel Day-ahead Electricity Price Forecasting Using multi-modal combined Integration via Stacked Pruning Sparse Denoising Auto Encoder. *Energy Rep.* **2021**, *7*, 2201–2213. [CrossRef]

40. Bhatia, K.; Mittal, R.; Varanasi, J.; Tripathi, M. An ensemble approach for electricity price forecasting in markets with renewable energy resources. *Util. Policy* **2021**, *70*, 101185. [CrossRef]

41. Australian Energy Market Operator. Data (NEM). Available online: https://www.aemo.com.au/energy-systems/electricity/national-electricity-market-nem/data-nem (accessed on 20 January 2020).

42. Unler, A.; Murat, A. A discrete particle swarm optimization method for feature selection in binary classification problems. *Eur. J. Oper. Res.* **2010**, *206*, 528–539. [CrossRef]

43. Renani, E.T.; Elias, M.F.M.; Rahim, N.A. Using data-driven approach for wind power prediction: A comparative study. *Energy Convers. Manag.* **2016**, *118*, 193–203. [CrossRef]

44. Amjady, N.; Hemmati, M. Energy price forecasting-problems and proposals for such predictions. *IEEE Power Energy Mag.* **2006**, *4*, 20–29. [CrossRef]

45. Jang, J.S.R. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [CrossRef]

46. Gao, W.; Moayedi, H.; Shahsavar, A. The feasibility of genetic programming and ANFIS in prediction energetic performance of a building integrated photovoltaic thermal (BIPVT) system. *Sol. Energy* **2019**, *183*, 293–305. [CrossRef]

47. Tavana, M.; Fallahpour, A.; Di Caprio, D.; Santos-Arteaga, F.J. A hybrid intelligent fuzzy predictive model with simulation for supplier evaluation and selection. *Expert Syst. Appl.* **2016**, *61*, 129–144. [CrossRef]

48. Svalina, I.; Galzina, V.; Lujić, R.; Šimunović, G. An adaptive network-based fuzzy inference system (ANFIS) for the forecasting: The case of close price indices. *Expert Syst. Appl.* **2013**, *40*, 6055–6063. [CrossRef]

49. Holland, J.; Goldberg, D. *Genetic Algorithms in Search, Optimization and Machine Learning*; Addison-Wesley Read: Boston, MA, USA, 1989.

50. Srinivas, N.; Deb, K. Muiltiobjective optimization using nondominated sorting in genetic algorithms. *Evol. Comput.* **1994**, *2*, 221–248. [CrossRef]

51. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [CrossRef]

52. McClymont, K.; Keedwell, E. Deductive sort and climbing sort: New methods for non-dominated sorting. *Evol. Comput.* **2012**, *20*, 1–26. [CrossRef]

53. Zhang, X.; Tian, Y.; Cheng, R.; Jin, Y. An efficient approach to nondominated sorting for evolutionary multiobjective optimization. *IEEE Trans. Evol. Comput.* **2015**, *19*, 201–213. [CrossRef]

54. Modiri-Delshad, M.; Rahim, N.A. Multi-objective backtracking search algorithm for economic emission dispatch problem. *Appl. Soft Comput.* **2016**, *40*, 479–494. [CrossRef]

55. Zitzler, E.; Thiele, L. Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Trans. Evol. Comput.* **1999**, *3*, 257–271. [CrossRef]

56. Coello, C.C. Evolutionary multi-objective optimization: A historical view of the field. *IEEE Comput. Intell. Mag.* **2006**, *1*, 28–36. [CrossRef]

57. Civicioglu, P. Backtracking search optimization algorithm for numerical optimization problems. *Appl. Math. Comput.* **2013**, *219*, 8121–8144. [CrossRef]

58. Xue, B.; Zhang, M.; Browne, W.N. Particle swarm optimization for feature selection in classification: A multi-objective approach. *IEEE Trans. Cybern.* **2013**, *43*, 1656–1671. [CrossRef]

59. Xue, B.; Zhang, M.; Browne, W.N. Particle swarm optimisation for feature selection in classification: Novel initialisation and updating mechanisms. *Appl. Soft Comput.* **2014**, *18*, 261–276. [CrossRef]

60. Zhang, C.; Zhou, J.; Li, C.; Fu, W.; Peng, T. A compound structure of ELM based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting. *Energy Convers. Manag.* **2017**, *143*, 360–376. [CrossRef]

61. Ahmed, T.; Mekhilef, S.; Shah, R.; Mithulananthan, N.; Seyedmahmoudian, M.; Horan, B. ASEAN power grid: A secure transmission infrastructure for clean and sustainable energy for South-East Asia. *Renew. Sustain. Energy Rev.* **2017**, *67*, 1420–1435. [CrossRef]

62. Zou, F.; Chen, D.; Li, S.; Lu, R.; Lin, M. Community Detection in Complex Networks: Multi-objective Discrete Backtracking Search Optimization Algorithm with Decomposition. *Appl. Soft Comput.* **2017**, *53*, 285–295. [CrossRef]

63. Alizadeh, M.; Jolai, F.; Aminnayeri, M.; Rada, R. Comparison of different input selection algorithms in neuro-fuzzy modeling. *Expert Syst. Appl.* **2012**, *39*, 1536–1544. [CrossRef]

64. Jang, J.-S. Input selection for ANFIS learning. In Proceedings of the Fuzzy Systems, 1996 Fifth IEEE International Conference, New Orleans, LA, USA, 8–11 September 1996; pp. 1493–1499.

65. AlRashidi, M.; El-Naggar, K. Long term electric load forecasting based on particle swarm optimization. *Appl. Energy* **2010**, *87*, 320–326. [CrossRef]

66. Gholami, A.; Bonakdari, H.; Ebtehaj, I.; Shaghaghi, S.; Khoshbin, F. Developing an expert group method of data handling system for predicting the geometry of a stable channel with a gravel bed. *Earth Surf. Process. Landf.* **2017**, *42*, 1460–1471. [CrossRef]

67. Ekonomou, L. Greek long-term energy consumption prediction using artificial neural networks. *Energy* **2010**, *35*, 512–517. [CrossRef]

68. Azadeh, A.; Saberi, M.; Gitiforouz, A.; Saberi, Z. A hybrid simulation-adaptive network based fuzzy inference system for improvement of electricity consumption estimation. *Expert Syst. Appl.* **2009**, *36*, 11108–11117. [CrossRef]

69. Ünler, A. Improvement of energy demand forecasts using swarm intelligence: The case of Turkey with projections to 2025. *Energy Policy* **2008**, *36*, 1937–1944. [CrossRef]

70. Azadeh, A.; Saberi, M.; Seraj, O. An integrated fuzzy regression algorithm for energy consumption estimation with non-stationary data: A case study of Iran. *Energy* **2010**, *35*, 2351–2366. [CrossRef]

71. Askarzadeh, A. Comparison of particle swarm optimization and other metaheuristics on electricity demand estimation: A case study of Iran. *Energy* **2014**, *72*, 484–491. [CrossRef]

72. Mostafavi, E.S.; Mousavi, S.M.; Hosseinpour, F. Gene Expression Programming as a Basis for New Generation of Electricity Demand Prediction Models. *Comput. Ind. Eng.* **2014**, *74*, 120–128.

73. Golbraikh, A.; Tropsha, A. Beware of q 2! *J. Mol. Graph. Model.* **2002**, *20*, 269–276. [CrossRef]

74. Roy, P.P.; Roy, K. On some aspects of variable selection for partial least squares regression models. *QSAR Comb. Sci.* **2008**, *27*, 302–313. [CrossRef]

75. Alavi, A.H.; Aminian, P.; Gandomi, A.H.; Esmaeili, M.A. Genetic-based modeling of uplift capacity of suction caissons. *Expert Syst. Appl.* **2011**, *38*, 12608–12618. [CrossRef]