Hourly Electricity Price Forecasting for the Next Month Using Multilayer Neural Network
Prédiction du tarif horaire de l'électricité pour le mois suivant utilisant les réseaux de neurones multicouches

Hossam Mosbah and Mohamed El-Hawary, *Life Fellow, IEEE*

***Abstract*—Load and price forecasting are key challenges for current electricity market participants. Load and price in electricity markets have complex peculiarities, such as nonlinearity, being nonstationary and irregular. Accurate short-term forecasting, such as hourly electricity price forecasting (EPF) for the next month gives pivotal information to power producers and consumers to enhance precise techniques to maximize their profit. This paper deals with short-term hourly EPF for the next month (January 2006), using the historical hourly data for the year 2005 as a training set. A new approach of multilayer neural networks is applied in composite topologies in order to improve forecasting accuracy. The intent is to study the behavior of diverse composite topologies to compare the best performance indices evaluated by the mean absolute percentage error and mean square error. The performance of different topologies is compared to identify the best connection architecture. The data used in the forecasting are hourly historical data of the temperature, electricity load, and natural gas price from the Australian electricity markets.**

***Résumé*—La prédiction du tarif et de la charge sont les principaux défis pour les acteurs actuels du marché de l'électricité. La charge et le prix sur les marchés de l'électricité ont des particularités complexes, telles que la non-linéarité, étant non stationnaire et irrégulière. Une prédiction précise à court terme, tels que la prédiction du tarif horaire de l'électricité (PTHE) pour le mois suivant donne des informations clés aux producteurs d'électricité et les consommateurs pour améliorer les techniques précises afin de maximiser leur profit. Cet article traite de la PTHE à court terme pour le mois suivant (Janvier 2006), utilisant les données historiques horaires de l'année 2005 comme ensemble d'apprentissage. Une nouvelle approche des réseaux de neurones multicouches est appliquée dans les topologies composites afin d'améliorer la précision de la prédiction. Le but est d'étudier le comportement des topologies composites divers pour comparer les meilleurs indices de performance évalués par le pourcentage d'erreur absolue moyen et l'erreur quadratique moyenne. La performance des différentes topologies est comparée pour identifier la meilleure architecture de connexion. Les données utilisées dans la prédiction sont les données historiques horaires de la température, la charge électrique, et le prix du gaz naturel à partir des marchés de l'électricité en Australie.**

***Index Terms*—Artificial neural networks, electricity price forecasting (EPF), load forecasting, multilayer neural networks.**

I. Introduction

IN DEREGULATED electricity markets, electricity prices will fluctuate as a result of competition among power suppliers. Profit maximization has become a major motivation in electric markets. The imbalance between supply and demand results in volatile electricity prices. In a regulated market, load forecasting was a main focus of the electric power industry. Subsequently, the more complex electricity price forecasting (EPF) has become more important in the deregulated electric power industry [1]–[5].

EPF is a critical factor in decision making in power systems. Its principal target is to lower the cost of electricity through rivalry, and augment efficient generation and consumption of electricity. In view of the nonstorable nature of electricity, all generated electricity must be consumed. Thus, both the producers and the consumers need accurate price forecasting so as to build their own particular methodologies for profit or utility augmentation [6].

EPF depends on input variables, such as available historical price and load data, system operating conditions, weather conditions and temperature values, fuel prices, time indices (including hours, week days, and seasons), demand, bidding strategies, operating reserves, imports, temperature effects, predicted power shortfalls, and generation outages [7], [8].

Bunn [9] reviewed several major innovative methodologies and techniques that focused on daily loads and prices forecasting in challenging electricity markets, such as variable segmentation, multiple modeling, combinations, and neural networks. He focused on forecasting strategy in both supply and demand side and discussed the benefit side of considering

Manuscript received August 27, 2015; accepted June 20, 2016. Date of current version December 6, 2016.

The authors are with the Department of Electrical and Computer Engineering, Dalhousie University, Halifax, NS B3H 4R2, Canada (e-mail: hs617117@dal.ca; elhawary@dal.ca).

Associate Editor managing this paper's review: Hilmi Turanli.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/CJECE.2016.2586939

more influential factors, such as weather. He suggested that further research and integration are needed in order to provide more accurate forecasts, so he concluded that "the forecasting of loads and prices are mutually intertwined activities and the economic perspective alone cannot be an accurate basis for daily."

Skantze and Ilic [10] classified power value models and existing significant distributions into six categories, and discussed their goals, attributes, and drawbacks. Shahidehpour *et al.* [11] discussed the basics of electricity demand estimation, including price, unpredictability, and exogenous variables, and proposed a value determining module focused on neural networks.

Koritarov [12] promoted using agent-based models to uncover and clarify the perplexing and total framework practices that arise out of the interaction of the heterogeneous individual elements. Ventosa *et al*. [13] discussed the diverse methodologies to study electricity markets, including the Nash–Cournot framework and the supply work balance approach.

Amjady and Hemmati [14] clarify the requirement for short-term price forecasts, survey issues identified with EPF, and set forward suggestions for such expectations. They contend that time series techniques, including autoregressive (AR), AR integrated moving average model (ARIMA), and generalized AR conditional heteroskedasticity (GARCH), are by and large just as effective in the regions where the frequency of the information is low, for example, week after week designs. Moreover, they advocate the utilization of computational insights and cross methodologies (neural systems, fuzzy regression, fuzzy neural networks, cascaded architecture of neural networks, and committee machines), which are fit for following the hard nonlinear behaviors of hourly load and particularly value signals [14].

Weron [15] provided an outline of modeling methodologies, and then focused on pragmatic applications of statistical strategies for day-ahead forecasting, including autoregressive integrated moving average (ARIMA), autoregressive Moving Average exogenous (ARMAX), and generalized autoregressive conditional heteroskedasticity (GARCH). Moreover, he discussed interval forecasts and proceeded to review quantitative stochastic models for subordinates estimating jump-diffusion models and Markov regime switching.

Zareipour [16] reviewed the time series of linear models, such as AR with exogenous input, ARIMA, and ARMAX, and nonlinear models, including regression splines and multilayer neural networks, and then used them to forecast hourly costs in the Ontario force market.

Amjady [17] briefly reviewed EPF techniques, then concentrated on artificial intelligence (AI)-based methods, and specifically emphasized the choice of procedures and hybrid forecast engines.

García-Martos and Conejo [18] investigated the short- and medium-term EPF, with a concentrate on time series models. Particularly, they consider ARIMA and regular ARIMA models aligned to hourly costs for day-ahead forecasts and vector ARIMA. Hong [19] discussed spatial load forecasting, short-term load forecasting, EPF, and two smart grid era research areas, including demand-response and renewable-generation forecasting. He categorized EPF models into three classes, i.e., the simulation systems of the power markets, load forecasts, and blackout data, and offers from business members, measurable techniques, and AI strategies [19], [20].

The latest study of structural models, distributed as a chapter in the book *Quantitative Energy Finance* by Carmona and Coulon [21], presented a point-by-point investigation of the structural methodology for power demonstrating, underlining its benefits with respect to customary diminished structure models. Expanding on a few latest articles, they advocate a wide and adaptable structural system at spot costs, consolidating demand, limit, and fuel costs in a few ways, while computing closed structure forward costs all through. The aforementioned articles, book chapters, and Ph.D. are supplemented by a couple of the overview meeting papers of changing quality [21].

In this paper, composite multilayer neural network topologies, including hybrid parallel and hybrid cascade topologies, are applied to enhance the hourly EPF for the next month in Australian electricity markets. The hourly temperature, hourly electricity load, hourly natural gas data, and other hourly historical data have been considered in forecasting. The neural network models are trained on hourly historical data for the year of 2005 from Australian electricity markets published by the Australian Energy Market Operator (AEMO) to predict the hourly EPF for January 2006. The simulation results obtained have shown that hybrid parallel topology is more accurate and less in computational time.

II. System Description

This paper presents an attempt to design the interconnections of neural networks with different connection topologies to improve the overall forecasting performance accuracy. A review of a number of successful implementations of multiple connection topologies is given first. Neural networks may be connected in many configurations, such as the following:

1) cascade or serial topology;
2) parallel topology;
3) cascade parallel topology in cascade and in parallel;
4) parallel–cascade topology in cascade and in parallel;
5) hybrid parallel topology;
6) hybrid cascade topology.

Fig. 1 shows the cascade topology interconnected in feedforward connection where the output of the first network is used as input for the second network. The neural network design is divided into stages, i.e., the training stage and the simulation stage. Each network is trained to optimize its inputs with respect to the target values.

As a result, Net 1 will be trained using x input value and t target value, while Net 2 will be trained with the output of Net 1, $y1$, as an input to it with respect to t target value, so connecting neural networks in cascade topology and finding the best cascade sequence. For example, speed and accuracy are evaluated for the cascade topology. The sequence (or order) is very important in this connection topology. It can be concluded that accurate network should be placed

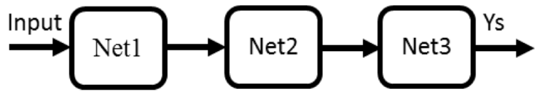

Fig. 1. Cascade topology.

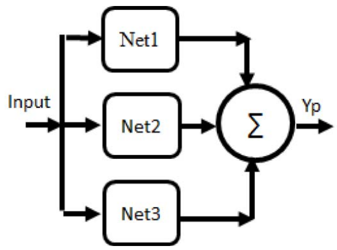

Fig. 2. Parallel topology.

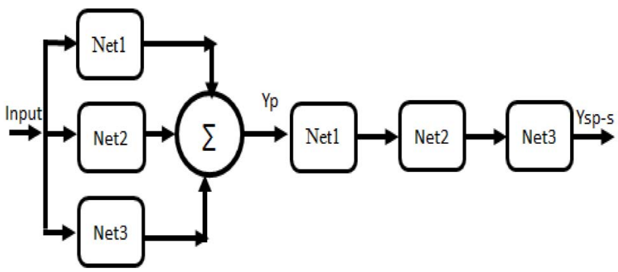

Fig. 3. Parallel–cascade-in-cascade connection.

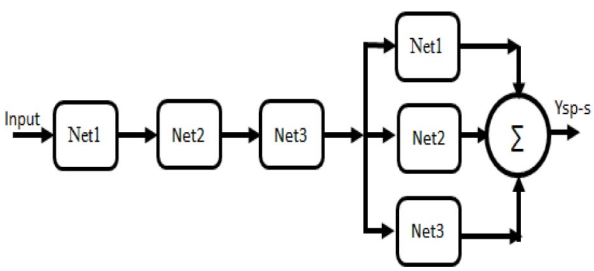

Fig. 4. Cascade–parallel-in-cascade connection.

first to minimize the overall fitting error and to enhance the overall performance.

In the parallel topology shown in Fig. 2, one starts using the same input applied to all networks and then the output of all the networks is aggregated into to a single output neuron. Different weights are assigned to each network output where the best performing network is assigned higher weight.

In connecting neural networks in parallel, there is no difference in connection sequence and the final output is improved compared with the output of each network resulting from the averaging mechanism. Enhanced results can be obtained by assigning a higher weight to the network with best performance and lower or even zero weight for the weak network. There are four connection combinations for cascade–parallel, such as cascade–parallel in cascade, cascade–parallel in parallel, parallel–cascade in series, and parallel–cascade in parallel. Figs. 3–6 show the difference between the cascade–parallel in cascade and the parallel–cascade in cascade, and also sequence between the cascade–parallel in parallel and the parallel–cascade in parallel.

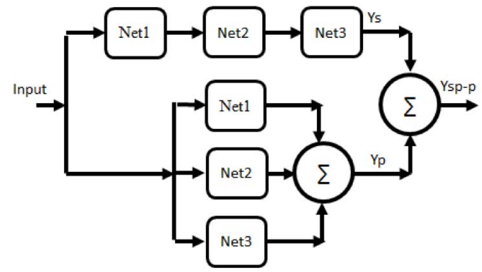

Fig. 5. Cascade–parallel-in-parallel connections.

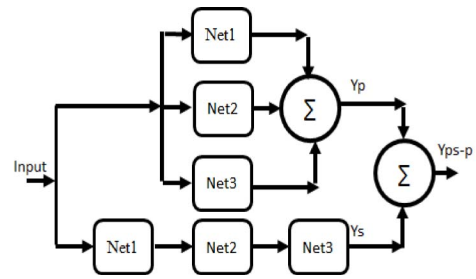

Fig. 6. Parallel–cascade-in-parallel connections.

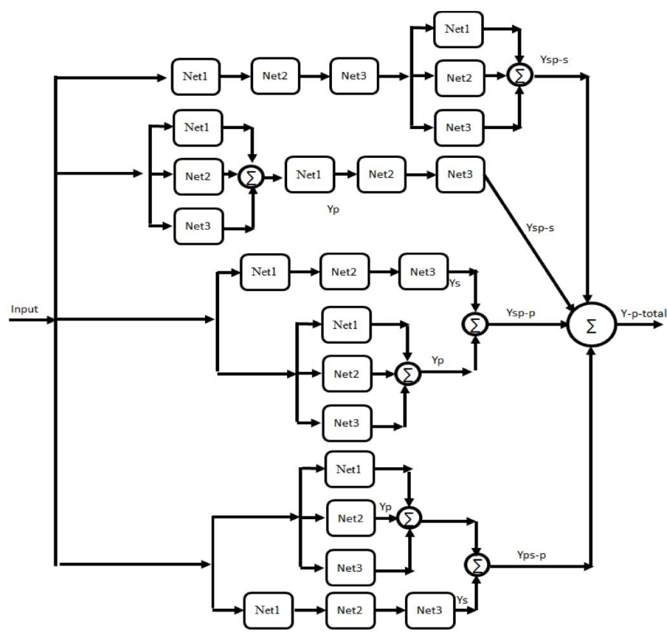

Fig. 7. Hybrid parallel topology.

Figs. 7 and 8 demonstrate hybrid parallel topology and hybrid cascade topology to determine the most efficient neural network interconnection topology in terms of accuracy and time management. The output and input of each hypertopology could be connected in layers, and the connection terminal of each one is connected to the other layers whether in cascade or parallel. For evaluating, each network topology with three different feedforward networks is trained with the same input–output data and each network has six hidden layers for Net 1, seven hidden layers for Net 2, and eight hidden layers for Net 3, which will result in a good fit, and then different hybrid connection topologies are evaluated.

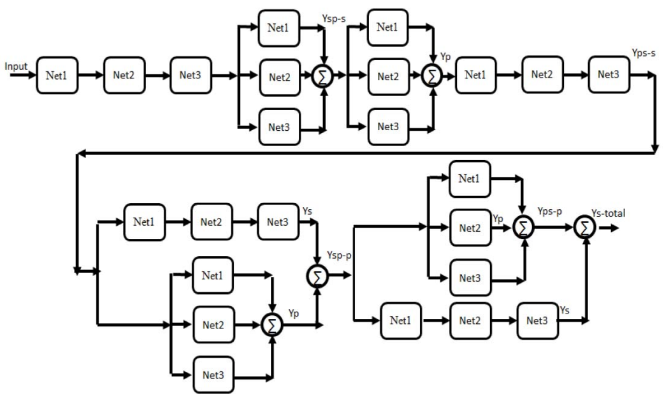

Fig. 8. Hybrid cascade topology.

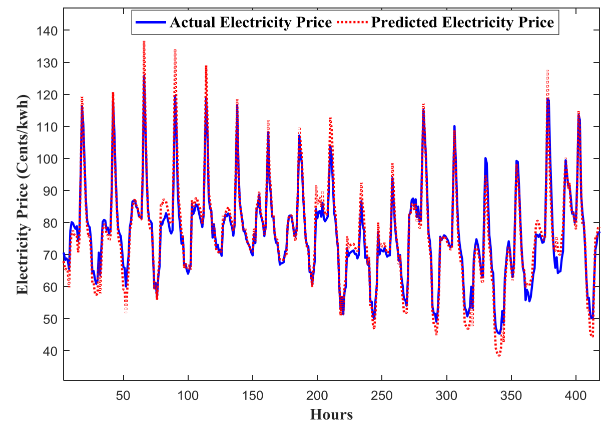

Fig. 9. Hybrid topology in parallel connection—actual and predicted hourly electricity prices for January 2006.

III. Simulation Results and Discussion

Our computational experiment involves the hourly historical data for the year of 2005 as training data to predict the hourly electricity price for January 2006. The data used are published by the AEMO in electricity load forecasting for the Australian market [22].

There are many impact factors that can be used to forecast the electric price, such as the following:

1) time in hours, including previous day same hour load, previous week same hour load, and previous 24 h average load;
2) weather conditions, including dry bulb and dew point;
3) electricity price, including previous day same hour price, previous week same hour price, and previous 24 h average price;
4) dates, including days of the week;
5) natural gas price, including previous day same hour price, previous week same hour price;
6) system loads, including previous load forecasting;
7) none of the business days working, including holidays.

Here the inputs include hourly load data, hourly weather conditions, hourly natural gas price, and hourly other factors for the year of 2005, and target includes the electricity price for year of 2005. We use the training data, which include inputs and targets, to predict the hourly electricity forecasting for January 2006

$$\mathrm{MSE} = \frac{1}{N_i}\sum_{i=1}^{n}\frac{1}{N_i}\sum_{I=1}^{N}[y_i - \hat{y}]^2 \tag{1}$$

$$\mathrm{MAPE} = \frac{1}{N_i}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}}{y_i}\right| \times 100 \tag{2}$$

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y} - \bar{y})}{\sum_{i=1}^{n}(y_i - \hat{y})} \tag{3}$$

where y_i and $\hat{y}$ are the actual and predicted prices, respectively, for January 2006, respectively, and N_i is number of predicted data. R^2 is the coefficient of determination, $\bar{y}$ is the average of the data [6].

A few comments on the selections made in our experiment are as follows.

1) We arbitrarily selected three different layers to compare performance and accuracy behavior between the proposed topologies. More layers require more computation, but their use might result in the network solving complex problems more efficiently.
2) We chose the number of neurons in each network based on the best performance, so we put six neurons for Net 1, seven neurons for Net 2, and eight neurons for Net 3. More neurons require more computation, and they have a tendency to overfit the data when the number is set too high, but they allow the network to solve more complicated problems.
3) We selected the initial weights of a neural network from the range $(-1/\sqrt{d},\ 1/\sqrt{d})$, where d is the number of inputs to a given neuron. It is assumed that the sets are normalized—mean 0 and variance 1.

We run the polynomial regression, different model selections include forward selection and backward elimination, to improve the full regression model whether by eliminating bad predictors (input factors) or by selecting good predictors (input factors). It is concluded that natural gas price has the high impact on the electricity price.

The weights in Table I are selected to improve the overall hybrid topology by giving high weights for the topology with the best contributed performance and low weights for the topologies with weak performance. The actual values of the weights were selected through trial and error.

The data are categorized into 75% used for training and 15% used for validation and testing of each network.

Figs. 9 and 10 show the predicted and actual electricity prices for January 2006. The total hours are 720 h for a month; however, we just take 400 h to see the variation between actual and predicted electricity prices. It can be seen that the hybrid parallel topology forecasting is less variable than the hybrid cascade topology forecasting.

Figs. 11 and 12 display the absolute percentage error (APE) of the hybrid parallel and cascade topologies for the actual and predicted electricity prices for January 2006. We have divided the total number of hours (720 h a month/60 h = 12 boxes).

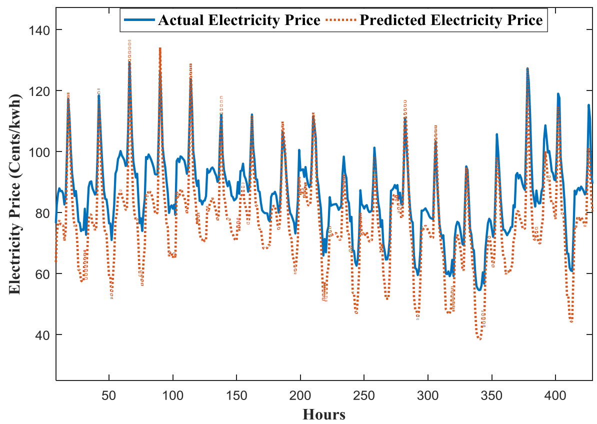

Fig. 10. Hybrid topology in cascade connection—actual and predicted hourly electricity prices for January 2006.

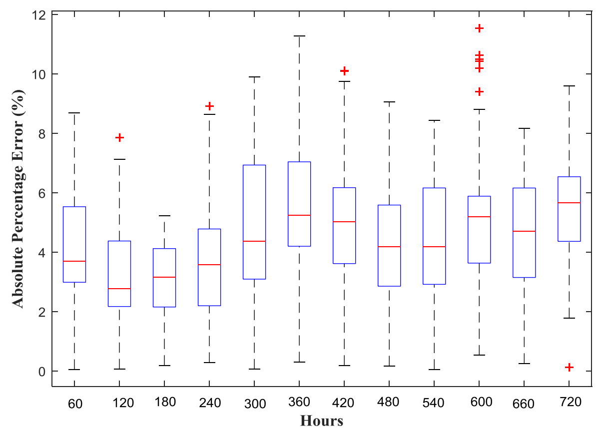

Fig. 11. Hybrid parallel topology—APE for hourly electricity price for January 2006.

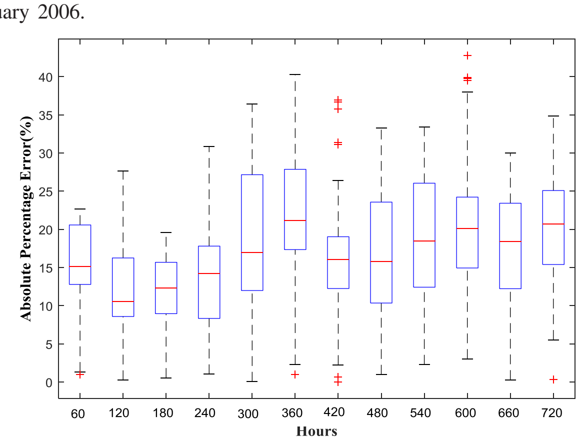

Fig. 12. Hybrid cascade topology—APE for hourly electricity price for January 2006.

Each box consists of 6 h of APE data and also each box represents different ranges that include maximum, average, and minimum for APE. The range of the APE of the hybrid parallel topology is less than that of the hybrid cascade topology.

Figs. 13 and 14 show the results of the hybrid parallel and cascade topologies' residuals between actual and predicted hourly electricity price for January 2006. It is evident that the residual range of hybrid parallel topologies is less variant than hybrid cascade topology.

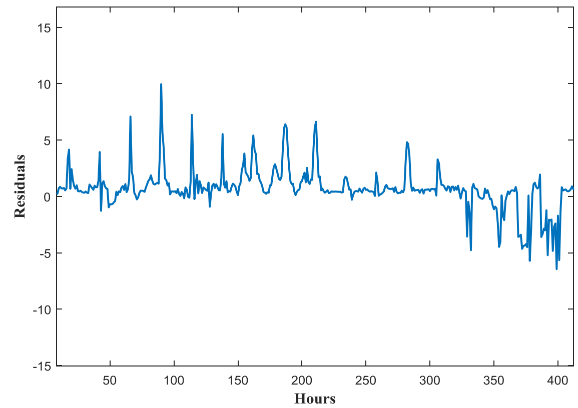

Fig. 13. Hybrid parallel topology—residual between the actual and the predicted hourly price forecasting for January 2006.

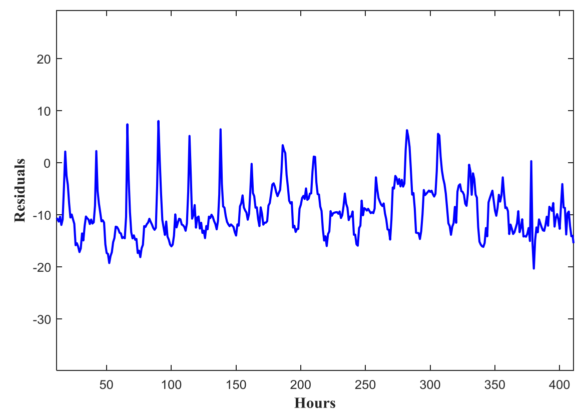

Fig. 14. Hybrid cascade topology—residual between the actual and the predicted hourly price forecasting for January 2006.

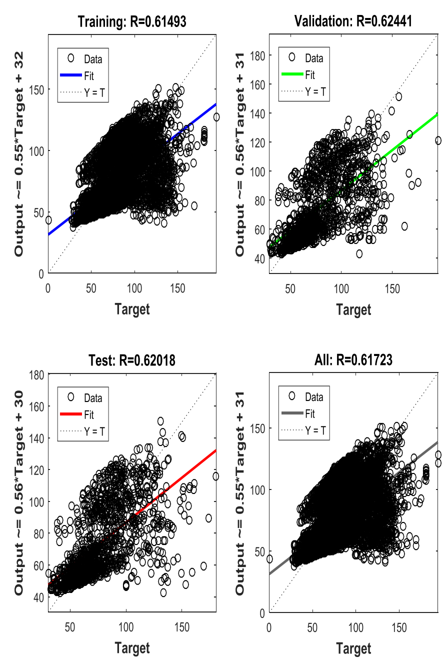

Fig. 15. Hybrid cascade topology—regression model for the actual and the predicted hourly price forecasting for January 2006.

Figs. 15 and 16 show the regression plots displaying the output of both the hybrid parallel and cascade topologies with respect to the targets for hourly data (January 2006),

TABLE I
WEIGHTS FOR THREE INTERCONNECTED NETS

TOPOLOGIES	Cascade-Parallel in Cascade W1	Parallel-Cascade in Parallel topology W2	Cascade-Parallel in Parallel topology W3	Cascade-Parallel in Parallel topology W4
HYBRID PARALLEL TOPOLOGY	0.50	0.95	0.70	1.0

TABLE II
FINAL COMPARISON OF CASCADE AND PARALLEL HYBRID CONNECTIONS IN DIFFERENT PERFORMANCES

HYBRID CONNECTIONS	Mean Square Error (MSE)	Mean Absolute Percentage Error (MAPE)	Coefficient of determination (R^2)	Time Seconds Training	Simulation
HYBRID TOPOLOGY IN PARALLEL	9.11	3.25	0.9815	228.21	0.086
HYBRID TOPOLOGY IN CASCADE	87.14	12.27	0.9420	912.82	0.065

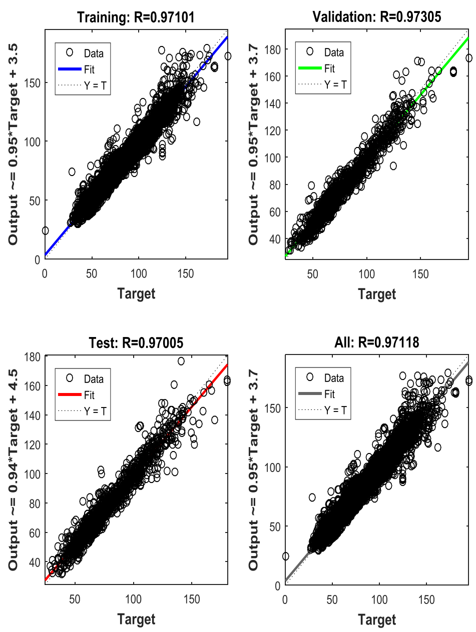

Fig. 16. Hybrid parallel topology—regression model for the actual and the predicted hourly price forecasting for January 2006.

validation, and test sets. The data should fall along a 45° line to satisfy the equality of both the output and the target. For this case, the regression value in each and total case is 0.97118 in hybrid parallel topology compared with hybrid cascade topology is 0.61723. To improve the regression value, we may retrain the network many times to update the initial weights and biases of the network. According to the value of the regression in hybrid cascade topology, $R = 0.61732$, and it can be seen that the hybrid cascade topology has a bad fitting model.

We can also increase the number of hidden layers or number of training vectors or using other algorithms for training.

The nonlinear relationships between electricity price output and the 16 input variables, including load data, natural gas price, crude oil price, temperature, humidity, and other factors, are important considerations.

Table I shows the weights assigned to multilayer neural networks in parallel cascade. There is no difference in interconnection sequence, because all the networks are in parallel and the final output is improved compared with the output of each network due to the averaging mechanism in the parallel analysis topology. A better outcome may be obtained by introducing a higher weight to the best performing network, such as $w2 = 0.95$ and $w4 = 1$. The higher weight is because these topologies contributed to improve hybrid parallel topology. Both of $w1 = 0.5$ and $w3 = 0.7$, the lowest weight is given to the network because the contribution of these topologies are lower and it makes the total, averaging of the parallel topology inefficient.

Table II offers a comparison of the hybrid cascade topology and hybrid parallel topology. In terms of training and simulating times, we conclude that the hybrid parallel topology has the lowest time in seconds. For instance, hybrid parallel topology takes the training time of 228.21 s and the simulating time of 86 ms. However, the hybrid cascade topology takes the training time of 912.82 s and the simulating time of 65 ms. In terms of accuracy, the hybrid cascade topology shows the mean square error (mse) of 87.4 and the hybrid parallel topology shows the mse error of 9.11. For the mean APE (MAPE), the hybrid parallel topology error is 3.25 compared with 12.27 for the hybrid cascade topology. We conclude that the hybrid parallel topology is efficient in comparison with the hybrid cascade topology.

Table III gives a better picture of which interconnection topology is the most efficient. If we take a training time and a simulating time, we can conclude that the parallel topologies have the lowest time in seconds to produce the most efficient solution. For instance, the cascade–parallel-in-cascade interconnected topology takes 480.26 s and in simulating time takes 250 ms to produce the optimum solution. However, in cascade–parallel-in-parallel interconnected topology takes just 100.48 s of training time and 84 ms of simulating time to do that. In terms of accuracy, it can be seen that the MAPE for the cascade–parallel-in-cascade

TABLE III
COMPARISON OF PERFORMANCE IN MULTILAYER TOPOLOGIES FOR ACTUAL AND PREDICTED ELECTRICITY PRICES FOR JANUARY 2006

Cascade	Parallel	Cascade-Parallel in Cascade	Cascade-Parallel in Parallel	Parallel-Cascade in Cascade	Parallel-Cascade in Parallel	Performance Indices
38.98	4.11	4.95	45.44	11.85	11.85	Mean Square Error (MSE)
6.93	1.42	1.54	7.66	3.76	3.76	Mean Absolute Percentage Error (MAPE %)
10 Iterations at the 6 validation checks	1000 Iteration at the 0 validation checks	181Iteration at the 0 validation checks	1000 Iteration at the 0 validation checks	11 Iteration at the 6 validation checks	1000 Iteration at the 0 validation checks	Maximum number of training Iteration (Epoch) at the validation checks
0.9439	0.9912	0.9898	0.9172	0.9745	0.9975	Coefficient of determination (R^2)
0.31	0.084	0.25	0.084	0.022	0.086	Simula tion Time (s)
102.91	50.32	480.25	100.48	127.65	204.43	Traini ng Time (s)

TABLE IV
RANKING TRAINING ALGORITHMS OF FEED FORWARD NEURAL NETWORK WITH TEN HIDDEN LAYERS IN HYBRID PARALLEL TOPOLOGY IN THE CASE OF ACCURACY

#	Algorith m	Training Time seconds	Operation Time seconds	Mean Square Error	Coefficient of determination ($\boldsymbol{R^2}$)	Mean Absolute Percentage Error (%)
1	LM	276.03	0.086	9.11	0.9815	3.25
2	BFG	237.06	0.083	14.95	0.9596	4.74
3	OSS	21.27	0.085	15.21	0.9554	4.93
4	RP	21.29	0.093	15.65	0.9536	5.07
5	SCG	18.32	0.092	22.30	0.9355	7.12
6	GDX	21.09	0.088	30.05	0.9032	10.39
7	GDM	13.15	0.091	50.21	0.7853	18.75

topology has 1.54%. However, the cascade–parallel-in-parallel topology shows 7.66%, and also the parallel–cascade-in-cascade and parallel–cascade-in-parallel topologies show 3.76%. Starting and ending with the cascade topology can enhance the total performance. In addition, starting with parallel and ending with parallel and cascade will not improve the performance according to MAPE = 3.76% in both the parallel–cascade in cascade and the parallel–cascade in parallel. It is affirmed that the parallel topologies have the least MAPE. Moreover, starting and ending with the parallel topology can minimize the training time. For instance, parallel–cascade-in-cascade topology requires less training time, 127.65 s, than the parallel–cascade-in-parallel topology, which required 204.43 s.

Table IV shows an evaluation of each network topology in the hybrid topology based on the actual and predicted electricity prices for January 2006; three different feedforward networks are trained. Each network has six, seven, and eight hidden layers, respectively.

In terms of accuracy, the Levenberg–Marquardt algorithm (LM) yields the most accurate performance exhibited by mse = 9.11 and MAPE = 3.25 compared with Gradient descent with momentum and adaptive learning rate backpropagation (GDX), showing mse = 10.39 and MAPE = 30.05%. In the case of operating time, the LM takes 276.03 s to compute the optimum solution compared with the variable learning rate backpropagation (GDX), which has the minimum operating time of 21.09 s. The gradient descent with momentum backpropagation (GDM) is the least efficient algorithm in terms of both accuracy and operating time.

Table V shows an evaluation of each network topology in hybrid topology based on the actual and predicted electricity price for January 2006, three different feedforward networks is trained. Each network has six, seven, and eight hidden layers, respectively. In terms of accuracy, the LM scores the most accurate performance exhibited by mse = 87.14 and MAPE = 12.27% compared with the resilient backpropagation, mse = 120.15 and

TABLE V
RANKING TRAINING ALGORITHMS OF FEEDFORWARD NEURAL NETWORK WITH TEN HIDDEN LAYERS IN HYBRID CASCADE TOPOLOGY IN THE CASE OF ACCURACY

Algorithm	Training Time seconds	Operation Time seconds	Mean Square Error	Coefficient of determination (R^2)	Mean Absolute Percentage Error
LM	509.74	0.086	87.14	0.9420	12.27
SCG	27.72	0.092	90.71	0.9154	16.15
OSS	33.51	0.085	100.86	0.8754	17.43
BFG	505.41	0.083	112.89	0.8356	25.12
RP	31.77	0.093	120.15	0.7785	27.05
GDX	28.38	0.088	127.34	0.7037	40.32
GDM	15.17	0.091	222.04	0.2300	105.75

MAPE $= 27.05\%$. The variable learning rate backpropagation (GDX) shows high error. In terms of execution time, the LM takes 509.74 s in training to compute the optimum solution compared to scaled conjugate gradient backpropagation, which has the lowest operating time of 18.32 s. GDM is the most inefficient algorithm in both accuracy and operating time.

The cascade-forward-net structure is less efficient and therefore, is not discussed further. We decided to set fit net and feedforward to use the different three multilayer nets and it seems to be working efficiently. But we are looking for more enhancement of neural network, so we set the feedforward for the three layers, which give the most efficient results with a small improvement in minimizing the errors with less training time and simulating time. We also attempted to increase the number of neurons, but this did not improve the performance of the multilayer neural network, so we have tried to change the number of hidden layers to two for each network and ten for each network; however, it achieved slightly higher error. We implemented several algorithms with backpropagation neural network and the best one is the Bayesian regulation backpropagation algorithm, but it takes a long time to complete training. We set feedforward network in the first and second layers and fit network in the third layer; however, we obtained less training time with some outliers of errors, which does not make our option as efficient as we expected. We tried to set the first two layers as a fit net and the third one as a feedforward net, but mse is high with good results in parallel connection in regression and mean absolute percentage performance and also less time training.

After testing all aforementioned options, we come up with the best option, we decided to set the LM for both cascade and parallel training functions to enhance the performance. Also we set Bayesian regulation backpropagation algorithm for the first net and the LM for the second Net 2 and third Net 3 to reduce the accumulated error.

Moreover, we set six neurons for Net 1, seven neurons for Net 2, and eight neurons for Net 3. We finally conclude that these options give the most accurate performance especially in hybrid parallel topology.

IV. CONCLUSION

This paper has demonstrated a new computational approach to applying composite backpropagation multilayer neural networks to forecast hourly electricity price for the next month based on hourly important factors, such as previous hourly load, hourly natural gas, and hourly weather conditions for electricity load forecasting for the Australian market in January 2006. We interconnected three separate nets in a cascade topology and a parallel topology to acquire the best performance based on simulation results.

A comparison of the cascade and parallel alone topology results in the performance superior to that of the parallel topology. A comparison of the cascade–parallel-in-cascade topology with the cascade–parallel-in-parallel topology reveals that starting the topology by a cascade connection and ending with a cascade connection results in the performance superior to that of starting the topology by a parallel connection and ending with a parallel connection. Therefore, the cascade–parallel-in-cascade topology is superior to that of the cascade–parallel in parallel. From a comparison of the performance of the cascade–parallel in parallel with that of parallel–cascade in parallel, we concluded that starting the topology with parallel connection yields superior performance to that starting with the cascade connection. Likewise, we found that there is no difference in accuracy between a parallel–cascade-in-cascade connection and parallel–cascade-in-parallel connection. From the earlier experiment, it could be reasoned that the performance of the cascade–parallel-in-cascade topology is superior to all different aforementioned topologies.

Following the analysis of the performance of each topology, it is concluded that the parallel topology, which averages the output errors, improves the performance of the systems. On the other hand, the cascade topology accumulates the error from each training stage, which mostly improves the overall performance in the systems with high data distribution.

REFERENCES

[1] N. Amjady and A. Daraeepour, “Mixed price and load forecasting of electricity markets by a new iterative prediction method,” *Electr. Power Syst. Res.*, vol. 79, no. 9, pp. 1329–1336, 2009.

[2] A. Karsaz, H. R. Mashhadi, and M. M. Mirsalehi, “Market clearing price and load forecasting using cooperative co-evolutionary approach,” *Int. J. Elect. Power Energy Syst.*, vol. 32, no. 5, pp. 408–415, 2010.

[3] R. Weron, “Electricity price forecasting: A review of the state-of-the-art with a look into the future,” *Int. J. Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014.

[4] J. P. S. Catalao, S. J. P. S. Mariano, V. M. F. Mendes, and L. A. F. M. Ferreira, “An artificial neural network approach for short-term electricity prices forecasting,” presented at the Int. Conf. Intell. Syst. Appl. Power Syst., Nov. 2007, pp. 1–6.

[5] V. Iyer, C. C. Fung, and T. Gedeon, “A fuzzy-neural approach to electricity load and spot-price forecasting in a deregulated electricity market,” Presented at the Conf. Convergent Technol. Asia-Pacific Region (TENCON), Oct. 2003, pp. 1479–1482.

[6] K. B. Sahay and M. M. Tripathi, “Day ahead hourly load and price forecast in ISO New England market using ANN,” in *Proc. Annu. IEEE India Conf. (INDICON)*, Dec. 2013, pp. 1–6.

[7] G. Li, C.-C. Liu, J. Lawarree, M. Gallanti, and A. Venturini, “State-of-the-art of electricity price forecasting,” presented at the Int. Symp. CIGRE/IEEE PES, Oct. 2005, pp. 110–119.

[8] A. Karsaz, H. R. Mashhadi, and R. Eshraghnia, “Cooperative co-evolutionary approach to electricity load and price forecasting in deregulated electricity markets,” presented at the IEEE Power India Conf., 2006.

[9] D. W. Bunn, “Forecasting loads and prices in competitive power markets,” *Proc. IEEE*, vol. 88, no. 2, pp. 163–169, Feb. 2000.

[10] P. L. Skantze and M. Ilic, *Valuation, Hedging and Speculation in Competitive Electricity Markets A Fundamental Approach*. Boston, MA, USA: Kluwer, 2001.

[11] M. Shahidehpour, H. Yamin, and Z. Li, *Market Operations in Electric Power Systems: Forecasting, Scheduling, and Risk Management*. New York, NY, USA: Wiley, 2002.

[12] V. S. Koritarov, “Real-world market representation with agents,” *IEEE Power Energy Mag.*, vol. 2, no. 4, pp. 39–46, Jul./Aug. 2004.

[13] M. Ventosa, Á. Baíllo, A. Ramos, and M. Rivier, “Electricity market modeling trends,” *Energy Policy*, vol. 33, no. 7, pp. 897–913, 2005.

[14] N. Amjady and M. Hemmati, “Energy price forecasting—Problems and proposals for such predictions,” *IEEE Power Energy Mag.*, vol. 4, no. 2, pp. 20–29, Mar./Apr. 2006.

[15] R. Weron, *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Chichester, U.K.: Wiley, 2006.

[16] H. R. Zareipour, *Price-Based Energy Management in Competitive Electricity Markets: Price Forecasting and Optimal Operation of Wholesale Customers*. VDM Publishing, 2008.

[17] N. Amjady, “Short-term electricity price forecasting,” in *Electric Power Systems: Advanced Forecasting Techniques and Optimal Generation Scheduling*, J. P. S. Catalão, Ed. Boca Raton, FL, USA: CRC Press, 2012.

[18] C. García-Martos and A. J. Conejo, “Price forecasting techniques in power systems,” in *Wiley Encyclopedia of Electrical and Electronics Engineering*. 2013, pp. 1–23. [Online]. Available: http://dx.doi.org/10.1002/047134608X.W8188

[19] T. Hong, “Energy forecasting: Past, present, and future,” *Foresight, Int. J. Appl. Forecasting*, no. 32, pp. 43–48, Winter 2014.

[20] C. P. Rodriguez and G. J. Anders, “Energy price forecasting in the Ontario competitive power system market,” *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 366–374, Feb. 2004.

[21] R. Carmona and M. Coulon, “A survey of commodity markets and structural models for electricity prices,” *Quantitative Energy Finance: Modeling, Pricing, and Hedging in Energy and Commodity Markets*, F. E. Benth, V. A. Kholodnyi, and P. Laurence, Eds. Springer.2014.

[22] (2015). *The Australian Energy Market Operator (AEMO)*. [Online]. Available: http://www.aemo.com.au/Electricity/Planning/Forecasting

Hossam Mosbah was born in Libya in 1984. He received the B.Sc. degree in electrical engineering from the College of Engineering Technology, Houn, Libya, in 2007, and the M.A.Sc. degree in electrical engineering from Dalhousie University, Halifax, NS, Canada, where he is currently pursuing the Ph.D. degree.

His current research interests include power system state estimation, and state estimation and optimization algorithms.

Mohamed El-Hawary (S’68–M’72–SM’76–F’90–LF’09) received the bachelor’s degree (Hons.) in electrical engineering from Alexandria University, Alexandria, Egypt, in 1965, and the Ph.D. degree from the University of Alberta, Edmonton, AB, Canada, in 1972.

He was a Killam Memorial Fellow with the University of Alberta. He has been a Professor of Electrical and Computer Engineering with Dalhousie University, Halifax, NS, Canada, since 1981. He served as a Faculty Member, and was the Chair of the Electrical Engineering Department with the Memorial University of Newfoundland, NL, Canada, for eight years. He was an Associate Professor of Electrical Engineering with the Federal University of Rio de Janeiro, Rio de Janeiro, Brazil, for two years. He was an Instructor with the University of Alexandria. He pioneered many computational and artificial intelligence solutions to problems in the economic/environmental operation of power systems.