[1]Mohammad Taghi Bagheri Shadbad

[2]*Mahdi Zeynali

[3]Mohammad Reza Ebrahimi Dishabi

A Combination Model of Artificial Bee Colony and Multilayer Perceptron Neural Network to Calculate the Cost of Electricity Distribution

***Abstract:* -** After the restructuring of the electricity market and turning it into a competitive market where the price of electricity is determined by the forces governing the market, the cost and supply and demand, price fluctuations in this market have increased. Today, issues related to predicting the cost of electricity are of great importance, and several methods have been proposed to predict it. Since the electricity market price model usually has complex features such as instability, nonlinear conditions and high fluctuations, mathematical and traditional methods are usually not suitable for solving these problems. As a result, artificial intelligence methods have higher accuracy in predicting the full price. Are in power distribution networks. In this paper, by integrating the artificial bee colony optimization algorithm and neural network, a new and efficient method for calculating and predicting the cost of electricity distribution network is proposed. In the proposed method, first the initial weighting of the neural network weights is determined using the particle optimization algorithm and then in the next step, based on the perceptron algorithm, the final weights of the neural network are updated to calculate the cost of electricity distribution. After implementing the prediction model developed in this paper, its performance was evaluated and the accuracy of the proposed method was compared with previous works, and the simulation results showed that the proposed method has a higher accuracy than previous works.

Keywords: Electricity, Electricity Distribution, Forecasting, Data Mining, artificial bee colony Algorithm, Neural Networks, Multilayer Perceptron.

I. INTRODUCTION

Today, energy consumption management and distribution is of particular importance among all countries in the world. Awareness of the energy consumption process is important for many managers and economic and political planning. Due to the social and economic sensitivities of energy and the high use of fossil resources, countries are looking for ways to reduce energy consumption and production costs and to manage and optimize consumption[1].

In the previous structure of the electricity industry in most countries of the world, the tasks of generating, transmitting and distributing electricity were the responsibility of complex and mostly state-owned electricity companies[2]. The increase in the demand for power in recent years in many countries has caused these companies to not be able to effectively respond to this increase in demand, as a result of blackouts, power outages and equipment malfunctions, etc. in many countries and consequently That cost of electricity went up sharply during the peak period. At the same time, with the economic growth of the countries, which led to an increase in the amount of energy they needed, the issue of quality of power and reliability became important. It was no longer possible to provide all electricity services, including production, transmission and distribution by one company, economically and competitively, and independent companies were formed in each sector of the electricity industry, including production, transmission and distribution partners[3].

In many countries of the world, the production sector in the electricity market has been competitive, and its future strategy in the retail sector is to become competitive. In the new structure of the electricity industry, electricity distribution companies, which are currently in charge of operating the distribution and energy sales network, must competitively purchase electricity from Regional Electricity and provide it to various customers, both domestic and commercial. And sell industrial. Creating competition in buying and selling electricity will increase efficiency in the distribution sector and help the country's electricity industry achieve its goals[4].

[1] [1]Department of Computer, Miyaneh branch, Islamic Azad University, Miyaneh, Iran.

[2] Department of Accounting, Tabriz branch, Islamic Azad University, Tabriz, Iran.

[3] Department of Computer, Miyaneh branch, Islamic Azad University, Miyaneh, Iran.

* Corresponding author

In systems with integrated vertical structure, power generation, transmission and distribution units are exclusively controlled by the government or an exclusive entity that is responsible for monitoring the monopoly market and even determining the exclusive prices. In unstructured power systems, consumers bought energy from distribution companies that were exclusive[5]. These consumers received their energy only at prices predetermined by the system operator. Manufacturers also produced the energy required by the power grid only at prices specified by the system operator and under traditional contracts and tariffs. In such power systems, network losses were very high. With the restructuring of the electricity industry and the entry of electricity as a commodity into competitive markets for electricity, consumers have direct access to the local distribution system and can choose their own manufacturer[6].

The presence of scattered production units has provided an economic opportunity for the distribution company to increase its profits by buying surplus electricity at a price lower than the purchase price of electricity in the wholesale market. On the other hand, these companies can reduce energy losses and make a financial profit by properly capaciting the distribution network in addition to improving the voltage profile and releasing the existing equipment capacity[7].

With the advent of smart power distribution networks, the simultaneous prediction of load and price in the market is a serious challenge[8]. In practice, price and demand forecasting in smart distribution networks is particularly complex due to the influence of various non-linear factors such as weather conditions, daily and weekly changes, and the dependence of price and demand on each other[9].

In competitive electricity markets, the price of electricity is a key factor in determining short-term operating plans and tender strategies. So many data-driven methods have been proposed to model electricity market forecasting. The results of these predictions have an error of about 5 to 36 percent, which is different due to the different methods used for forecasting and the different analyzes that are done from the market. This error range is large compared to the short-term load prediction error range, which is about 1 to 3 percent.

Various factors cause the accuracy of electric energy price prediction models to decrease. The unpredictable forced exit of production units, variable and complex price models, the integration of energy production resources, the implementation of load response programs to increase reliability all make it impossible to offer an accurate model for fluctuations and changes in electricity prices[10].

Previous research has shown that existing models try to predict the price of electricity in the coming hours and close to the formation of real (instantaneous) prices. However, not all market participants need to know the exact price in the future in their decision-making processes. For example, after the introduction of smart grid technology and new market designs, it is expected that with the interaction of the demand side, home, commercial and industrial levels will have a wide presence in the market. Due to the inactive nature of many electrical loads, especially at the residential level, when the price is higher than a certain level (threshold), participants in the demand side market may react at first, but beyond if the price is higher. It is considered insignificant for the consumer and the consumer still consumes electricity despite the fact that the price is very expensive. In addition, many load response programs may have a specific price for certain hours. Another decision based on the minimum price is when the facilities at the power plants buy energy from the network for domestic use, and the market price is less than the cost of their production.

In this type of application, where the exact value of the price is not a priority, the price forecast can be reduced as a sub-problem to determine the price. Some research is based on the premise that short-term price classification methods can replace numerical price forecasting. In the classification, according to the initial data and price targets, the price is classified into two categories above the threshold or below the threshold.

Data mining in developed and developing countries is one of the hottest and most used topics. Also, many developing countries are using intelligence from developing countries to intelligentize their organizations with data mining. New knowledge of data mining as the newest and best method of problem solving, using advanced combined algorithms and through intelligent decision-making process in the evolution of the power industry and water and electricity in developed countries has a special and vital place. The use of new data analysis methods such as data mining is very important for the databases of this industry and therefore today the knowledge of data mining plays an effective role in all small and large issues and planning of water and electricity industry.

Numerous data mining methods have been proposed to predict the cost of a power distribution network. One effective way to calculate the cost of electricity is to use neural network-based methods. Due to the fact that in the MLP neural network learning method, the initial weighting is done randomly, it is possible that if the initial weights are far from the optimal weights, or the convergence time is prolonged, or it is possible to achieve the optimal weight at all. Do not provide a place. As a result, primary weights need to be produced non-randomly and purposefully. Therefore, in this dissertation, it is suggested that instead of randomly generating primary weights, artificial bee colony algorithm be used for this purpose.

The reminder of the paper is organized as follows: Section 2 gives a brief review of previous works. The proposed method is presented in Section 3. The experimental results on well-known dataset are reported in section 4 and finally, section 5 concludes the overall work.

II. RELATED WORKS

Electricity price forecasting is crucial for decision makers to arrange bidding strategies and develop appropriate products, it has been one of the main efforts of energy market researchers and practitioners. However, electricity price prediction has volatility, it can be affected by multiple factors, such as holidays, temperature, festival and season. Moreover, due to fluctuations, unbalanced power resources may lead to increased power loads and higher electricity prices [11, 12]. In recent years, many researchers have demonstrated that power-related predictions are a daunting task because power demand depends to a large extent on many factors [13-15]. Weron [16] pointed out that electricity price forecasting methods can be roughly classified into five categories: statistical approaches, computational intelligence techniques multi-agent models, fundamental methods and reduced-form models, among which statistical methods and computational intelligence techniques are most commonly used.

Previous statistical methods used in electricity price forecasting include exponential smoothing methods, AR-type time series models just like the well-known AutoRegressive Moving Average (ARMA), AutoRegressive Integrated Moving Average (ARIMA), and Generalized AutoRegressive Conditional Heteroskedastic (GARCH). Tan et al. [17] proposed a method of electricity price prediction using WT combined with ARIMA and GARCH, and verified in the PJM and Spanish power markets. Dong et al. proposed a hybrid model based on Empirical Mode Decomposition (EMD), Seasonal Adjustment (SA) and Autoregressive Integrated Moving Average (ARIMA) and tested it on electricity price dataset from New South Wales in Australia [18]. However, many statistical methods always focused on combining historical data with current data using a mathematical combination approach and limited to handle complex or nonlinear time series problems.

Computational intelligence methods have elicited widely attention for its outstanding performance in handle complex and nonlinear problems. In terms of electricity price forecasting, traditional machine learning has achieved many impressive results, of which extreme learning machine (ELM), support vector machine (SVM) and artificial neural network (ANN) are widely used. The author of [19] used artificial neural network (ANN) and hybrid model of ANN combined with clustering algorithm for Day-ahead price forecasting. ELM coupled with wavelet technique was used by Shrivastava et al. [20] to improve the forecasting accuracy as well as reliability based on the dataset from Ontario, PJM, New York and Italian Electricity markets. A multiple support vector machine (SVM) based mid-term electricity market clearing price forecasting model is proposed by Yan et al. [21] and the PJM interconnection data are used to test the proposed model. With the rapid development of artificial intelligence, deep learning has attracted considerable attention for its outstanding performance in language modeling, speech recognition, and natural language inference. Compared to traditional machine learning methods, deep neural networks can analyze deep and complex nonlinear relationships through hierarchical and distributed feature representations. Recurrent neural network (RNN) is a powerful deep learning method for processing sequential data such as sound and stock market, which differ from the traditional feedforward networks in the sense that they don't only have neural connections on a single direction, in other words, neurons can pass data to a previous or the same layer. However, due to the existence of “long-term dependencies”, the traditional RNNs encountered obstacles to the study of time series analysis.

Long Short Term Memory networks e usually just called “LSTM” are a special kind of RNN, capable of learning long-term dependencies. They were introduced by Hochreiter & Schmidhuber (1997), and were refined and popularized by many people in many areas. They work tremendously well on a large variety of problems, and are now widely used in sequence analysis, speech recognition and natural language processing. The author of [22] has forecasted the volatility of stock price index with hybrid model of LSTM and multiple GARCH, In Ref. [23] , Liu et al. predicted the wind speed with the hybrid model of empirical wavelet transform (EWT), Elman neural network and LSTM. Long Short Term Memory networks are explicitly designed to avoid the long-term dependency problem. Remembering information for long periods of time is practically their default behavior, not something they struggle to learn.

The actual electricity price data is nonlinear and non-stationary. Peng et al. [24] has adopted differential evolution (DE) algorithm to identify parameters for LSTM while predicting electricity prices, and experimentally verified that Long Short Term Memory networks have a clear advantage in accuracy of electricity price forecasting over some statistical approaches and traditional computational intelligence methods, mainly because it can better capture features and deal with the irregularities in electricity prices. The method based on stochastic gradient has core practical importance in the field of deep learning optimization, which plays a key role in the accuracy of prediction. Adam is one of efficient stochastic optimization that only requires first-order gradients with little memory requirement it has combined the advantages of two popular methods: AdaGrad [25], which works well with sparse gradients, and RMSProp [26], which has an excellent performance in non-line and non-stationary settings. It's easy implementation, little memory requirements and appropriateness for non-stationary make Adam an effective and efficient algorithm that has been used successfully in previous studies [27-29]. It can achieve the goal of optimizing the deep learning model by finding a series of parameters to minimize the objective function. Therefore, the Adam optimized LSTM neural network is considered as a powerful tool to predict electricity price.

The wavelet transform, denoted as WT, is a data processing method that provides very useful decomposition information in the time domain and frequency domain, making it suitable for analyzing non-stationary signals such as time series. It combines many other models to form a hybrid model to improve the performance of the model, which is reflected in many previous literatures [30, 31]. Yang et al. [32] presented a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods for electricity price forecasting on the dataset from PJM, Australian and Spanish markets; The author of [33] suggested a hybrid method that combined the WT, the Gravitational Search Algorithm and LSSVM and evaluated by using electricity price data from Iran's, Ontario's and Spain's price markets.

III. PROPOSED METHOD

In this section, a novel data mining method is proposed by integrating the Neural Network with Artificial Bee colony (NNPSO) to predict and calculate the cost of electricity in a distribution network. In the reminder of this section, the details of the proposed method are explained.

preprocessing

The importance of data processing is very important and vital in data mining processes, so that in many cases, a good pre-processing can make it easier to continue data mining processes. In fact, since data is usually sourced from sources that produce or store data regardless of the data mining process, the data needs to be prepared according to the circumstances and the problem, to the appropriate data for injection into data mining algorithms. To prepare data, we need to take it out of its original form and turn it into a form suitable for the algorithm. Also, existing data usually has different redundancies that may cause the algorithm to fail.

Data preprocessing is a data mining technique involving the conversion of raw data into comprehensible data and is in fact an important step in the process of solving machine learning problems. In the real world, data is often incomplete, incompatible, and without certain behaviors or tendencies, and is likely to be accompanied by a large number of errors. Data preprocessing is a proven way to solve these problems. Most machine learning datacenters need to be processed and formatted so that machine learning algorithms can be taught by them. Data preprocessing prepares raw data for subsequent processing.

In some of these data sets, different attributes have different values. In this case, features that have a larger range of values may override features that have a smaller range of values and may be more likely to be selected. To solve this problem, all different data sets are normalized before the feature selection process begins. In this dissertation, nonlinear scaling technique is used to normalize data sets. Using this technique, all data values in the dataset are normalized from zero to one. Normalization of features is done using the following formula:

$$\hat{w}_{ij} = \frac{1}{1+\exp(-\frac{w_{ij}-\bar{w}}{\sigma})} \quad (1)$$

where, w_{ij} is the *j-th* features of *i-th* data is dataset, and $\bar{w}$ and σ show the mean and standard deviation for all data in the dataset, respectively.

Generate neural network

Pattern recognition and classification is one of the most important applications of statistical methods in various sciences. One of the main goals of modeling and classification in statistics is to predict based on existing facts and variables and available information on a specific topic. In statistics, this task is mainly based on methods such as regression, audit analysis, time series, classification, tree regression.

Statistical classical methods for modeling relationships between variables have a number of assumptions and limitations. Consideration of a default distribution such as normal distribution for response variables, linearity of the proposed relationship, uniformity of variance of errors, etc. are among the limitations of classical methods that when using these methods, if real data, conditions Assuming they do not have the model, the use of these methods is not possible or is accompanied by a significant error. In addition, none of these methods have the ability to model complex nonlinear relationships and high-degree interactions. The greater sensitivity of these models to missing observations and discarded data is another limitation of these methods.

Therefore, there is a need for methods that face fewer limitations in this area. Meanwhile, neural network models can be one of the most suitable methods. Artificial neural networks, or more simply neural networks, are new computational systems and methods for machine learning, demonstration of knowledge, and applied knowledge to overstate the output responses of complex systems. The main idea of such networks is inspired by the way the nervous system works, to process data, and information in order to learn and create knowledge. The key element of this idea is to create new structures for the information processing system.

In this paper, a multi-layered neural network will be used to predict and calculate the cost of electricity in a distribution network. Multilayer neural networks are one of the most widely used types of neural networks that have a great ability to solve many problems. One of the most important and widely used of these networks is the Multilayer perceptron neural network, which uses a post-release training algorithm to train the network. In the following, we will briefly introduce the structure and training algorithm of this network. The general structure of a post-diffusion multilayer network is shown in Figure 1.

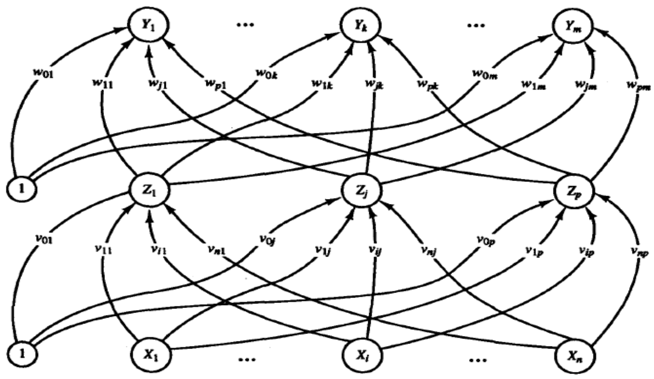

Fig 1. The structure of a multi-layer neural network with a hidden layer

The neural network shown in Figure 1 shows an input layer (X neurons), a hidden layer containing hidden units (Z neurons), and an output layer (Y neurons). As can be seen in this figure, the output units and the hidden units can also have bias. The number of neurons in our input layer in this dissertation will be equal to the number of input data in our problem. For example, if the number of input data in our problem is 5 attributes, the first layer of the neural network will also have 5 neurons. Also in the middle layer we have hidden neurons. In most cases, the prediction of 10 to 20 neurons will be appropriate. In this dissertation, the initial performance of the network was examined in a number of initial repetitions, and in the end, based on the initial answers, it was concluded that 10 neurons in the hidden layer will have the best performance. Also, in the output layer, we will have the number of outputs of the neuron problem. Due to the structure of our problem and its output, the hidden layer will have a neuron.

Learning a perceptron neural network involves three steps: feed forward training pattern, then back forward related error, and adjusting the weights. In the forward step, each X_i input unit receives an input signal and sends this signal to each of the $Z_1,\ldots, Z_p$ units. Each hidden unit then calculates its activation and sends its signal, Z_i, to all output units. Each Y_k output unit calculates its activation (equal to y_k) to form the network response for the input pattern provided.

The first step in learning the neural network is the initial weighting of the grids in the network. Due to the fact that in the MLP neural network learning method, the initial weighting is done randomly, it is possible that if the initial weights are far from the optimal weights, or the convergence time is prolonged, or it is possible to reach the optimal weight at all. Do not provide a place. As a result, primary weights need to be produced non-randomly and purposefully. Therefore, in this dissertation, it is suggested that instead of randomly generating primary weights, an evolutionary optimization algorithms be used for this purpose.

Evolutionary algorithms have advantages over other optimization algorithms that have led to their widespread use. These algorithms, for example, do not require a complete introduction to the problem and can only work with a little information about the definition of the problem. They also have no restrictions on the competency function and do not need to be derivative.

In addition, because evolutionary algorithms have a population of organisms and work in parallel on different parts of the population, they are less likely to be in local optimizations. This feature of evolutionary algorithms allows optimization work to be performed in parallel on multiple sections of the population.

The artificial Bee Colony algorithm (ABC) is an optimization algorithm based on swarm intelligence and intelligent behavior of the bee population that simulates the food search behavior of bee groups. In the early version of this algorithm, it performs a kind of local search that is combined with a random search and can be used for hybrid optimization or functional optimization. This SI-based algorithm has been utilized in many studies to search for the optimization problem.

An important difference between the ABC algorithm and other algorithms based on collective intelligence is that in the ABC algorithm the possible solutions represent the same food sources, whereas in algorithms such as the PSO, each solution may represent an individual from a group.

The implementation of the ABC algorithm to predict and calculate the cost of electricity in a distribution network mainly includes the particle encoding, initial population generation, fitness function calculation, particle position and velocity updating, and stopping criterion. The additional details of these steps are described in the following subsection.

Food Sources encoding:

In the proposed method, each food source in the population represents a candidate solution to the initialization of neural network weights problem. In representing the problem of the initial value of the weights in the neural network, each food sources in the ABC algorithm must specify all the initial weights. As a result, the dimensions of each food sources will be the same as the initial weights of the network.

Initialization:

At this step, F food sources should be generated as the initial population. The initial population generation is essential for the convergence speed and the quality of the solution of the ABC algorithm. During the initialization of ABC, as no prior knowledge is available, the initial food sources are generally determined by random initialization.

Fitness function calculation:

At this stage, the feedforward error in the neural network will be used as the measurement fit of each particle.

Food source updating:

At each step, the position and velocity of each particle are updated. Two criteria are used to do this. The first is the best position the particle has ever reached, known as the pbest, and the second is the best position ever obtained by the particle population. This position is displayed with gbest. In the particle optimization algorithm, each particle is displayed in the search space with the x_i vector. It also has a velocity of v, which is indicated by the vector v_i. In each repeated iteration, the velocity and location of each particle are measured using the following equations.

$$x_{ij} = x_{ij} + \emptyset_{ij}(x_{ij} - x_{kj}) \quad)2($$

Here, $\emptyset_{ij}$ is a random number in the interval [1,1-] that controls the production of the position of neighboring food sources around x_{ij}. In this equation, the variable K is also generated randomly and must be different from i.

According to this equation, as the difference between x_{ij} and x_{kj} decreases, the deviation from the position x_{ij} will also decrease. In fact, in this regard, we try to select one of the dimensions of one of the positions and move towards it or in the opposite direction according to the value of $\emptyset_{ij}$.

Stopping criterion

At this step, the termination condition is checked. If the number of iterations is not more than the maximum allowed iteration, fitness function calculation, food source updating will be repeated. Otherwise, these steps will be stopped, and the global best position reported as the final feature set.

IV. EXPERIMENTAL RESULTS

In this section, four cases are enumerated to verify the performance of the proposed hybrid model. In case 3, we use the latest price data to verify the superiority of the proposed model, and the other two cases are compared with the recently published publication using proposed method. As shown in Table 1.

Table 1. Case study

	Data Sources	Date of price
Case 1	New South Wales	January 2014-December 2014
Case 2	New South Wales	January 2014-December 2014
Case 3	France, New South Wales	September 2017-September 2018

Case 1:

After determining the parameters of the LSTM model, it is crucial to choose a suitable optimizer for it. There are many optimizers available to optimize the LSTM neural network for better performance. Among them, SGD, RMSprop and Adam are more commonly used for time series prediction. To find the optimizer that is most suitable for the LSTM neural network, three optimizers SGD, RMSprop and Adam are used for comparison, and electricity price data for New South Wales from January 2014 to December 2014 is used to demonstrate the fitness of different optimizers for the LSTM model. During the 12 months, the last day of each month data is used as a test set to verify the quality of the model, all data containing 30 days or 29 days or 27 days before the test set is used as the training set. Three loss functions, RMSE, MAE and MAPE, are used to evaluate the accuracy of the LSTM model under different optimizer optimizations for electricity price prediction. The average results are shown in Table 2. As a kind of stochastic gradient-based optimizer, Adam shows better fitness with LSTM than other optimizers.

Table 2. Comparison of different optimizers for LSTM neural network models.

	RMSE	MAE	MAPE		
SGD-LSTM	2.87	2.10	2.34		
RMSProp-LSTM			2.94	2.16	2.41
ABC Neural Network			2.63	1.86	2.21

Case 2:

In this case, electricity price data from New South Wales for 12 months in 2014 is adopted, which is exactly the same as the data used in Ref. [32], in which, a hybrid model, combining wavelet transform, statistical method ARMA and kernel-based extreme learning machine is considered to predict the electricity price on the last day of each month in NewSouthWales 2014. During the 12- month period, the last day of each month data is used as a test set quality of the model, and all data containing 30 days, 29 days or 27 days prior to the test set is used as the training set. Just like what [32] did, MAPE, U1 and U2 is used to enable a fair comparison. As the Case 1 did, the training set is smoothed and standardized before all data is used, and the Adam optimizer is used to optimize the performance of the proposed model in terms of electricity price forecasting. As shown in Table 3, the proposed model is compared to Yang's model [32].

From the results in Table 3, we can see that the 12-month MAPE averages from Yang's model [32] and the proposed model are 3.74, and 2.26, respectively. Compared with Yang's method, the error is reduced by 1.48%. In addition, Table 3 also shows two statistical measures U1 and U2 calculated from the proposed method and the comparison methods. U1 values by proposed methods are close to zero, denoting good prediction accuracy of proposed model, moreover, in all the comparison methods, the proposed model has the smallest average value of U1. In addition, the value of U2 of the proposed model is much smaller than 1. All results show that the proposed method has the best predictive performance.

Table 3. Comparison of different optimizers for LSTM neural network models.

	Proposed method	**Yang's model [32]**
Ave (MAPE)	2.26	3.74
Ave (U1)	0.0145	0.0225
Ave (U2)	0.1845	0.2908

Case 3:

To fully illustrate the outstanding performance of the proposed model in electricity price forecasting, we have compared the proposed model with the recently published publication [34], which used differential evolution (DE) to optimize LSTM models, denoted as DE-LSTM, to predict electricity prices and validate them in the Australian and French electricity markets. Just as what the author of [34] did, the first 100 points of electricity price in France in September 2017 and electricity price of New South Wales in May 2013 are obtained to verify the performance of proposed model.

To enable a fair comparison, as the dataset appeared in Ref. [34], the first 100 points of electricity price in France in September 2017 are divided two subsets: training set (first 80 points) and test set (20 points). Then the dataset is normalized and Adam is selected as the optimizer for LSTM (The training set is not smoothed because there are no outliers in the training set) three loss functions, RMSE, MAE and MAPE, are used to evaluate the accuracy of proposed model and DE-LSTM [34]. The comparison between the proposed method, DE-LSTM [34] and Adam-optimized LSTM is shown in Table 4, and the actual value and predicted price are shown in Fig. 5. As can be seen in Table 4, the author [34] lists different cases based on this dataset for electricity price forecasting, the minimum error of these cases are 3.65% (MAPE), 1.76 (RMSE) and 1.43 (MAE), which is higher than the proposed model by 1.56 (MAPE), 0.68 (RMSE), 0.69 (MAE) respectively.

Table 4. Methods comparison of MAPE (%), RMSE and MAE in France, September 2017.

	Proposed method	**DE-LSTM [46]**	**Adam-optimized LSTM**
Ave (MAPE)	2.09	3.65	3.64
Ave (RMSE)	1.08	1.76	1.67

Ave (MAE)	0.74	1.43	1.38

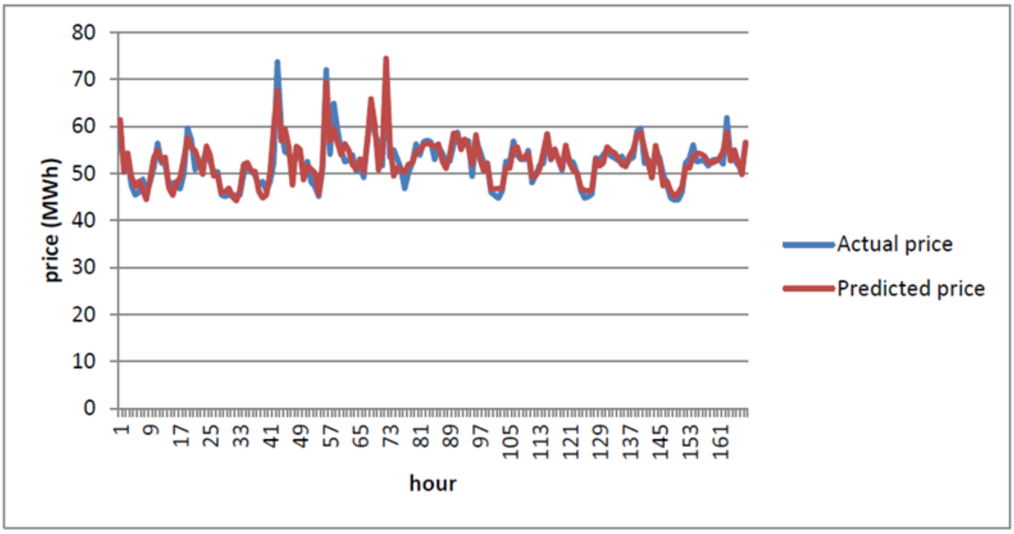

Fig 2. Actual electricity price and predicted price of NSW in May 2013 (24-step-ahead (1 day ahead)).

V. CONCLUSION

In recent decades, the world's electricity industry has undergone many changes. The performance of this industry in different countries is gradually giving way from a monopolistic structure to a competitive one. In some countries, this restructuring has begun with the competitiveness of the manufacturing sector, and in others, such as India, with the competitiveness of the distribution sector, and gradually other sectors of the electricity industry have become competitive. In competitive electricity markets, the price of electricity is a key factor in determining short-term operating plans and tender strategies. So many data-driven methods have been proposed to model electricity market forecasting. Numerous data mining methods have been proposed to predict the cost of an electricity distribution network. One effective way to calculate the cost of electricity is to use neural network-based methods. Due to the fact that in the MLP neural network learning method, the initial weight gain is done randomly, it may be that if the early initial weight is far from the optimal weights, or the convergence time is prolonged, or the optimal weight gain is not at all spatial. As a result, primary weights must be produced non-randomly and purposefully. Therefore, in this dissertation, it is suggested that an artificial bee colony algorithm be used for this purpose instead of randomly generating primary weights. Implementing the proposed method and comparing its performance on three different data sets showed that the accuracy of calculating and predicting the proposed method was superior to other methods and has less error.

VI. REFERENCES:

[1] Chaâbane, N., A hybrid ARFIMA and neural network model for electricity price prediction. International journal of electrical power & energy systems, 2014. **55**: p. 187-194.

[2] Peng, L., et al., Effective long short-term memory with differential evolution algorithm for electricity price prediction. Energy, 2018. **162**: p. 1301-1314.

[3] Ebrahimian, H., et al., The price prediction for the energy market based on a new method. Economic research-Ekonomska istraživanja, 2018. **31**(1): p. 313-337.

[4] Kuo, P.-H. and C.-J. Huang, An electricity price forecasting model by hybrid structured deep neural networks. Sustainability, 2018. **10**(4): p. 1280.

[5] Gholipour Khajeh, M., et al., Electricity price forecasting using neural networks with an improved iterative training algorithm. International Journal of Ambient Energy, 2018. **39**(2): p. 147-158.

[6] Wang, F., et al., Daily pattern prediction based classification modeling approach for day-ahead electricity price forecasting. International Journal of Electrical Power & Energy Systems, 2019. **105**: p. 529-540.

[7] Gao, W., et al., Combination of fuzzy based on a meta-heuristic algorithm to predict electricity price in an electricity markets. Chemical Engineering Research and Design, 2018. **131**: p. 333-345.

[8] Ouyang, T., A. Kusiak, and Y. He, Modeling wind-turbine power curve: A data partitioning and mining approach. Renewable Energy, 2017. **102**: p. 1-8.

[9] Zhou, Y., R. Arghandeh, and C.J. Spanos, Partial knowledge data-driven event detection for power distribution networks. IEEE Transactions on Smart Grid, 2017. **9**(5): p. 5152-5162.

[10] Shao, Z., et al., A new electricity price prediction strategy using mutual information-based SVM-RFE classification. Renewable and Sustainable Energy Reviews, 2017. **70**: p. 330-341.

[11] He K, Yu L, Tang L. Electricity price forecasting with a BED (Bivariate EMD Denoising) methodology. Energy 2015;91:601e9.

[12] Elamin N, Fukushige M. Modeling and forecasting hourly electricity demand by SARIMAX with interactions. Energy 2018;165:257e68.

[13] Singh N, Mohanty SR, Shukla RD. Short term electricity price forecast based on environmentally adapted generalized neuron. Energy 2017;125:127e39.

[14] Lago J, et al. Forecasting day-ahead electricity prices in Europe: the importance of considering market integration. Appl Energy 2018;211:890e903.

[15] Nowotarski J, Weron R. On the importance of the long-term seasonal component in day-ahead electricity price forecasting. Energy Econ 2016;57: 228e35.

[16] Weron R. Electricity price forecasting: a review of the state-of-the-art with a look into the future. Int J Forecast 2014;30(4):1030e81.

[17] Tan Z, et al. Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. Appl Energy 2010;87(11): 3606e10.

[18] Dong Y, et al. Short-term electricity price forecast based on the improved hybrid model. Energy Convers Manag 2011;52(8e9):2987e95.

[19] Panapakidis IP, Dagoumas AS. Day-ahead electricity price forecasting via the application of artificial neural network based models. Appl Energy 2016;172: 132e51.

[20] Shrivastava NA, Panigrahi BK. A hybrid wavelet-ELM based short term price forecasting for electricity markets. Int J Electr Power Energy Syst 2014;55: 41e50.

[21] Yan X, Chowdhury NA. Mid-term electricity market clearing price forecasting: a multiple SVM approach. Int J Electr Power Energy Syst 2014;58:206e14.

[22] Kim HY, Won CH. Forecasting the volatility of stock price index: a hybrid model integrating LSTM with multiple GARCH-type models. Expert Syst Appl 2018;103:25e37.

[23] Liu H, Mi X-w, Li Y-f. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. Energy Convers Manag 2018;156: 498e514.

[24] Peng L, et al. Effective long short-term memory with differential evolution algorithm for electricity price prediction. Energy 2018;162:1301e14.

[25] Duchi J, Hazan E, Singer Y. Adaptive subgradient methods for online learning and stochastic optimization. J Mach Learn Res 2011;12(Jul):2121e59.

[26] Tieleman T, Hinton G. Lecture 6.5-RMSProp, COURSERA: neural networks for machine learning. University of Toronto, Technical Report; 2012.

[27] Peng Y, et al. Chemical-protein relation extraction with ensembles of SVM, CNN, and RNN models. 2018. arXiv preprint arXiv:1802.01255.

[28] Bansal T, Belanger D, McCallum A. Ask the gru: multi-task learning for deep text recommendations. In: Proceedings of the 10th ACM conference on recommender systems. ACM; 2016.

[29] Arik SO, et al. Convolutional recurrent neural networks for small-footprint keyword spotting. 2017. arXiv preprint arXiv:1703.05390.

[30] Reddy SS, Jung C-M. Short-term load forecasting using artificial neural networks and wavelet transform. Int J Appl Eng Res 2016;11(19):9831e6.

[31] Mandal P, et al. A novel hybrid approach using wavelet, firefly algorithm, and fuzzy ARTMAP for day-ahead electricity price forecasting. IEEE Trans Power Syst 2013;28(2):1041e51.

[32] 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;190:291e305.

[33] Shayeghi H, Ghasemi A. Day-ahead electricity prices forecasting by a modified CGSA technique and hybrid WT in LSSVM based scheme. Energy Convers Manag 2013;74:482e91.

[34] Peng L, Liu S, Liu R, et al. Effective long short-term memory with differential evolution algorithm for electricity price prediction. Energy 2018;162:1301e14.