Electricity Price Short-Term Forecasting Using Artificial Neural Networks

B.R. Szkuta, L.A. Sanabria, T.S. Dillon
Applied Computing Research Institute
La Trobe University, Melbourne, Australia

Abstract: This paper presents the System Marginal Price (SMP) short-term forecasting implementation using the Artificial Neural Networks (ANN) computing technique. The described approach uses the three-layered ANN paradigm with back-propagation. The retrospective SMP real-world data, acquired from the deregulated Victorian power system, was used for training and testing the ANN. The results presented in this paper confirm considerable value of the ANN based approach in forecasting the SMP.

Keywords: Artificial Neural Networks, System Marginal Price, Electricity Market.

1. Introduction

World wide reforms of the power industry in the last 10 years have opened the electricity markets to alternative producers and alternative purchasers. As a result of these, the electricity supply industry general model, made up of a number of individual generators supplying energy to a number of individual distributors through the transmission network, is operated by an Independent System Operator (ISO). The generators bid quantities of energy to the ISO at their prices. Based on the generators bids, the ISO dispatches blocks of energy to match the system demand at a minimum cost subject to a number of constraints related mainly to the system security. The ISO is responsible for ensuring that schedules for use of the transmission system are feasible, operating the transmission system in real time, and settling financially with parties who use the transmission system. As the ISO serves also as a daily spot market operator, one of the ISO services rendered to the market is the electricity SMP forecasting. In the deregulated Victorian power system, which has provided our model with the training and testing data sets, the effective energy unit price is set half-hourly by the last generator (highest successful bid) selected to run [3].

An up-front proper assessment of the energy price is of prime importance to the deregulated electricity industry. Decisions on transmission augmentation, generation expansion, distribution planning and energy exchange between regions are largely determined by the *long-term* forecast of the electricity price. Sophisticated models, to simulate the system operation under a large variety of conditions for predicting both SMP and delivery costs, have been developed [4], [5]. The energy trading levels between market participants, however, is largely dependent on the *short-term* price forecasts.

In the Victorian wholesale electricity market, three levels of trading are recognised: bilateral long-term contracts, short term forward trading and spot trading [6]. The bulk of trading is based on a combination of the three. In long-term contracts, the sellers and buyers contract the supply of fixed amounts of energy for specified periods under set prices. The contract must include provisions for payment of the price difference between SMP and set price. Short-term forward trading is used to complement long-term contracts and to lower exposure to variations of energy price around the contracted price. Price for the short-term purchases depends on the offer-demand energy price. This option, however, is usually available in the one or two days ahead trading. The spot trading is used to cover energy included in neither, the long-term nor short-term contracts, and the price of this energy is effectively the SMP. Based on the *short-term* SMP forecast, market analysts decide on the most profitable way of meeting obligations not covered by the long-term contracts. The main tokens in the game are reduction/increase of generation or demand, buying energy in either short-term forward market or in the spot market.

The Victorian generation capacity of about 7000MW is based largely on brown coal, with important contributions from natural gas, Victoria's local hydro generation and a share of the Snowy Mountains hydro scheme. Number of the market participants does not exceed 10.

The Victorian electricity market weekly turn-over during the summer of 1997 reached AU$60.8 million in the third week of January [7]. It is obvious that the *short-term* SMP forecast directly influences the energy trading market and any improvement in the forecast may significantly change the trading scenario of a market participant with tangible financial benefits.

2. Problem Description

In general, the commodity market prices are compelled by the supply and demand relationship. In the electricity market where the traded commodity cannot be stockpiled, the constraints are defined by the system total capacity to satisfy the demand at any

PE-149-PWRS-0-08-1998 A paper recommended and approved by the IEEE Power System Analysis, Computing and Economics Committee of the IEEE Power Engineering Society for publication in the IEEE Transactions on Power Systems. Manuscript submitted December 2, 1997; made available for printing August 14, 1998.

given time during any day of operation. The reserves, in turn, accommodate the system ability to fulfil the uninterrupted supply commitments. Such scenarios make the constraints binding. The total supply to the local electricity market is an aggregate of the local power generation and the interstate power flow. The demand forecast, together with a knowledge of the existing reserves, or an *intelligent guess* on their expected level (normally, such information would not be published by the ISO), determines the supplier's prospects of selling a block of energy by an auction process.

To buy an access to the network, besides the timing and energy amount, a generator also quotes its price. In other words, the closer the bid price and the final energy market values are, the better are the bidder's chances of getting the 'job' without compromising commercial gains. This implies that apriori knowledge of the energy price must be of prime importance to the business performance of the electricity market participants.

In general, the anticipated SMP value is strongly tied to the generators bidding pattern, which in turn is dependent on the anticipated System Power Reserves (SPR) and System Potential Demand (SPD). The neural network model should include these parameters in order to capture this effect. In other words, the SPD and the SPR decide the generator's bidding strategy and, eventually, shape the behaviour of the spot market SMP. However, it is anticipated that SPD and SPR would influence a generator's bidding not only in the current but also in future time intervals.

In this paper we limit this influence to two time intervals ahead (time lag). The input data temporal set, used for training the ANN, that contains retrospective values of variables is encapsulated in the vector **X(t)**, see definition (1). All input parameters, therefore, including the SMP(**t**), are known for all the time intervals **t** included in the training.

Definition (1)

X(t) = {SPD(t-2), SPD(t-1),SPD(t), SPR(t-2), SPR(t-1),SPR(t),SMP(t-2),SMP(t-1), SMP(t)} 6 calendar related parameters.
Where t ranges from 3 (time lags) to 48, spanning 203 epochs (days).

The temporal relationship between all the training variables was encapsulated in the input data tuples. The model training variables, therefore, comprised (9+6+1)-tuple time series (see the equation 3), where 9 of the components represented the data inputs, 6 components represented the relevant calendar input parameters (day type, holiday code, month, DST/EST clock change, Xmas code, time of year) and 1 represented the desired output.

In order to validate our selection of the input variables and assure their significance to the SMP 'prediction' process, a sensitivity test has been carried out on the model by varying one variable at a time, whilst holding the others constant. The results of this are presented in Section 4.

After the selection of the input parameters has been validated, the 15-15-1 network model has been constructed (ie. 15-input parameters, 15-hidden layer units, and 1- output).

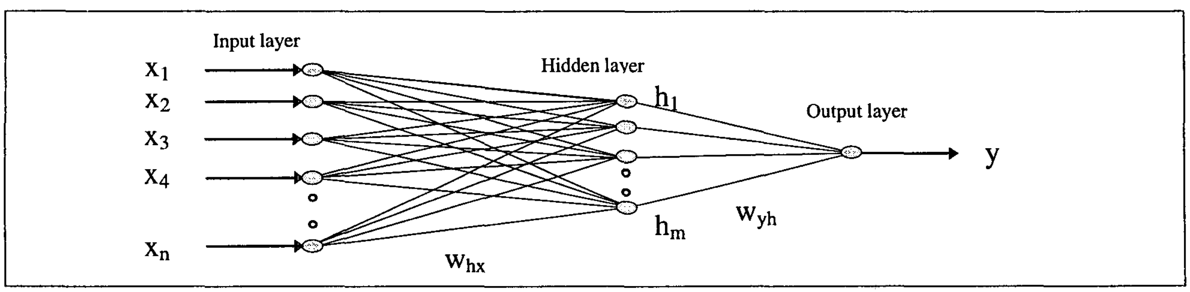

Fig.1. Generic example of a 3-layered feed-forward ANN model with **n** input, **m** hidden and **1** output node.

3. Artificial Neural Networks

The term ANN is used to describe various constructs of highly interconnected simple processing units that deliver an alternative to conventional computing techniques. The difference from the traditional approach is that they represent the related objects through learning from sample data rather than through modelled computing procedures. The ANN have been applied successfully to a variety of problems.

The techniques based on ANN [1], [2] are especially effective in the solution of high complexity problems for which a traditional mathematical model is difficult to build (where the nature of the input-output relationship is neither well defined nor easily computable). Problems of forecasting, noise filtering, pattern recognition, credit-worthiness assessment, etc., fall into such a category. In power systems the ANN have already been used to solve problems such as load forecasting, component and system fault diagnosis, security assessment, unit commitment, etc. [8].

The most popular ANN architecture is a three-layered, feed-forward system with back-propagation, (see Fig.1). The success of this configuration dwells in the fact that it can learn from retrospective information in a process called supervised learning. In supervised learning the network is trained using historical data derived from the system the relationship between input and output is to be determined.

A set of data (*training time series*) is presented to the network a number of times (*training cycles*), and its output is compared with some known output or target. The epoch Mean Square Error (MSE_E), (see equation (2)), a measure of distance between the network output (O_k) and the known target (D_{Ek}), is back propagated to the input layer, and the process repeated.

$$MSE_E = 1/2 \sum_{\forall k} (D_{Ek} - O_k)^2 \qquad (2)$$

$$MSE_A = 1/2 \sum_{\forall E} \sum_{\forall k} (D_{Ek} - O_{Ek})^2 \qquad (3)$$

where E - epoch presented to the ANN,
k - layer index.

The new output is again compared with the target and the new error is back-propagated. At each iteration, the network changes its weights in order to minimise the error. The error minimisation process is repeated until the error converges to a predefined small value. The overall MSE_A is given by the average mean square error (3).

In training, the network learns by adjusting the weights connecting the input and hidden layer and the weights connecting the hidden layer and output, by the gradient multiplied by the *learning rate* parameter.

In order to accelerate the learning process, two parameters of the back-propagation algorithm, the *learning rate* and another parameter, *momentum*, can be adjusted [9]. Both affect the error gradient minimisation process. The *learning rate* is the proportion of error gradient by which the weights should be adjusted. Large values can give a faster convergence to the gradient minimum but also may produce oscillation around the minimum, hence positive values lower than 1 are recommended. The *momentum* determines the proportion of the change of past weights that should be used in the calculation of the new weights.

The major advantage of ANN is that the training, which is the most time consuming exercise, is done off-line.

4. Model Sensitivity Test For Input Variable Selection Validation

The ANN based SMP model sensitivity test, on some of the parameters chosen for inclusion, has been carried out. The purpose of the test was to validate the preconception that the selected input variables actually influence the object.

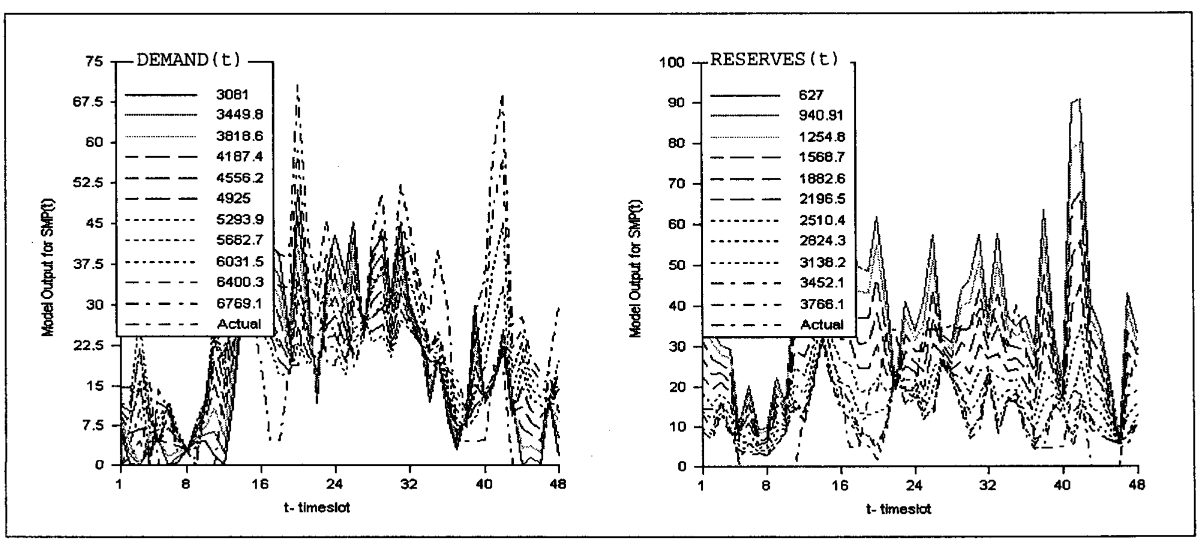

Fig. 2. SMP as a function of the current SPD and SPR (model sensitivity test for an arbitrarily selected epoch).

The plots in Fig.2 to Fig.4 are for illustration only and are shown here to demonstrate the SMP sensitivity to SPD, SPR and to previous values of the SMP itself. The authors understand that, due to the density of the represented data, the plots may look unreadable. However, to sustain the choice of input parameters, they decided to show them.

The plots in Fig.2 demonstrate a high variance of the SMP when related to the immediate SPD and SPR values. Particularly notable is the SMP strong dependence on the SPR.

Retrospective knowledge of the SPD and SPR have also a decisive impact on the current SMP, as demonstrated by the plots in Fig.3.

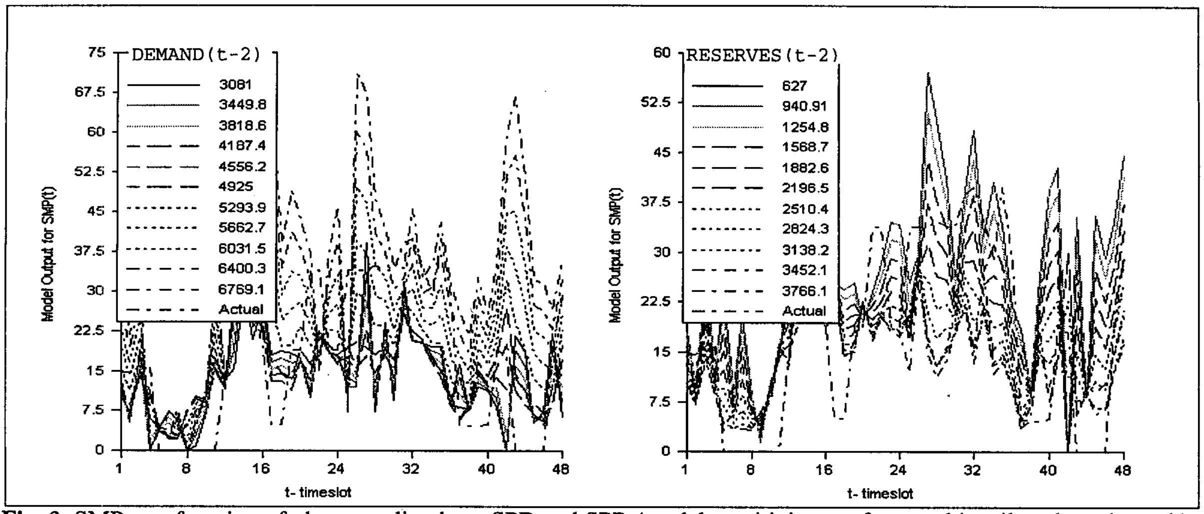

Fig. 3. SMP as a function of the preceding hour SPD and SPR (model sensitivity test for an arbitrarily selected epoch).

The plots on Fig.4 demonstrate the high volatility of the SMP when related to its own preceding values. In general, the results obtained demonstrate that all these variables do indeed influence the SMP.

For some time-slots, the sensitivity becomes low, thus showing negligible significance of the input to the object tested. However, the epochs are wholly presented to the ANN and the low sensitivity time-slots cannot be excluded from the model.

The important information for the ANN training is carried in the periods where the sensitivity is high. The high sensitivity periods reflect the market participants bidding behaviour, who would enter the electricity market auction at established times of an epoch and would bid high when the reserves are low or when the demand is high, or both.

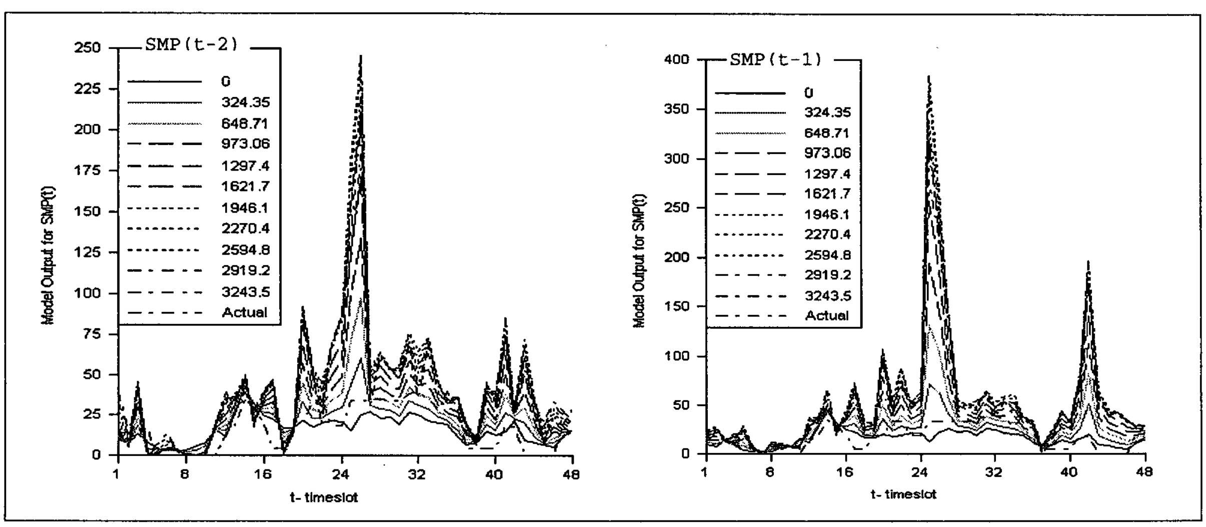

Fig. 4. SMP as a function of two preceding half-hourly values (model sensitivity test for an arbitrarily selected epoch).

5. Neural Network Training

In a series of experiments, the multivariate time series was presented to the network and the network was then made to 'predict' the next value in the time sequence. The comparison of the just 'predicted' value and the known value for the processed time-slot, resulted in a value of the calculated mean square error (MSE) which was then propagated back through the feed-forward links. The back-propagation algorithm globally modified the weights of all internal links to minimise the error. If the error gradient still exceeded a predefined small value, a new epoch was started. Finally, after the training phase has processed all epochs, its performance could be judged in terms of the globally calculated MSE. After all epochs of the training phase were processed the network was ready for prediction and generalisation.

The back-propagation learning algorithm, as mentioned earlier, is characterised by two parameters, the *training rate* and the *momentum.* These were used by the process of progressive *refining* weights based on the RMS error minimisation.

One of the objectives was to find the optimal number of the training cycles performed on the training data. To do this, the network was subjected to a number of training sessions with varying number of training cycles. After each experiment the network was tested for its ability to correctly classify the test data. The experiments included training sessions of 40,000 through 200,000 training cycles.

With the increased number of the applied cycles, the results gradually showed improvement in the network ability to 'predict' the input parameter used in training. However, such better performance also indicated the network's symptoms of over-fitting the training data, that is, of being over-trained. The number of 80,000 training cycles has been elected as not compromising the network ability to correctly generalise on new (previously unseen) input parameters and the network was considered ready for generalisation.

The data-driven model included temporal data gathered from the Victorian power system for a period of over eight months, ranging from 23 Oct 96 to 13 May 97. The input data time series included half-hourly values of the SPD, SPR and SMP. The training data time series, therefore, consisted of 203 epochs (days) containing 48 time-slot values each. Further 7 epochs, ranging from 14 to 20 May 97, were used for generalisation .

6. Discussion Of Results

The real data that was used for the ANN training and testing was derived from the Victorian power system. The training data spanned very hot summer, mild autumn and relatively ordinary winter; the daily temperatures reached 40 degrees in summer (when the SMP exceeded $3,000 per MWh) and 29 degrees in autumn. The nightly temperatures were also respectively high. These have caused large variations in the power demand and, consequently, a large variation of the SMP. The data used in the experiments reflected the variation of both the SMP and SPD, thus making the forecasting task even more challenging for the ANN data-based model.

The ISO plans the system reserves, to achieve the required generation capacity, and also regularly produces the system demand forecasts. The predicted system demand and planned reserves are presented to the trained ANN to produce the SMP forecast values. The plots presented in Fig.5 show individual days (epochs) of the weekly forecast produced by the ANN after the 80,000-cycle training session. As demonstrated by the plots, despite volatility of the actual SMP daily curves, the ANN generated forecasts appeared to fit the curves relatively closely (measured by the relatively low RMS_E error level). Table 1 presents values of several statistical indicators calculated individually for each epoch of the weekly forecasts.

The RMS_E, that measures the distance between the ANN output and the known target for an epoch E, was calculated using equation (4).

$$RMS_E = 1/N_E \sqrt{\sum_{k}^{N_E} (Pred_k - Targ_k)^2} \qquad (4)$$

Where: N_E - number of time-slots per epoch E,
$Pred_k$ - time-slot k predicted value,
$Targ_k$ - time-slot k target value.

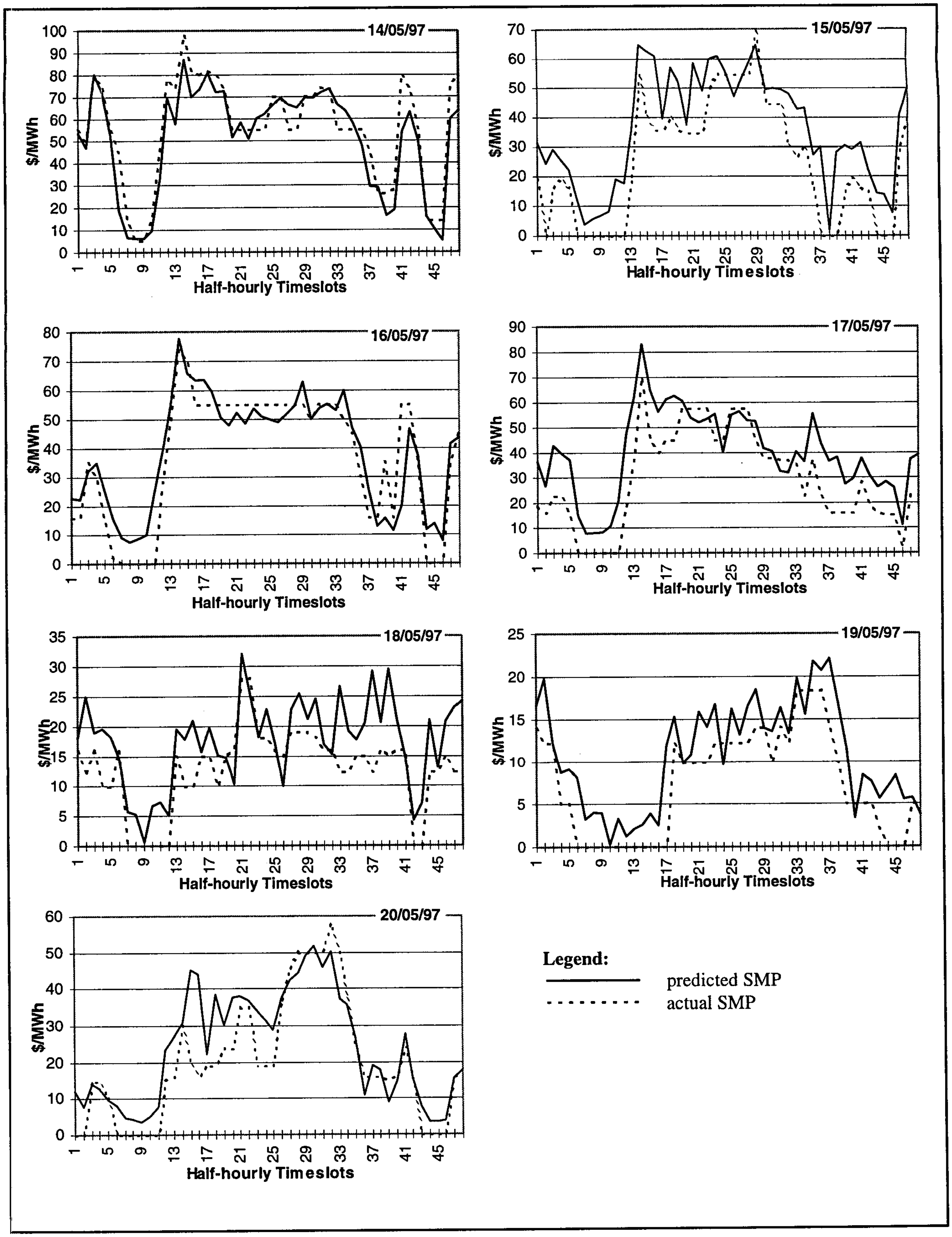

Fig. 5. Summary of the resultant plots.

Epoch Date	14/05/97	15/05/97	16/05/97	17/05/97	18/05/97	19/05/97	20/05/97
Min-Max (Actual SMP)	5.04-98.12	0.00-70.00	0.00-73.90	0.00-70.00	0.00-28.00	0.00-18.34	0.00-57.62
Min-Max (Predicted SMP)	5.29-87.16	2.00-64.92	7.51-77.82	7.81-83.05	0.60-32.07	0.30-22.10	3.38-51.90
Average Error	4.16	-11.09	-2.18	-10.61	-4.88	-3.27	-4.17
RMS Error	1.33	1.98	1.38	2.00	0.99	0.64	1.27
Error Standard Deviation (σ)	8.25	8.10	9.32	8.90	4.84	2.95	7.73
Min-Max (APE[1])	0.56-58.06	1.80-84.02	0.12-671.02	1.98-140.93	0.60-11.31	0.25-179.60	0.98-175.59

Table 1. Statistical analysis of the ANN errors ($/Mwh). (1) Absolute Percentage Error.

7. Conclusions

Notwithstanding the fact that the achieved results were very satisfactory, the authors acknowledge that the training data set was relatively small which, together with the large daily variation in SMP, made the ANN model rather difficult to evaluate. The model response appeared highly sensitive to the training parameters like momentum, training rate and also to the number of training cycles applied. The true nature of this sensitivity, in reference to the presented model have not been investigated as yet. This, however, together with the impact of other input parameters (not included, like market bidding patterns, generation planned outages, etc.) are planned to be the subject of a follow-up work.

The ANN, where the RMS error minimisation implies approximation of the target object, are used extensively to learn and predict time series. The authors understand that such implementations aim at making only point predictions. The authors also acknowledge, that if noise characterises the model, as for the presented model it certainly did, the forecasting may become of a lesser value unless the confidence level associated with it can be predicted as well.

The Victorian electricity market, which started operation in late 1994 and still undergoes changes, is considered relatively new. Until the participants ascertain their operational strategies, the market may remain subject to a wide variation in the electricity price and, therefore, the relevant data model will contain varying levels of distorting noise. It is expected, however, that as soon as the market reaches reasonable stability, the presented ANN based SMP forecasting model shall deliver less volatile results.

The Victorian ISO publishes the weekly SMP forecasts in the TV teletext pages. The official forecasts are produced by methods based on conventional linear regression techniques. A comparison of the official results to the forecast made with ANN, related to the same periods, has clearly demonstrated the ANN's advantage. The space limitation of the paper, however, does not allow the authors to present the statistical analysis of such comparisons, or to show the relevant plots.

8. References

[1] Vemuri V.R. and Rogers R.D. "Artificial Neural Networks. Forecasting Time Series". IEEE Computer Society Press. 1994.
[2] Baumann T., Thurner E. and Deisenhofer D. "Real World Implementation of a Short-term Load Forecasting Method using Modular Neural Networks". Proc 12th PSCC. Dresden, August 19-23, 1996, pp 231.
[3] VICPOOL. Victoria's Competitive Electricity Market. Victorian Power Exchange (VPX). May 1995.
[4] Willis L., Finney J. and Ramon G. "Computing the Cost of Unbundled Services". IEEE Computer Applications in Power. Vol 9, No. 4, October 1996.
[5] Rivier M., Perez-Arriaga I.J. and Luengo G. "JUANAC: A Model for Computing of Spot Prices in Interconnected Power Systems". Proc 10th PSCC. Graz, Austria. August 1989, pp 254.
[6] NEM/VICPOOL. "Market Simulation Workshop". Victorian Power Exchange (VPX). March 1996.
[7] VICPOOL Information Bulletin. Vol 4, No. 4. VPX. January 1997.
[8] Dillon T.S. and Niebur D. "Neural Networks Applications in Power Systems". CRL Publishing Ltd. London 1996.
[9] Smith M. "Neural Networks for Statistical Modelling". Van Nostrand Reinhold. New York 1993.