A New Electricity Price and Load Uncertainty Prediction Method based on Optimal Neural Networks for Deregulated Electricity Power Markets

FARSHID KEYNIA[1] **AND MOHAMMAD BAHRAMPOUR**[2]

[1,2] *Department of Energy, Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran.*
[*] *Corresponding author: keynia@yahoo.com*

Manuscript received June 13, 2017; revised November 4: accepted November 4, 2017. Paper no. JEMT1706-1020

In recent years, Short-term load and price forecast has always been a key issue for power system operation. In regulated power systems, Short-term load forecast is an important tool for reliable and economic operation of power systems. Many operating decisions are based on Short-term load forecasting, such as dispatch scheduling of generating capacity, reliability analysis, security assessment and maintenance plan for the generators. On the other hand, electricity price variation is more important and effective factors for all power market participants. Bidding strategy, risk control, investment decisions, demand and supply balancing and power system reliability and other power market applications are highly depended to load and price uncertainty. In this paper a new intelligent hybrid method has been proposed to price and load uncertainty prediction. The proposed method consists of an improved version of particle swarm optimization algorithm to fine tuning the main predictor system's adjustable parameters. The price and load variation intervals have been predicted by predictor system based on multi-layer neural networks. The proposed method has been examined in some well-known power markets.

keywords: *Prediction intervals, Price and Load Uncertainty, Mutual Information Feature Selection, Electricity Load and Price, Particle Swarm Optimization.*

http://dx.doi.org/10.22109/jemt.2017.89366.1020

1. INTRODUCTION

Load forecasting is a vital task for power system operation. The balance between actual demand and generated power is an important factor for the reliability of power system, and load forecasting helps this balance. Short term load forecasting (STLF) is one of the aspects of load forecasting which is necessary for generation planning, maintenance scheduling and security analysis. Various methods has been presented for STLF including: regression-based approach [1], ARIMA models [2], fuzzy systems [3], bagged neural networks (BNNs) [4], Bayesian neural network (BNN) [5], support vector machine [6], and also combined methods such as in [7] that authors have used an extreme learning machine (ELM) for short term load forecasting, the model parameters are optimized by a modified artificial bee colony (MABC) algorithm, they also have used wavelet transform to decompose the load signal. In [8] a load forecasting method based on the combination of support vector regression and recurrent neural network has been proposed, the optimal parameters of the model is obtained by a bee colony algorithm. In [9] the electricity load is forecasted by support vector machine (SVM), and ant colony optimization is utilized as a feature selection technique. Authors in [10] have forecasted the electricity load by support vector machine, the model parameters are optimized by simulated annealing optimization algorithm.

Price forecasting is an important issue for both market participants and market operators in wholesale electricity markets. There are two markets for electricity trading: pool based market and bilateral market. In both of them price forecasting is an essential requirement. Short-term price forecasting can help GenCos for bidding strategies to maximize their profit. Consumers can minimize their utilization cost by price forecasting. Time series and intelligence systems are the most methods used for price forecasting. Typical time series and intelligence system methods include auto-regressive integrated moving average (ARIMA) [11,12], generalized auto-regressive conditional heteroskedastic (GARCH) [13–15], neural networks [16,17], support vector machine (SVM) [18], and a new hybrid approach combines ARIMA with a cascaded NN models [19]. As mentioned above, accurate load and price forecasting are the important issues in deregulated power markets. But all the methods have been presented so far have focused on point forecasting. It means that they forecast the numerical values of electricity price or load for a given forecast horizon. Point forecasts don't have any probability of correct predictions of the objective function and also they can't handle the data's uncertainty correctly [20]. Point forecasts have no more information about sample errors and prediction accuracy. Renewable energies are sources of uncertainty in power system. The uncertainty existence in data causes the reduction of prediction accuracy. Prediction intervals (PIs) are common tools for quantifying the uncertainty levels related to the price or load forecast and evaluating decision risk for participants. PIs include the actual price or load with a predefined probability named confidence level. They predict the distribution of future prices and loads. Sometimes participants are more attribute to know the intervals for future samples instead of points. PIs are useful for bidding strategy, risk control and investment decisions. They are also useful for applications which are needed to balance the demand and supply such as power markets. In pool based power markets ISO forecasts the demand and matches the suppliers bidding with this forecast, if it over predicts the demand, the operating cost increases, and if it under predicts the demand, the reliability of supply will compromise. Therefore supply optimization is the utilities needed. If they knew the intervals they can improve the bidding strategies and investment decisions. As literatures show, neural networks have been used widely for forecasting problems and also they indicate an acceptable performance. Some of the commonly known techniques construct PIs based on NNs are Bayesian [21], delta [22], bootstrap [23]. These methods, first make point forecasts and then construct PIs based on this forecast. In [24] a novel hybrid method has been introduced to construct high quality PIs. In this ref. A Generalized Auto Regressive Conditional Heteroscedasticity (GARCH) model is used for PIs construction and the parameters of model have been optimized based on a single objective function (CWC) and SA as an optimization algorithm is used. Recently a new method based on neural networks, named LUBE (lower, upper bound estimation), has been presented in [25] by Khosravi et al. This method constructs PIs directly and also has less computational burden than traditional methods. The main concept of LUBE method is to train an NN with two outputs to directly construct the upper and lower bound of PI.

From the feasibility view of constructing PIs, intervals should cover the main target with high probability and also have width as narrow as possible. These two aspects are conflicting. It means that the high coverage probability is obtained by wide PIs and narrow PIs may not have a large coverage probability. In [25] this initial multi objective optimization problem has become a single objective problem with one cost function called CWC. Ref. [25] minimizes this objective function with simulated annealing (SA) for real-world case studies. In [26,27] minimization was performed by particle swarm optimization (PSO) for six synthetic and real-world case studies. In [28–30] a new constrained single objective problem has been proposed and PIs has constructed for short term load. Authors in [31] extended LUBE method and used mutual information as feature selection for selecting neural network inputs. In this Ref. a single stage feature selection is used. Ref. [32] generates the lower and upper bounds of the future electricity prices by using support vector machines (SVM), in this ref. model parameters were obtained by minimization CWC using PSO.

In this paper, we are going to construct PIs for electricity load and price with LUBE (lower upper bound estimation) method.

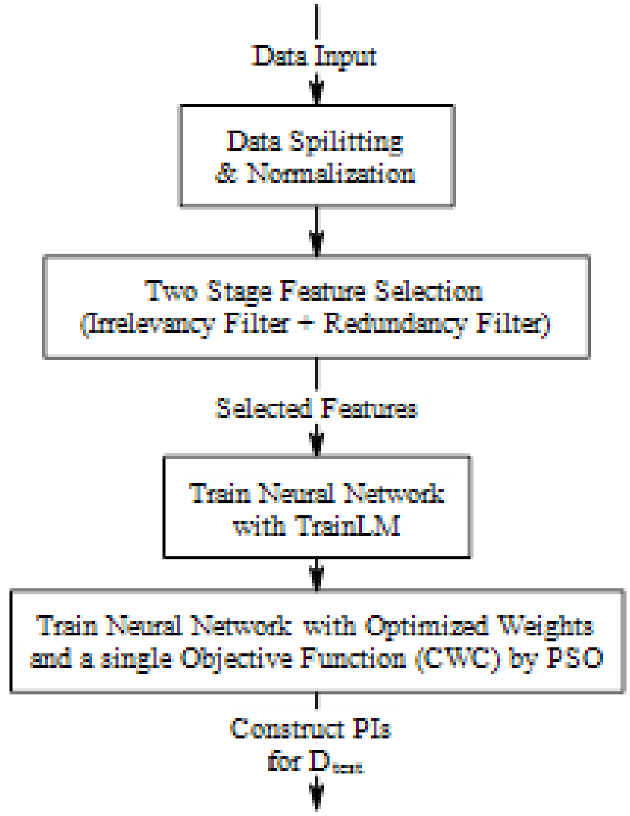

Fig. 1. overall procedure of the proposed method to construct PIs

We minimize the CWC with PSO, which is a dominant evolutionary algorithm. In all references mentioned above, the evolutionary algorithms start from a random point, but we first train an NN with two outputs, which are repeating the target. It means that the target, which could be either load or price, repeats again and it will have two rows with the same values. Then, with this target and also a train matrix, an NN will be trained by an analytical algorithm, levenberg-Marquad (trainLM). Trained network parameters, weights and biases, taken and all PSO particles are filled with these weights and biases. Then PSO starts the optimization from this point. Fig.1 shows the overall procedure of our proposed method for PIs construction. Once the network is trained by trainLM, it gives two outputs that are equal to upper and lower bounds respectively, at this stage the coverage probability is near zero, because the upper and lower bounds are the same as each other and the width of upper lower bound is too narrow. The evolutionary algorithm makes the bound wider and increases the coverage probability as the width of bound is narrow. This task helps the algorithm to achieve a better performance and obtains a better optimal point. This paper is the first one which constructs PIs for loads and prices with a new intelligent hybrid method and minimization of single objective function.

The method contributions can be summarized as:

1) a two stage feature selection method is used, first stage for

refining the irrelevant features, and the second one for redundant features. Mutual information as a suitable tool for feature selection is also used.

2) A new cost function is defined to construct PIs for electricity price and load values.

3) A new hybrid particle swarm optimization and neural network based method has been proposed to predict best PI values.

4) The obtained results show that proposed method outperforms all of the well-known methods in this field.

The remaining of the paper is as follows. In section 2 PIs indices are described. The cost function is formulated in section 3. In section 4 the stages of the proposed algorithm are explained. Section 5 discusses case studies and implementation of the algorithm. Numerical results and discussions are presented in section 6. Finally the paper concluded in section7.

2. PREDICTION INTERVAL ASSESSMENT INDICES

By definition, future observations should lie within the upper and lower bounds with a predefined probability called the confidence level $((1-\alpha)\%)$. In literatures, two indices have been introduced for PIs evaluation, which are based on coverage probability and the width of intervals.

A. PICP (Prediction Interval Coverage Probability)

Prediction interval coverage probability (PICP) is the main index for evaluation of PIs. It indicates that what percentage of the target is covered by upper lower bounds. It is defined as follows:

$$PICP = \frac{1}{n}\sum_{i=1}^{n} c_i \tag{1}$$

where n is the number of samples, and ci=1 if the target y_i lies between the upper bound U_i and lower bound L_i of the ith PI; otherwise $c_i = 0$, [25]. If all of the target values lie within the upper and lower bounds, PICP is 100%. To have a valid PIs, PICP should be equal or greater than the nominal confidence level of PIs $((1-\alpha)\%)$. The best case for PICP is $PICP = 100\%$, which means that all of the target values lie within the PIs.

B. PINAW (Prediction Interval Normalized Average Width) and PINRW (Prediction Interval Normalized Root-mean-square width)

If the upper and lower bound of PIs are too wide the PICP is high, but these PIs don't have more information about the target. Therefore the width of PIs should be considered. PINAW formulated this aspect of PIs [26]:

$$PINAW = \frac{1}{nR}\sum_{i=1}^{n}(U_i - L_i) \tag{2}$$

where R is the difference of maximum and minimum of the target values. Originally this index is prediction interval average width $(PIAW)$, but for comparing goals it is normalized by dividing to R.

In point forecast problems there are two main indices which are mean square errors (MSE) and mean absolute percentage errors (MAPE). MSE is used usually as the cost function in learning algorithms such as back propagation neural networks, and MAPE is used for performance evaluation of the trained network. That is because MAPE gives equal weights to forecasting errors, but MSE magnifies bigger errors and so the train stage will perform better. In contrast to this, we use PINRW index for training and PINAW for testing. The PINRW index is defined as follows [26]:

$$PINRW = \frac{1}{R}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(U_i - L_i)^2} \tag{3}$$

The format of PINAW is similar to MAPE in point forecast while PINRW is similar to MSE in point forecast. In the rest of this paper, PINRW is used for training NN and PINAW is used for testing.

C. SCORE

Another index for evaluating the quality of PIs, proposed in [33] by Winkler, which is a scoring rule for interval prediction. This score rewards narrow PIs and gives a penalty if target values y_i is not covered by PIs. This score is calculated as follows:

$$S_i = \begin{cases} -2\alpha v_i - 4(L_i - y_i); y_i < L_i \\ -2\alpha v_i; y_i \in [L_i, U_i] \\ -2\alpha v_i - 4(y_i - L_i); y_i > U_i \end{cases} \tag{4}$$

where α is related to the confidence level ($\alpha = 0.1$ for 90% nominal confidence level) and v_i is the size of the intervals:

$$v_i = U_i - L_i \tag{5}$$

The average score indicates overall performance:

$$\bar{S} = \frac{1}{n}\sum_{i=1}^{n} S_i \tag{6}$$

As is clear, the score in (6) is negative. The score with lower absolute value indicates better performance of PIs.

3. COST FUNCTIONCOST FUNCTION

As mentioned before, the PI construction problem is a multi-objective optimization problem. From an optimization perspective, PIs with high coverage probability and narrow width is suitable. The primary problem can be formulated as follows [26]:

Objectives: Finding optimal weights w^* to:

Maximize: $PICP(w)$;

Minimize: $PINAW(w)$.

Constraints: $PINAW(w) > 0$;

$\mu \leq PICP(w) \leq 1$.

where μ is the nominal confidence level related to PIs and can set to $(1-\alpha)$, in this paper in training stage it set to a value that is a little above $(1-\alpha)$ and in testing it set to $(1-\alpha)$. As the above problem indicates, two objectives are conflicting, improving one may decline another. A PI-based cost function has been introduced in [25] and improved in [26]. It is a single objective function that combines two objectives (PICP and PINAW), and gives a high penalty to cost when PICP is less than μ, and when PICP is equal or more than μ the exponential term becomes zero and the algorithm focuses on PINAW minimization. This coverage width-based criterion (CWC) is defined as follows:

$$CWC = PINAW * (1 + \gamma(PICP)e^{-\eta(PICP-\mu)}) \tag{7}$$

where $\gamma(PICP) = 1$ in the training stage. For testing $\gamma(PICP)$ is a step function which depends on PICP value. If PICP is equal or more than μ, this function is zero and CWC equals to PINAW, otherwise this function is one and make CWC high.

$$\gamma(PICP) = \begin{cases} 1; PICP < \mu \\ 0; PICP \geq \mu \end{cases} \tag{8}$$

η and μ are two constant parameters that indicate how much penalty is related to PIs with low coverage probability. μ is same as mentioned above and η exponentially magnifies the difference between PICP and μ, by adjusting it to a large value (10-100) we give high penalties to the cost function when PICP is less than μ. At the start of the optimization, because the network is already trained, PICP is near to zero, so the penalty factor is high, after a few iterations, when PICP is equal to μ the penalty factor becomes zero and the algorithm focuses on minimization of PINAW. In the training PINRW is utilized instead of PINAW and in test PINAW.

4. CONSTRUCT PIS FOR SHORT TERM LOAD AND PRICE FORECASTING

The main method for construction of PIs in this paper is LUBE (lower, upper bound estimation) method. In this method one neural network with two outputs will be trained to construct PIs directly in one step [25]. The first NN output corresponds to upper bound and the second one corresponds to lower bound. The implementation of this method is easy and has less computational burden than traditional methods [34].

A. Implementation

The flowchart in Fig.2 explains the proposed method which is used in this paper.

A.1. Data splitting

The total dataset is split into two sets: training set (D_{train}), and test set (D_{test}). Then training set is normalized to [11], and the same setting is applied to test set for normalization.

A.2. Feature Selection (Irrelevancy and Redundancy Filter)

In [35, 36] discussed that price signal has daily and weekly periods, so for considering this issue, lagged values until 200 hours ago is considered. Also for the price, exogenous data, including load are considered, but for load no such data are considered. Therefore, price train matrix has 400 rows (200 hours ago of price and 200 hours ago of load), and the load train matrix has 200 rows. Input data sets candidate for price and load are as follows:

$$\begin{aligned} S_p &= [P(t-1), P(t-2), ..., P(t-200), \\ &L(t-1), L(t-2), ..., L(t-200)] \\ S_l &= [L(t-1), L(t-2), ..., L(t-200)] \end{aligned} \tag{9}$$

Network training with these matrices takes the high computational burden and time, while the existence of irrelevant and redundant inputs increases forecasting error. Author in [37] has been introduced a two stage feature selection with mutual information for electricity price forecasting. Here this method is used to reduce the size of these matrices and selects the best inputs for training network.

A.3. Determine NN optimal structure and parameters

For construction of the forecast method, we simply begin with a MLP neural network due to its flexibility as a nonlinear predictor and ease of implementation. According to Kolmogorov's theorem, the MLP can solve a problem by using one hidden layer,

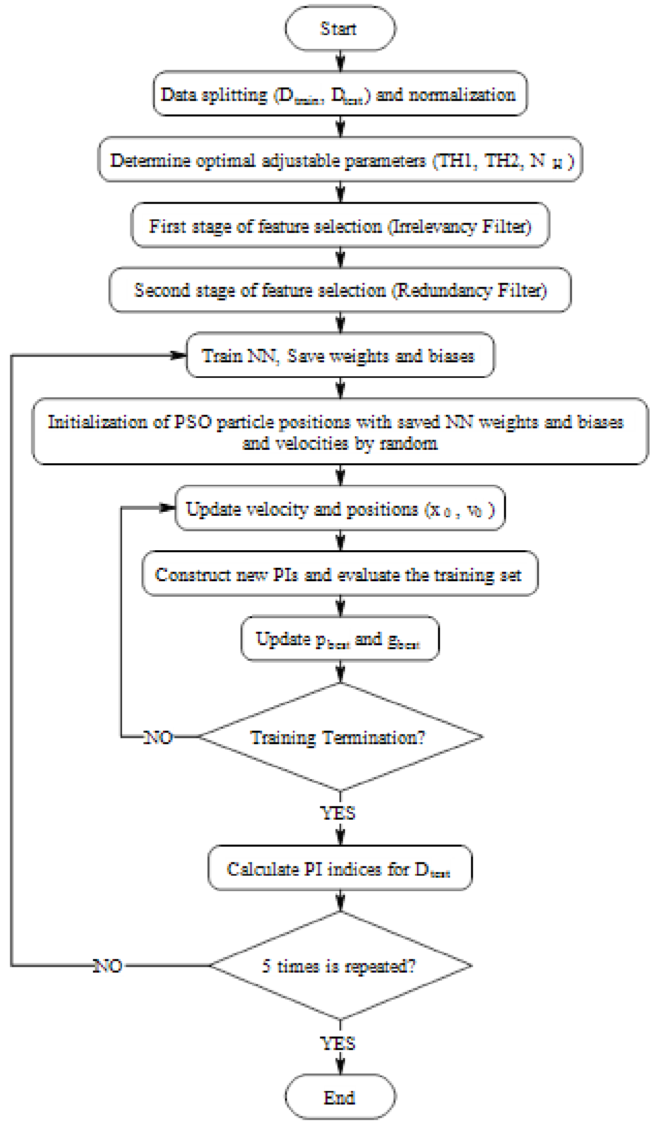

Fig. 2. The flowchart of the proposed method

provided it has the proper number of neurons [45]. So, one hidden layer has been considered in the MLP structure of the NN. Here a multi-layer perceptron (MLP) structure with one hidden layer is used. After determination of input candidates by feature selection technique, the neurons in hidden layer should be determined. To determine the neurons of hidden layer and also the thresholds for the last step, which are mentioned in ref. [37], a cross-validation mechanism which is introduced in [37] is applied. Number of neurons in input layer is determined by inputs which are come from feature selection technique, and number of neurons in the output layer is two, one for upper and the other for lower bound.

A.4. Train NN

Training the NN is performed by an analytical algorithm early. Here Levenberg Marquardt (LM) learning algorithm is chosen. This algorithm is more appropriate for prediction tasks and it trains the neural network faster than the usual gradient descent back propagation methods [38]. A network with two outputs is required. For this purpose, both outputs are set to actual load or price and this matrix is used for network training. After training, the weights and biases are stored in a vector. In this step, two outputs that are related to upper and lower bounds are extremely narrow and the coverage probability of PIs is near zero. Then weights of this trained network are given to an evolutionary algorithm such as PSO to train the network and give a valid PICP.

A.5. PSO parameters initialization

In PSO, there are two parameters that must be initialized at the beginning: position and velocity of each particle. For positions the weights and biases stored in the last step, is utilized. All particles positions are initialized by the stored vector of weights. For velocity random initialization is used.

A.6. Update velocity and positions

Position and velocity update is the fundamental step in PSO algorithm. By this step, useful information of particles is exchanged. Equations (10) and (11) [39] indicate the classic update process of particle velocity and positions:

$$v_i^n(t+1) = W * v_i^n(t) + C_1 * rand(p^i_{best,n} - x^i_n(t)) \\ +C_2 * rand(g_{best,n} - x^i_n(t)) \quad (10)$$

$$x^i_n(t+1) = x^i_n(t) + v^i_n(t+1) \quad (11)$$

where $v_i^n(t)$ and $x^i_n(t)$ are the nth dimensional of ith particle velocity and position in iteration t respectively. W is the inertia weight that decreases here linearly from W_{max} to W_{min}, rand is a random value is generated between [0, 1], C_1 and C_2 are self and social recognition coefficients respectively, pbesti is the best experiment of ith particle until iteration t, gbest is the best experiment of all particles in the population until iteration t [40]. There are some boundary limitations that apply in the updating process, such as limitation of velocity and position that are V_{max} and X_{max} respectively.

A.7. Construct new PIs and evaluate the training set

After updating the position of particles, these positions are taken, which are the weights and biases of the network, to construct the network and the LUBE method is applied for constructing PIs. Then PIs indices such as PICP, PINRW and CWC are calculated by equations (1), (3) and (5), where $\gamma(PICP)$ is set to 1 for training. These indices are further used for p_{best} and g_{best} updating.

A.8. Updating p_{best} and g_{best}

The cost function is CWC, and the goal is to minimize it. For this purpose, the new CWC of each particle is calculated, if the $CWC_{new} < CWC_{pbest}$, then pbest is updated; further if $CWC_{pbest} < CWC_{gbest}$, then g_{best} is updated.

A.9. Training termination

If the maximum number of iterations reached, or the gbest doesn't improve for certain iterations, the training process terminates, otherwise it goes to step 4.1.6. The maximum number of iterations is different in case studies, which based on decreasing and the converging process of CWC. In this paper 1000 is used for maximum number of iterations, but the algorithm doesn't reach this number of iterations anytime, because it terminate when g_{best} doesn't improve for twenty iterations.

A.10. Calculate PIs indices for D_{test}

After training, the gbest position is used to construct PIs for the test set. As mentioned before, in test phase PINAW is used instead of PINRW and $\gamma(PICP)$ is a step function defined in Eq. (6). The evaluation indices are PICP, PINAW, CWC and SCORE that are measured and recorded. This process repeated for 5 times.

5. CASE STUDIES

A. Datasets

For evaluating the performance of the proposed method, case studies from real world electricity markets has been considered. In the case of electricity load data sets from Texas (TX) (USA) [41], PJM [42] and New South Wales (NSW) (Australia) [43] and in the case of price, data sets from PJM [42], Ontario [44] and Victorian electricity Market (VIC) (Australia) [43] are used. These datasets have been chosen to compare the results with references [24, 28–30, 32]. Experiments have been performed for four month data sets of year 2010 for each of the PJM and Ontario markets. The four months are January, April, July, and October. The last week of each month has been separated as test set. In the case of TX, 23rd to 29th, September, 2011 is the test week. For NSW two weeks of 2011, 21st to 27th March and 9th to 15th April are the test weeks. For VIC price three months, May, August and December are the test sets. Historical time lags as mentioned in section 4.1.2 are 200 hours and they are refined by two stage mutual information technique. The data corresponds to 50 historical days before the test week is used to train the neural network; it means 50*24=1200 hours before test week are built the training samples.

B. Feature selection and determination of adjustable parameters

Because of the large scale of training matrices (400*1200 for price, 200*1200 for load), input filter and best ones selection are required. Feature selection methods like mutual information are appropriate to this end. Here two thresholds (TH1 and TH2) are required for two stage feature selection [37]. Another adjustable parameter that should be determined is the number of neurons in the hidden layer of the neural network. In [37] a cross-validation technique is introduced to tune these parameters. The first training dataset is separated into two sets, training and validation sets. The last week of training set or the week before test set is utilized as validation set. It means that from 1200 training sample (50 days before test week), 43 days (1032 samples) is for training and last 7 days (168 samples) is for validation. The PIs construction algorithm is executed for 5 times with a set of candidate of adjustable parameters and median validation error (median of the CWC) is observed. Note that in each execution of the algorithm valid indices for PIs, means $\mu_{nominal} \leq PICP$ and $PINAW > 0$, should be obtained. The set of parameters that give a minimum CWC for validation set, are selected.

For an example, Figs. 3 and 4 show this process for a sample week, which is the week of April 2010 in the Ontario electricity market. Assume we would construct PIs for the price of the test week of April 2010 in the Ontario market. First the adjustable parameters should be determined $(TH1, TH2andNH)$, for this purpose the training set which is from 5^{th} March to 23^{rd} April, is separated into two subsets, training and validation. As mentioned above, 5^{th} March to 16^{th} April is for training and 17^{th} to 23^{rd} April is for validation. In the beginning $TH1$ and $TH2$ are

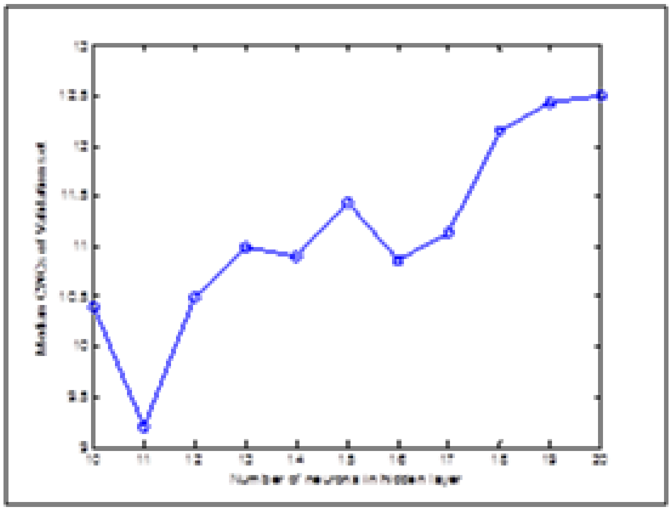

Fig. 3. curve of validation error (CWC) with respect to $NH(TH1 = 0.07, TH2 = 0.58)$

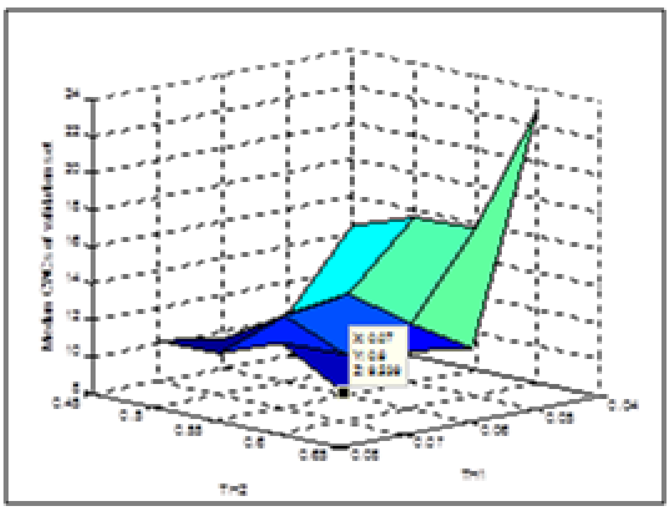

Fig. 4. curve of validation error (CWC) with respect to$TH1andTH2(NH = 11)$

fixed to 0.07 and 0.58 and NH is changed from 10 to 20 for this example. For each value of NH the algorithm in Fig.1 is executed for 5 times and median CWCs of these times are recorded as the validation error. Fig.3fig.3 shows this process for determination of NH. In this Fig the horizontal axis is the values of NH and the vertical one is median CWCs. Consequently, 11 neurons in hidden layer is a good choice. Then the NH is fixed and $TH1$ and $TH2$ are changed. For a set of $TH1$ and $TH2$, again the algorithm is executed for 5 times and median CWCs of each run are recorded as a validation error. Fig.3 shows this process.

In Fig. 4, the number of neurons in the hidden layer is fixed ($NH = 11$) and $TH1$ and $TH2$ are varied. It indicates that the best values for TH1 and TH2 are near 0.07 and 0.6. The minimum validation error or CWC for validation set in Figs 3 and 4 is about 9.539%. Note that CWC here is in percentage. Because the results are recorded for validation set when the value of PICP is more than the nominal probability (90%), and when PICP is more than 90% the (PICP) is a step function and it is set to 1, so, CWC is equal to PINAW, because PINAW is between [0, 1] so CWC is between [0, 1] too, therefore it reports in percentage in Figs. 3 and 4.

Table 1. Obtained features from two stage feature selection

Rank	Selected feature	Rank	Selected feature
1	$P(t-1)$	11	$L(t-24)$
2	$P(t-2)$	12	$L(t-168)$
3	$P(t-24)$	13	$L(t-167)$
4	$P(t-25)$	14	$P(t-48)$
5	$P(t-3)$	15	$P(t-5)$
6	$L(t-1)$	16	$P(t-144)$
7	$P(t-23)$	17	$L(t-169)$
8	$P(t-22)$	18	$P(t-47)$
9	$P(t-26)$	19	$P(t-49)$
10	$P(t-4)$	20	$P(t-168)$

6. NUMERICAL RESULTS AND DISCUSSIONS

A. Training process

By determination of adjustable parameters, the training process is running and PIs are constructing for the test set. This process is executed for 5 times to confirm the results. Figs.4 and 5 indicate the training process. They show the variation of the PICP and PINRW index of gbest in the PSO iterations with respect to the CWC. As these Figs show, in the beginning of the optimization, PICP is above 0.95 and it is more than the $\mu_{train} = 0.93$, so the algorithm focuses on minimization of the PINRW. The aim of the algorithm is to minimize the CWC, therefore it minimizes the PICP, because when PICP is equal to μ_{train} the penalty factor is zero, here PICP is more than µtrain and the penalty factor exist, so the algorithm minimizes the PICP and PINRW together. The variation of the PICP and PINRW is the same, when PINRW decreases a little, PIs become narrow and coverage probability decreases, so PICP decreases. After variation of PICP and PINRW due to minimize CWC, indices come to a steady point and CWC becomes stable. As mentioned before, the algorithm terminates in 90 iterations due to gbest hasn't been improved for twenty iterations. By determination of optimal weights, indices are calculated for test samples. Note that the first iteration that is shown in the Figs is after a perturbation in velocity of particles, in the initialization (before first iteration) the PICP is too low (near zero) and PINRW is also near zero, which is because of the NN training at first. In the training process PINRW index is used instead of PINAW, and PINAW for test samples. It is like that utilizing MSE for training process in point forecast and MAPE for test. Tables 2 display the CWC and PSO parameters. In training process µtrain is used which is more than the $\mu_{nominal}$ and the PICP of the gbets particle in PSO pay more attention to this index. This action ensures that the PICP of test samples is above the nominal μ. The maximum number of iterations in PSO is 1000, but if the gbets cost function (CWC) not improves for 20 iterations, the training process terminates.

Obtained results in each run, median of them and also comparison between them and references are displayed in tables 4 to 8. For a better comparison with other references, the improve-

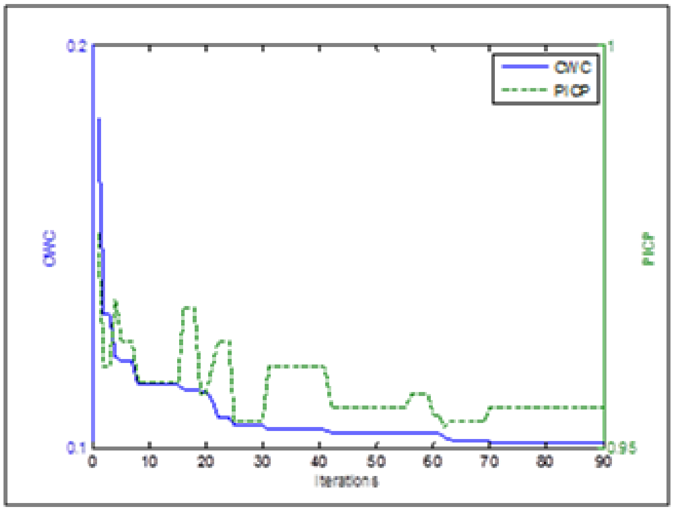

Fig. 5. variation of CWC and PICP for g_{best} with PSO iterations

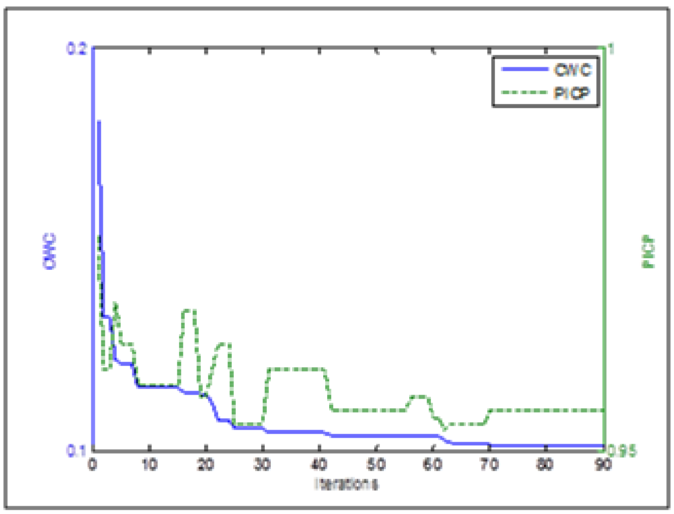

Fig. 6. Variation of CWC and PINRW for gbest with PSO iterations

ment percentage is formulated as follows [29]:

$$\frac{\text{ComparedResults} - \text{NewResults}}{\text{ComparedResults}} * 100\% \qquad \textbf{(12)}$$

PIs are constructed for one week of Texas in [29], one week of NSW in [28] and another week of NSW in [30] with PSO+LUBE technique and are compared with ARIMA, ES and Naïve techniques. The obtained results with proposed technique and the results of these references are shown in table 4. The standard deviation of CWC for TX, NSW (Mar.) and NSW (Apr.) are 0.4999, 0.2603 and 0.2165 respectively, that show the stability of the proposed method. As the stability of the results are obvious (low standard deviations), proposed method show a good improvement. In the case of Texas in [29] the PSO+LUBE is the best technique, if we consider median of CWC as the comparative index, proposed method indicates 81.44%, 67.87% and 67.07% improvement in cases of Texas, NSW (March) and NSW (April) respectively. In the case of Texas in [29] the SCORE index is absolute, but it doesn't absolute here, in cases of NSWs they

Table 2. Parameters that utilize in this paper

	Parameters	Numerical values
PSO	C_1	1.49
	C_2	1.49
	W_{max}	0.7
	W_{min}	0.1
	Num. Particles	50
CWC	α	0.1
	$\mu_{nominal}$	90 %
	μ_{train}	93 %
	η	90

Table 3. PI assessment indices for STLF, 23rd to 29th Sept., 2011 in Texas and NSW electricity markets

~~proposed~~	TX Load				NSW load (21^{st} to 27^{th}, Mar)				NSW load (9th to 15th, Apr)			
	PICP(%)	PINAW(%)	CWC(%)	SCORE	PICP(%)	PINAW(%)	CWC(%)	SCORE	PICP(%)	PINAW(%)	CWC(%)	SCORE
1	92.26	7.15	7.15	-686.91	96.43	7.78	7.78	-138.43	94.94	8.96	8.96	-108.55
2	92.26	7.31	7.31	-727.92	96.43	7.58	7.58	-140.3	94.05	8.58	8.58	-103.39
3	92.26	6.19	6.19	-566.53	96.73	7.23	7.23	-130.26	94.35	8.45	8.45	-111.81
4	91.07	6.93	6.93	-741.42	97.92	7.35	7.35	-129.53	95.54	8.77	8.77	-99.44
5	92.86	6.31	6.31	-599.18	97.92	7.82	7.82	-136.85	95.24	8.46	8.46	-103.06
Median	92.14	6.78	6.78	-664.39	97.08	7.55	7.55	-135.07	94.82	8.64	8.64	-105.25
PSO+LUBE	90.81	36.53	36.53	4725.06	90.34	23.50	23.50	—	91.28	26.24	26.24	—
ARIMA	84.26	28.63	1794.11	4553.45	88.80	20.79	58.74	—	—	—	—	—
ES	85.80	29.94	846.93	4769.64	88.75	22.02	63.15	—	—	—	—	—
Naïve	84.41	32.12	1671.34	5296.59	88.65	24.24	71.92	—	—	—	—	—

don't report SCORE index.

PIs are constructed for the load of PJM electricity market in four weeks of four months, and the obtained results are reported in table 5. The standard deviations of CWC for these weeks are 0.5462, 0.4512, 0.4613, and 0.4588 respectively. As the results show April has better indices, the lower absolute of median SCORE in this week. This is because that this week has a more stable signal than the other weeks. Fig.5 shows this test week signal and the upper, lower bounds which are constructed by the proposed method. Also Fig. 6. shows the bounds (PIs) for this test week and the test points. The heights of intervals confirm the results.

In [24] PIs are constructed for the VIC electricity market for three months with a hybrid method named bootstrap-GARCH and the results compared with delta technique. PIs are also constructed for electricity prices for these months. The obtained results compared with the mentioned reference and are shown in table 5. The standard deviations of CWC for these three months are 0.2965, 0.3817, and 0.277 respectively. Moreover the stability of the proposed method, the comparison of median CWCs with the bootstrap-GARCH method, proposed method indicate 84.03%, 77.23% and 87.15% improvement in May, August and December respectively.

PIs are constructed for four weeks of four months in PJM and Ontario electricity market for price forecasting. In [32] for the test week of January in Ontario the results with the SVM-PSO method compared with several methods. These results and the results of the proposed method for this week are reported in table 6. It shows a better performance of the proposed method. Also for four weeks the results of the proposed method in PJM and Ontario are compared with the method in [32]. For PJM market the results are reported in table 7. The standard deviations of CWC for four weeks are 0.63, 0.3701, 0.5585 and 0.3989

Table 4. PI assessment indices for test samples of the load for PJM electricity market

proposed		PICP (%)	PINAW (%)	CWC (%)	SCORE
January	1	92.26	9.94	9.94	-498.38
	2	92.26	9.47	9.47	-466.85
	3	92.26	10.78	10.78	-488.53
	4	91.07	9.88	9.88	-516.68
	5	92.26	10.62	10.62	-468.24
	Median	92.02	10.14	10.14	-487.73
April	1	97.02	9.18	9.18	-336.38
	2	97.02	9.18	9.18	-332.78
	3	97.02	8.32	8.32	-309.61
	4	96.43	9.52	9.52	-370.24
	5	97.02	8.86	8.86	-351.67
	Median	96.9	9.01	9.01	-340.14
July	1	92.26	9.29	9.29	-875.2
	2	92.86	8.44	8.44	-762.97
	3	92.26	8.93	8.93	-754.24
	4	93.45	9.56	9.56	-843.61
	5	92.26	8.63	8.63	-762.31
	Median	92.62	8.97	8.97	-799.67
October	1	96.43	9.93	9.93	-589.33
	2	97.62	9.3	9.3	-527.98
	3	97.62	8.99	8.99	-533.88
	4	97.62	9.11	9.11	-547.02
	5	97.02	8.7	8.7	-504.73
	Median	97.26	9.21	9.21	-540.59
Average of the year	94.7	9.33	9.33	-542.03	

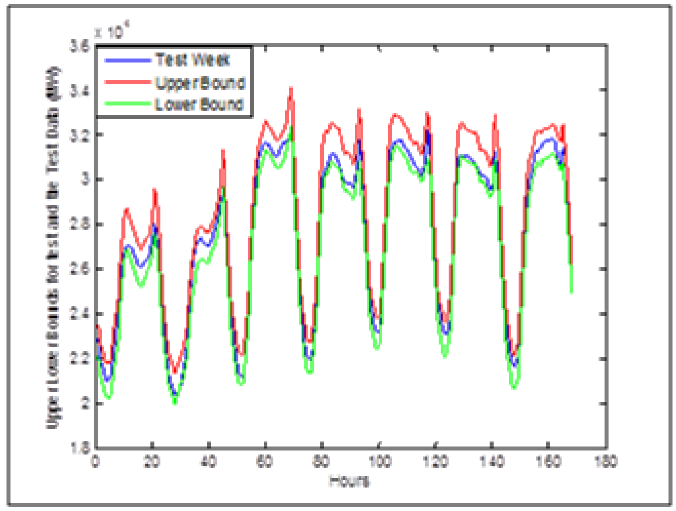

Fig. 7. The test week of load for PJM market, 24^{th} to $30^{t}h$, April, 2010

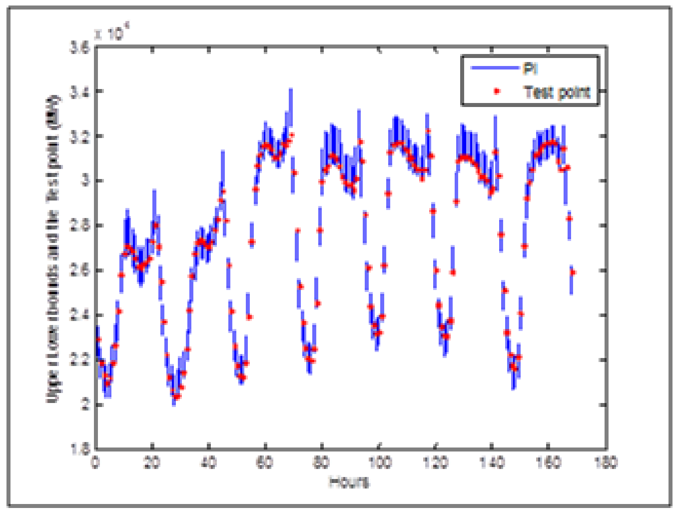

Fig. 8. PIs for test week of load in PJM, 24^{th} to 30^{th}, April, 2010

Table 5. PI assessment indices of electricity price for VIC power market

proposed	May				August				December			
	PICP(%)	PINAW(%)	CWC(%)	SCORE	PICP(%)	PINAW(%)	CWC(%)	SCORE	PICP(%)	PINAW(%)	CWC(%)	SCORE
1	95.23	2.21	2.21	-47.46	95.9	4.03	4.03	-19.8	96.91	3.78	3.78	-26.05
2	95.36	2.19	2.19	-47.53	95.23	3.02	3.02	-22.36	96.98	3.1	3.1	-21.24
3	94.76	2.35	2.35	-50.72	95.83	3.43	3.43	-18.62	96.24	3.15	3.15	-26.49
4	94.09	2.57	2.57	-55.47	94.35	3.19	3.19	-18.56	96.24	3.31	3.31	-26.71
5	94.42	2.9	2.9	-66.05	94.02	3.43	3.43	-27.33	96.03	3.49	3.49	-26.93
Median	94.77	2.44	2.44	-53.45	95.07	3.42	3.42	-21.33	96.48	3.37	3.37	-25.49
bootstrap-GARCH	94.26	15.28	15.28	—	94.93	15.02	15.02	—	91.89	26.23	26.23	—
Delta	98.65	37.82	37.82	—	98.31	33.92	33.92	—	100	439.3	439.3	—

respectively. For the January test week PICP of the SVM-PSO is less than the nominal μ, which is 90

For Ontario market the obtained results are reported in table 8. The results are compared with [32]. The standard deviations of CWC for the four weeks are 0.338, 0.2188, 0.5655 and 0.2868 respectively. The comparison shows better performance of the proposed method. The improvement for the average of total year is 55.24%. During these four weeks, the week of April has the least absolute SCORE. Figs. 10 and 11 show the upper, lower bounds and PIs constructed by the proposed method for this

Table 6. Compared results in 30 for the test week of January in Ontario and the proposed method

	PICP	PINAW	CWC	SCORE
Naïve	86.90	50.18	54.88	-2030.20
QR	11.90	9.20	9.08E+16	-3917.00
CART	88.10	36.49	39.087	-1774.30
LR	94.64	35.61	35.61	-1604.40
Bootstrap	95.83	35.56	35.56	-1541.40
MVE	26.79	3.82	5.33E+15	-4582.00
LUBE NN	97.62	104.38	104.38	-3179.80
SVM-PSO	90.48	26.58	26.58	-1226.00
proposed	96.79	9.04	9.04	-11.04

Table 7. PI assessment indices of electricity price for PJM power market

proposed		PICP(%)	PINAW(%)	CWC(%)	SCORE
January	1	97.62	22.28	22.28	-4.89
	2	94.64	21	21	-4.7
	3	97.62	21.39	21.39	-4.74
	4	96.43	21.5	21.5	-5.06
	5	96.43	20.59	20.59	-5.02
	Median	96.55	21.35	21.35	-4.88
	SVM+PSO	89.88	15.87	16.93	-594.2
April	1	94.64	15.46	15.46	-1.65
	2	94.05	15.96	15.96	-1.77
	3	94.64	15.79	15.79	-1.89
	4	94.64	15.25	15.25	-1.74
	5	94.05	16.17	16.17	-1.72
	Median	94.4	15.72	15.72	-1.75
	SVM+PSO	90.48	14.69	14.69	-232.6
July	1	94.05	13.25	13.25	-6.04
	2	94.64	12.49	12.49	-5.69
	3	94.05	12.58	12.58	-5.95
	4	95.83	12.09	12.09	-5.12
	5	95.24	13.44	13.44	-5.51
	Median	94.76	12.77	12.77	-5.66
	SVM+PSO	90.48	10.35	10.35	-416.8
October	1	94.05	8.41	8.41	-2.03
	2	94.05	8.78	8.78	-2.09
	3	95.83	9.47	9.47	-2.35
	4	94.05	8.83	8.83	-2.23
	5	94.64	8.6	8.6	-2.08
	Median	94.52	8.82	8.82	-2.16
	SVM+PSO	91.67	25.60	25.60	-433.3
Average of the year	proposed	95.06	14.67	14.67	-3.61
	SVM+PSO	90.63	16.63	16.63	-419.23

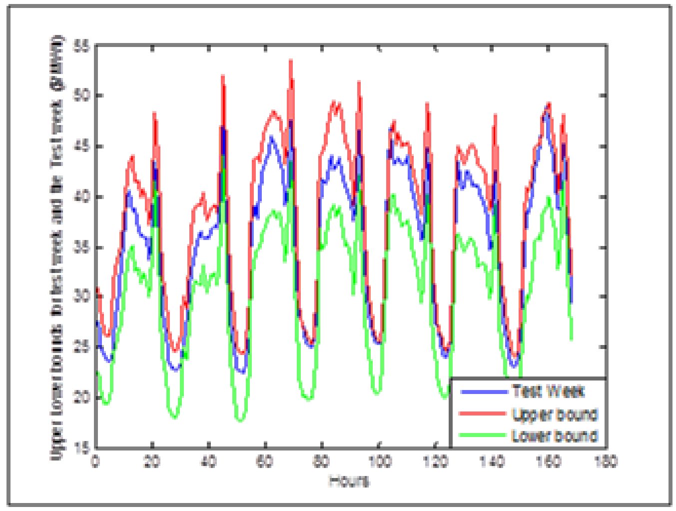

Fig. 9. The test week of price for PJM market, 24^{t}h to 30th, April, 2010

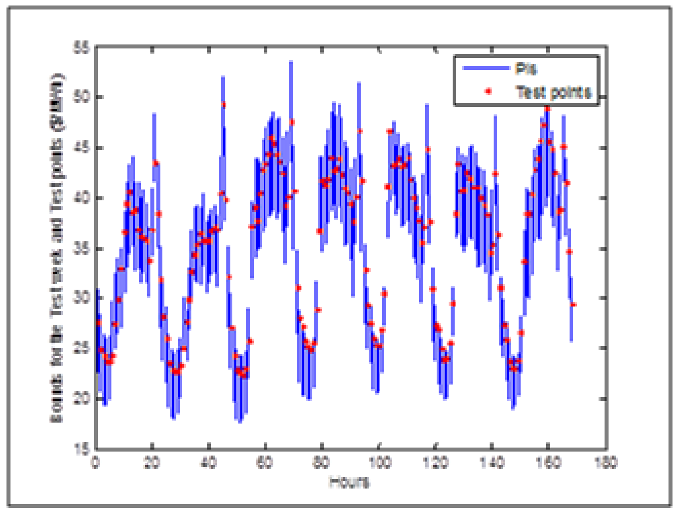

Fig. 10. PIs for test week of price in PJM, 24th to 30th, April, 2010

test week.

7. CONCLUSION

Accurate load and price forecasting is an important factor for market participants. As literatures show point forecasting is still an application, but it cannot quantify the uncertainties in load and price signals due to some uncertainty sources such as renewable energy. Prediction intervals are suitable tools for this purpose. To overcome the troubles of traditional methods for PIs construction, a new method named LUBE is developed to predict intervals for future observes. In this paper a novel technique is utilized to improve the performance of LUBE method. A neural network with two outputs is trained with trainLM first, then the optimized weights are given to the PSO to improve the indices of PIs. A single objective function CWC is minimized with PSO. Comparisons with several references of load and price prediction intervals indicate better performance of the proposed technique. Test weeks are similar to the references due to a fair comparison. The methods proposed in these references have

Table 8. PI assessment indices of electricity price for Ontario power market

proposed		PICP(%)	PINAW(%)	CWC(%)	SCORE
	1	96.43	8.63	8.63	-10.75
	2	96.43	8.83	8.83	-11.64
	3	96.43	9.17	9.17	-10.91
January	4	97.02	9.51	9.51	-11.23
	5	97.62	9.06	9.06	-10.66
	Median	96.79	9.04	9.04	-11.04
	SVM+PSO	90.48	26.58	26.58	-1226.0
	1	90.48	9.3	9.3	-8.6
	2	90.48	9.43	9.43	-8.42
	3	91.07	9.52	9.52	-7.48
April	4	90.48	9.06	9.06	-7.79
	5	91.67	9.62	9.62	-7.84
	Median	90.83	9.39	9.39	-8.03
	SVM+PSO	92.26	33.17	33.17	-1136.2
	1	92.26	16.79	16.79	-33.93
	2	92.86	17.32	17.32	-34.14
	3	92.26	16.51	16.51	-30.27
July	4	92.86	17.61	17.61	-34.38
	5	92.86	17.87	17.87	-34.34
	Median	92.62	17.22	17.22	-33.41
	SVM+PSO	90.48	7.58	7.58	-4797.8
	1	99.4	8.61	8.61	-10.38
	2	99.4	8.2	8.2	-9.64
	3	99.4	8.5	8.5	-10.28
October	4	99.4	8.97	8.97	-10.64
	5	99.4	8.74	8.74	-10.22
	Median	99.4	8.61	8.61	-10.23
	SVM+PSO	90.48	31.49	31.49	-1000.7
Average of the year	proposed	94.91	11.06	11.06	-15.68
	SVM+PSO	90.93	24.71	24.71	-4790.18

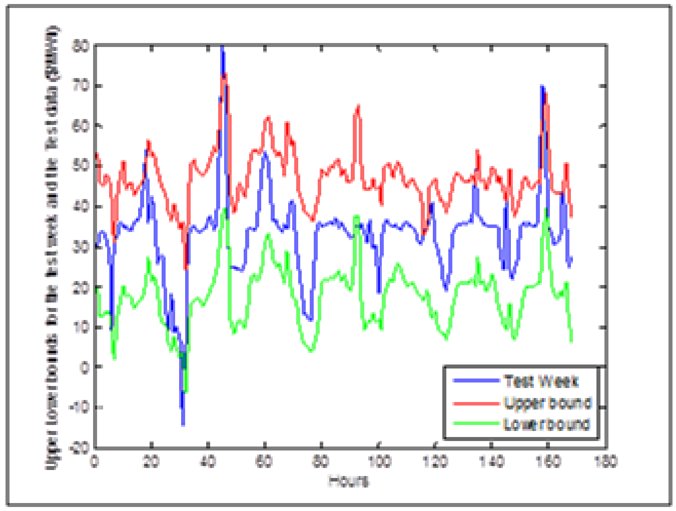

Fig. 11. upper lower bounds for price signal, 24^{th} to 30^{th}, April 2010, Ontario market

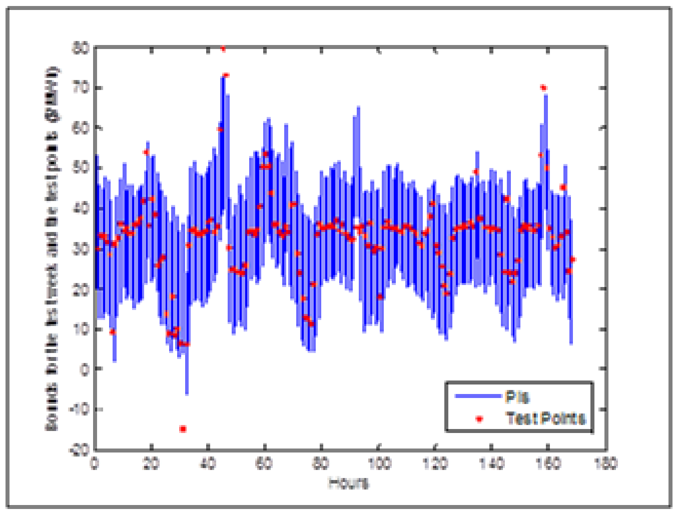

Fig. 12. PIs for the price test week in Ontario, 24^{th} to 30^{th}, April 2010, Ontario market

been compared with several traditional methods and proposed method is better than these references. This proposed technique constructs reliable PIs for both load and price signals.

ACKNOWLEDGMENTS:

The financial support of the Graduate University of Advanced Technology of Kerman under the grant number of 2590 is greatly appreciated.

REFERENCES

1. A. D. Papalexopoulos and T. C. Hesterberg, "A regression-based approach to short-term system load forecasting," Power Systems, IEEE Transactions on, vol. 5, pp. 1535-1547, 1990.

2. C. Sigauke and D. Chikobvu, "Prediction of daily peak electricity demand in South Africa using volatility forecasting models," Energy Economics, vol. 33, pp. 882-888, 9// 2011.

3. A. Bakirtzis, J. Theocharis, S. Kiartzis, and K. Satsios, "Short term load forecasting using fuzzy neural networks," Power Systems, IEEE Transactions on, vol. 10, pp. 1518-1524, 1995.

4. A. S. Khwaja, M. Naeem, A. Anpalagan, A. Venetsanopoulos, and B. Venkatesh, "Improved short-term load forecasting using bagged neural networks," Electric Power Systems Research, vol. 125, pp. 109-115, 8// 2015.

5. M. Ghayekhloo, M. B. Menhaj, and M. Ghofrani, "A hybrid short-term load forecasting with a new data preprocessing framework," Electric Power Systems Research, vol. 119, pp. 138-148, 2// 2015.

6. B.-J. Chen, M.-W. Chang, and C.-J. Lin, "Load forecasting using support vector machines: A study on EUNITE competition 2001," Power Systems, IEEE Transactions on, vol. 19, pp. 1821-1830, 2004.

7. S. Li, P. Wang, and L. Goel, "Short-term load forecasting by wavelet transform and evolutionary extreme learning machine," Electric Power Systems Research, vol. 122, pp. 96-103, 5// 2015.

8. W.-C. Hong, "Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm," Energy, vol. 36, pp. 5568-5578, 9// 2011.

9. D. Niu, Y. Wang, and D. D. Wu, "Power load forecasting using support vector machine and ant colony optimization," Expert Systems with Applications, vol. 37, pp. 2531-2539, 2010.

10. P.-F. Pai and W.-C. Hong, "Support vector machines with simulated annealing algorithms in electricity load forecasting," Energy Conversion and Management, vol. 46, pp. 2669-2688, 2005.

11. A. J. Conejo, M. A. Plazas, R. Espinola, and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," Power Systems, IEEE Transactions on, vol. 20, pp. 1035-1042, 2005.

12. J. Contreras, R. Espinola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," Power Systems, IEEE Transactions on, vol. 18, pp. 1014-1020, 2003.

13. E. Hickey, D. G. Loomis, and H. Mohammadi, "Forecasting hourly electricity prices using ARMAX–GARCH models: An application to MISO hubs," Energy Economics, vol. 34, pp. 307-315, 1// 2012.

14. H. Liu and J. Shi, "Applying ARMA–GARCH approaches to forecasting short-term electricity prices," Energy Economics, vol. 37, pp. 152-166, 5// 2013.

15. H. Qu, W. Chen, M. Niu, and X. Li, "Forecasting realized volatility in electricity markets using logistic smooth transition heterogeneous autoregressive models," Energy Economics, vol. 54, pp. 68-76, 2// 2016.

16. Y.-Y. Hong and C.-Y. Hsiao, "Locational marginal price forecasting in deregulated electricity markets using artificial intelligence," IEE Proceedings-Generation, Transmission and Distribution, vol. 149, pp. 621-626, 2002.

17. B. Szkuta, L. Sanabria, and T. Dillon, "Electricity price short-term forecasting using artificial neural networks," Power Systems, IEEE Transactions on, vol. 14, pp. 851-857, 1999.

18. T. Papadimitriou, P. Gogas, and E. Stathakis, "Forecasting energy markets using support vector machines," Energy Economics, vol. 44, pp. 135-142, 7// 2014.

19. J. C. R. Filho, C. d. M. Affonso, and R. C. L. de Oliveira, "Energy price prediction multi-step ahead using hybrid model in the Brazilian market," Electric Power Systems Research, vol. 117, pp. 115-122, 12// 2014.

20. A. Khosravi, S. Nahavandi, and D. Creighton, "Construction of optimal prediction intervals for load forecasting problems," Power Systems, IEEE Transactions on, vol. 25, pp. 1496-1503, 2010.

21. C. M. Bishop, "Neural networks for pattern recognition," 1995.

22. J. G. Hwang and A. A. Ding, "Prediction intervals for artificial neural networks," Journal of the American Statistical Association, vol. 92, pp. 748-757, 1997.

23. T. Heskes, "Practical confidence and prediction intervals," Advances in neural information processing systems, pp. 176-182, 1997.

24. A. Khosravi, S. Nahavandi, and D. Creighton, "A neural network-GARCH-based method for construction of Prediction Intervals," Electric Power Systems Research, vol. 96, pp. 185-193, 2013.

25. A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Lower upper bound estimation method for construction of neural network-based prediction intervals," Neural Networks, IEEE Transactions on, vol. 22, pp. 337-346, 2011.

26. H. Quan, D. Srinivasan, and A. Khosravi, "Construction of neural network-based prediction intervals using particle swarm optimization," in Neural Networks (IJCNN), The 2012 International Joint Conference on, 2012, pp. 1-7.

27. H. Quan, D. Srinivasan, and A. Khosravi, "Particle swarm optimization for construction of neural network-based prediction intervals," Neurocomputing, vol. 127, pp. 172-180, 2014.

28. H. Quan, D. Srinivasan, and A. Khosravi, "Short-term load and wind power forecasting using neural network-based prediction intervals," Neural Networks and Learning Systems, IEEE Transactions on, vol. 25, pp. 303-315, 2014.

29. H. Quan, D. Srinivasan, and A. Khosravi, "Uncertainty handling using neural network-based prediction intervals for electrical load forecasting," Energy, vol. 73, pp. 916-925, 2014.

30. H. Quan, D. Srinivasan, A. Khosravi, S. Nahavandi, and D. Creighton, "Construction of neural network-based prediction intervals for short-term electrical load forecasting," in Computational Intelligence Applications In Smart Grid (CIASG), 2013 IEEE Symposium on, 2013, pp. 66-72.

31. M. Rana, I. Koprinska, A. Khosravi, and V. G. Agelidis, "Prediction intervals for electricity load forecasting using neural networks," in Neural Networks (IJCNN), The 2013 International Joint Conference on, 2013, pp. 1-8.

32. N. Shrivastava, A. Khosravi, and B. Panigrahi, "Prediction Interval Estimation of Electricity Prices using PSO tuned Support Vector Machines." Industrial Informatics IEEE Transaction on, Vol 11, No 2, pp: 322-330, April 2011.

33. R. L. Winkler, "A decision-theoretic approach to interval estimation," Journal of the American Statistical Association, vol. 67, pp. 187-191, 1972.

34. A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Comprehensive review of neural network-based prediction intervals and new advances," Neural Networks, IEEE Transactions on, vol. 22, pp. 1341-1356, 2011.

35. N. Amjady and F. Keynia, "Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method," International Journal of Electrical Power Energy Systems, vol. 30, pp. 533-546, 2008.

36. N. Amjady and F. Keynia, "Day-ahead price forecasting of electricity markets by mutual information technique and cascaded neuro-evolutionary algorithm," Power Systems, IEEE Transactions on, vol. 24, pp. 306-318, 2009.

37. F. Keynia, "A new feature selection algorithm and composite neural network for electricity price forecasting," Engineering Applications of Artificial Intelligence, vol. 25, pp. 1687-1697, 2012.

38. N. Amjady, "Short-term bus load forecasting of power systems by a new hybrid method," Power Systems, IEEE Transactions on, vol. 22, pp. 333-341, 2007.

39. J. Kenndy and R. Eberhart, "Particle swarm optimization," in Proceedings of IEEE International Conference on Neural Networks, 1995, pp. 1942-1948.

40. A. E. Munoz-Zavala, A. Hernandez-Aguirre, E. R. Villa-Diharce, and S. Botello-Rionda, "PESO+ for constrained optimization," in Evolutionary Computation, 2006. CEC 2006. IEEE Congress on, 2006, pp. 231-238.

41. Hourly load data archives Online. Available: www.ercot.com/gridinfo/load/load-hist.

42. PJM Web Site [Online]. Available: www.pjm.com

43. AEMO (Australian Energy Market Operator) [Online]. Available: http://www.aemo.com.au/Electricity/Data/Price-and-Demand/Aggregated-Price-and-Demand-Data-Files

44. Independent electricity system operator (IESO) website [Online]. Available:http://ieso.ca/imoweb/marketData/marketData.asp/.

45. Tsekouras GJ, Hatziargyriou ND, Dialynas EN. An optimized adaptive neural network for annual midterm energy forecasting. IEEE Trans Power Syst 2006;21:385–91.