The Sixth International Power Engineering Conference
(IPEC2003), 27-29 November 2003, Singapore

Support Vector Machine Based Electricity Price Forecasting For Electricity Markets utilising Projected Assessment of System Adequacy Data.

D. C. Sansom (sansom@itee.uq.edu.au) , **T. Downs and T. K .Saha**

School of Information Technology and Electrical Engineering
University of Queensland

Abstract

In this paper we present an analysis of the results of a study into wholesale (spot) electricity price forecasting with Support Vector Machines (SVM) utilising past price and demand data and Projected Assessment of System Adequacy (PASA) data. The forecasting accuracy was evaluated using Australian National Electricity Market (NEM), New South Wales regional data over the year 2002. The inclusion of PASA data shows little improvement in forecasting accuracy.

Keywords
Price Forecasting, Support Vector Machines, electricity markets.

1 INTRODUCTION

Electrical Supply Industries (ESI) worldwide have been restructured (deregulated) with the intention of introducing levels of competition into energy generation and retail energy sales. In any market with levels of competition information of future market conditions can contribute to giving market participants a competitive advantage over their fellow market participants.

In an open auction style electricity market such as the Australian National Electricity Market (NEM) [1] a large volume of information on historical and predicted market conditions is available to all market participants. As the ESI is a large volume industry all market participants can gain advantages from even a small increase in the accuracy of their electricity price forecasts.

As Electrical Power Engineers with experience in electrical load forecasting [2] a logical starting place for electricity price forecasting was to utilise the same methods as we used for load forecasting. This provided a fruitful starting place as variations in electricity price depends on and so mirrors the variations in electrical demand[3, 4]. However electricity prices are far more volatile than electrical demand as prices are also a market function of supply and demand.

Electrical loads vary in a stable periodic way with seasonal and climate variations and weekly and daily human activity patterns. Thus loads could be forecasted by utilising a knowledge of these periodic variations however the electrical load forecasts can be improved by including predictions of future weather data. Loads could be forecasted by examining only past demand data however the forecast can be improved by considering a wider range of data.

Electricity prices are based on the demand and so also vary in similar stable periodic ways as the demand however as the NEM regional electricity price is determined in an auction style market based on the economic principles of supply and demand. From economic principles we hypothesis that electricity prices would be influenced by the difference between available supply and the required demand at each instant in time. In previous studies we have only utilised past demand and price data [5, 6,14] in this research we hope to improve our price forecasts by utilising a wider range of data.

Some data that gives an indication of future available supply or generation capacity and the projected required demand is found in the short-term PASA files provided to market participants by NEMMCO [13].

The results of these tests are being used to investigate the following hypothesis, over the period tested the electricity price forecasting accuracy for the NSW regional electricity price will be improved by the inclusion of the PASA data variables into the input data set presented to the SVM forecasting model.

2 SHORT-TERM PASA DATA

The short-term Projected Assessment of System Adequacy Data files are produced for the NEM by NEMMCO every two hours. The files contain projected half-hourly data for the next six days starting at 04:30

the day after the PASA file was published. In this research the data variables utilised from the PASA data are:

1) projected capacity required
2) projected reserve required
3) projected reserve surplus
4) projected regional demand 10% POE
5) projected regional demand 50% POE
6) projected regional demand 90% POE

where POE is probability of being exceeded.

The projected capacity required is an approximation of the total regional generation capacity that is required for that half-hour. The capacity required is equal to the 10% POE regional demand forecast plus the set reserve required.

The projected reserve required is the Minimum level of reserve required in the region as determined by the Reliability panel. Usually set at approximately 5 to 7% of the expected total regional demand. Through out the majority of the period in this study the reserve required was set at 660MW for the NSW region, which has a total demand from 7000 to 11000MW.

Projected reserve surplus is the surplus (positive value) or deficiency of available reserve (negative value) compared to the capacity required.

The 10% POE regional demand forecast is the regional demand forecast produced with a 10% probability of being exceeded (POE). Similarly the 50% POE forecast has a 50% chance of being less than the actual demand at that half-hour.

3 SUPPORT VECTOR MACHINE THEORY

With the goal of reducing the time and expertise required to construct and train price forecasting models we considered the next generation of NNs called support vector machines (SVM). SVM have fewer obvious tuneable parameters than NNs and the choice of parameter values may be less crucial for good forecasting results. The SVM is designed to systematically optimise its structure (tune its parameter settings) based on the input training data. The Training of a SVM involves solving a quadratic optimisation, which has one unique solution and does not involve the random initialisation of weights as training NN does. So any SVM with the same parameter settings trained on identical data will give identical results. This increases the repeatability of SVM forecasts and so greatly reduces the number of training runs required to find the optimum SVM parameter settings when compared to NN training.

The following explanation of SVM is the combination of information from sources [7] [8], more information regarding SVMs can be obtained from the kernel machines web site[9].

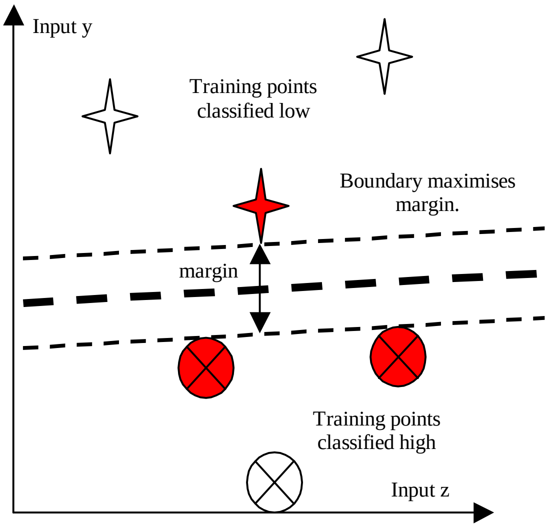

Figure 1 Maximum Margin of Support Vector Machine

To explain the principles of SVM we begin with an explanation of the application of a SVM to classify data points as high or low in a two dimensional input space. The basic principal of SVM is to select the support vectors (shaded data points) that describe a threshold function (boundary) for the data that maximises the classification margin (as in Figure 1) subject to the constraints that at the support vectors the absolute value of the threshold function must be greater than one as in Equation 1 (see Figure 2). The non-support vector data points (unshaded points) do not effect the position of the boundary.

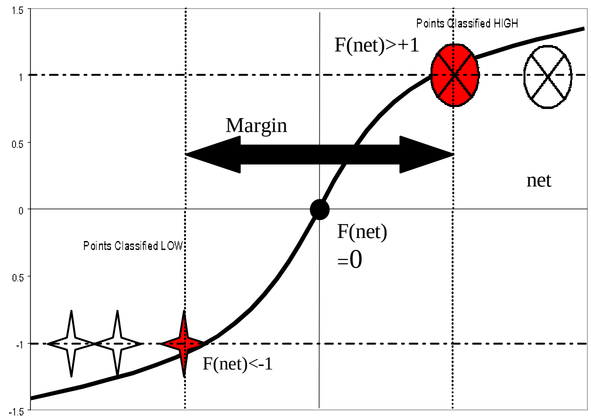

Figure 2 Threshold function for SVM

Equation 1 optimisation to minimise margin

$$\text{minimise} \quad F(W) = \frac{1}{2}(W \bullet W^T)$$
$$\text{subject to} \quad y_k(W \bullet X_k + b) - 1 \geq 0$$
$$\text{for data points} \quad k = 1,..., l$$
$$\text{where} \quad y_k \quad \text{is target of data point k}$$

To overcome the limitation that the SVM only applies to linearly separable systems the inputs (X_k) are mapped through a transform function Φ(X) into a higher dimensional space where the system is linearly separable. This can be understood with the help of the very simple example in Figure 3 where the one-dimensional system is not linearly separable however if the system is mapped by a dot product into two-dimensional space the system becomes linearly separable.

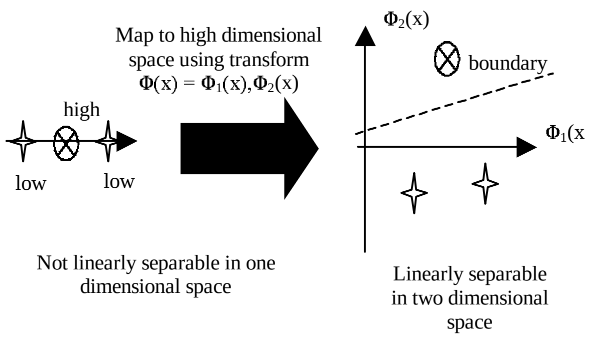

Figure 3 Example of mapping to higher dimension to make linearly separable

This method of mapping to higher dimensions to make the system linearly separable creates two challenges; how to choose a valid mapping transform Φ(X) and that it may be impractical to perform the dot product required for the margin optimisation in higher dimensional space. To overcome these two challenges a Kernel function is used as shown in Equation 2. This Kernel function can implement the dot product between two mapping transforms without needing to know the mapping transform function itself.

Equation 2 Kernel function to perform dot product of two mapping functions

$$K(X_k, X_j) = \Phi(X_k) \bullet \Phi(X_j)$$

Once the Kernel function has been included the SVM training can be written as the quadratic optimisation problem in lagrangian multiplier form as:

Equation 3 lagrangian formulation

$$\max \quad W(\Lambda) = \Lambda^T \bullet \tilde{1} - \tfrac{1}{2}\left[\Lambda^T D \Lambda\right]$$

Where

$$D_{k,j} = d_k y_j K(x_k, x_j)$$

and the vector of lagrangian multipliers is

$$\Lambda = (\lambda_1, \lambda_2, ..., \lambda_l)$$

Solving this quadratic optimisation gives the vector of lagrangian multipliers (shadow prices). Support vectors are the only data points with non-zero lagrangian multipliers so only support vectors are required to produce a forecasting model (i.e. describe the boundary in Figure 1).

$$s \text{ support vectors} \quad S_s = X_s \quad only\ if\ \lambda_s \neq 0$$

To produce forecast implement Equation 4 below as in Figure 4

Equation 4 output of SVM

$$f(X_k) = sign(net(X_k))$$
$$net(X_k) = \sum_s \lambda_s d_s K(S_s, X_k) + b$$

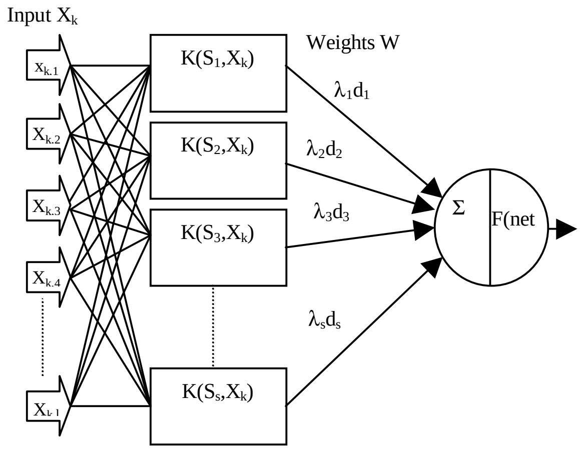

Figure 4 Structure of SVM

To apply SVM to regression forecasts a slack variable ξ_k is applied for each data point, which allows for an error between the target price y_k and the output of the SVM. The optimisation then becomes:

Equation 5 SVM training for regression

$$\text{minimise} \quad F(W) = \frac{1}{2}(W \bullet W^T) + C\sum_k \xi_k$$
$$\text{subject to} \quad y_k(W \bullet X_k + b) - 1 + \xi_k \geq 0$$
$$\text{for data points} \quad k = 1,..., l$$
$$\text{where} \quad y_k \quad \text{is price of data point } k$$

C is a parameter chosen by the user to assign penalties to the errors. A large C assigns more penalty to the errors so the SVM is trained to minimise error, can be considered lower generalisation. A small C assigns less penalty to errors so SVM is trained to minimise margin while allowing errors, higher generalisation. From previous studies a C between 0.1 and 10 was found best for electricity price forecasting models. In this paper all SVM models are trained with C set to 0.5.

4 PROCEDURE

The SVM training and forecasts were performed with the mySVM program developed by Stefan Rüping [10]. The program was designed to solve the dual of the optimisation in Equation 5 by dividing the training set into small working sets or chunks [11].

In this study all forecasts were seven days into the future forecasts utilising real NEM data obtained from the NEMMCO web site [13]. Note no data was omitted not even very large price spikes.

The forecasting tools were designed to produce a practical forecast and so no data was used that would not be available to all market participants at the time of producing the forecast.

Timing terminology used within this paper:

- As standard for the NEM the time t is defined as the trading half-hour. The half-hour is defined as the half-hour ending at that time. So the 48th half-hour of the day is defined as the 0:00 half-hour which covers the trading period from 23:30 to 0:00(midnight). So a day starts at the 1st half-hour 00:30 covering the period from 00:00 to 00:30.
- NOW is at t=0. The time at which the forecast is produced.
- Forecast time. The time in half-hours the forecast is for. A forecast time of 8/3/02 14:30 means the forecast price is for the trading half-hour ending at 14:30 on 8/March/02. (note UK date format used)
- The delay. Is the time t in half-hours before NOW. So a negative delay is in the future compared to NOW.
- Forecast ahead. Is the time in half-hours for which the forecast is made into the future. Thus a one week ahead forecast has a forecast ahead of 336 half-hours.

To allow the user time to obtain and process the short-term PASA data files a minimum delay of one hour was always used in processing the data for producing these forecasts.

In our early price forecasting studies it was assumed that a very accurate forecast of future regional demand was available and so the actual demand for the forecast time was used in producing the price forecasts. In this study no data after NOW is used. The demand forecasts used are from the short-term PASA files provided on the NEMMCO web site. So the forecasts produced in this study are more practical results than in our past studies. In previous studies we found that using a demand forecast instead of the actual demand reduced the accuracy of the price forecasts by 1 to 4% (average of 2.3%) depending of the accuracy of the demand forecast for that week.

All SVM price forecasting models were trained with 28 days of data and tested by forecasting the next seven days of NSW regional electricity price. The results were obtained by testing over 25 weeks of data from the 12th of February to the 30th of July 2002. This data was obtained from the NEMMCO web site.

The SVM forecasting models utilising PASA data were presented with all 15 variables in Table 1. The models not using PASA data were presented with variables 1 to 4 and 11 to 15 only.

Table 1 Inputs Variables

	Inputs to SVM		
Input	Input Name	Half-hour delay. t=0 NOW	Comment
Target	spot price	t=-336	Cents/MW
1	spot price	t=3	1 hour
2	regional demand	t=3	1 hour
3	daily half-hour	t=-336	
4	weekly half-hour	t= -336	
5	capacity required	N/a	PASA File at delay t=2 Data read at time delay t=-336
6	reserve required	N/a	
7	reserve surplus	N/a	
8	PASA demand 10%	N/a	
9	PASA demand 50%	N/a	
10	PASA demand 90%	N/a	
11	spot price	48	1 day
12	spot price	96	2 days
13	spot price	144	3 days
14	spot price	336	1 week
15	spot price	672	2 weeks

5 RESULTS

5.1 Value of PASA data

The SVM price forecasting model utilising no PASA data gave forecasting with a Mean Absolute Error (MAE) of 28.6% and a Root Mean Squared (RMS) error of 251. The addition of PASA data offered no substantial improvement in forecasting accuracy to MAE 28.0% and RMS 254. The plots of MAE and RMS are shown in Figure 8 and Figure 9. The plots for the model not using PASA data were almost identical to these plots. Both MAE and RMS error plots shown in this paper are sliding window averages, with widow widths of 48 and 336 half-hours.

5.2 Analysis of results

Before the winter load and pricing patterns began around the 20th of May (such as in Figure 5) the forecasting results were more acceptable with a MAE of 22.0% and a RMS of 12.1. After the 20th of May the winter pattern began with the price spiking most weekdays at 18:00 and/or 18:30 as shown in Figure 6. These large price increases were predictable as they occurred between 17:00 and 19:30, mostly at 18:00 on weekdays. In the NSW region over the winter period 18:00 to 18:30 is the peak load period of the day and so is expected to have the highest prices of that day. However the magnitudes of these daily price spikes did not have any obvious pattern and so will be a focus of our next study.

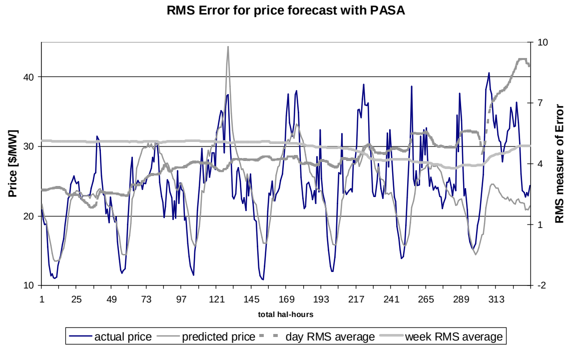

Figure 5 Good accuracy week 3rd to 9th of March 02

One possible solution is to use two separate forecasting models one for the very important peak demand and therefore price periods and another model to forecast the price for the remainder of the time. When the half-hours 18:00 and 18:30 were removed from model the error of the results improved (marginally) to 27.1% MAE and (significantly) to 86 RMS as the RMS measure emphasises larger errors.

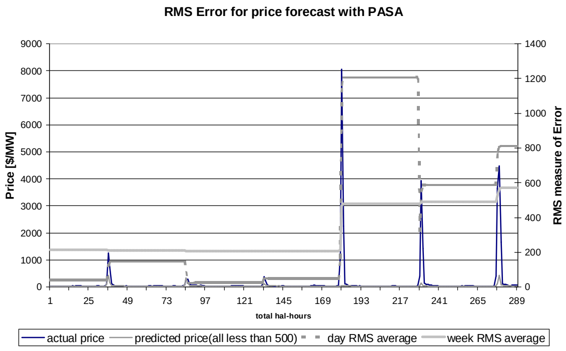

Figure 6 Poor accuracy week 26/6/02 to 1/7/02

5.3 Importance of generation Capacity

Our hypothesis in performing this research was that the price based on a supply and demand market would depend on the difference between generation capacity and required demand and therefore the difference between NEMMCO's demand forecast of required generation and the actual required demand at the time of supply. Figure 7 shows results for the error in demand forecasting and the price for week 19 of the forecasting period. This was typical for the period under study with only a weak correlation found between the errors in demand forecasting and the changes in electricity prices. Thus based on our results our hypothesis would seem to be less important than other factors such as absolute demand magnitudes and generator bidding strategies. Price spikes not caused by system failures appeared to occur at 18:30 on winter weekday nights. The timing of these spikes was independent of whether the required demand was in excess or less than the expected demand obtained from load forecasts. Were these price spikes a result of the demand or of generator bidding behaviour?

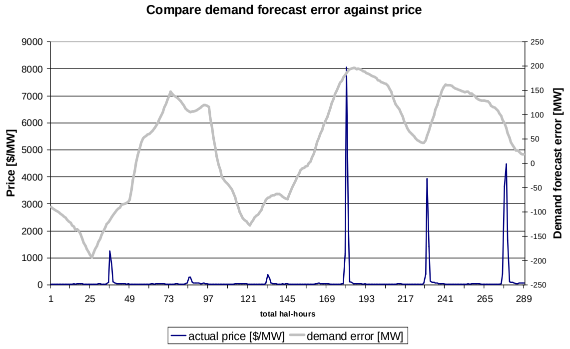

Figure 7 Demand error and price 26/6 to 2/7/02

6 CONCLUSIONS

The PASA data provided only a small improvement in the accuracy of the SVM price forecasting model. Based on these results the cost and time in collecting and processing the PASA data is not justified by the improvement in forecast accuracy.

The accuracy of the demand forecast or knowing the difference between expected generation capacity and the required demand was not as crucial in price forecasting as knowing the time of day and magnitude of the peak demand. However the magnitude of peak demand did not correlate with the magnitude of the price spikes.

Our future research needs to explore and verify our growing belief that understanding generator bidding strategies and regulations and regulatory changes would be more beneficial to electricity price forecasting than historical statistical based methods. The question for the ESI is, "has electricity pricing evolved into a dynamic market where the actions and strategies of participants are of equal or more importance than the deterministic ideas and methods of power system analysis and load forecasting?"

7 REFERENCES

[1] Published by NEMMCO," *An Introduction to Australia's National Electricity Market*: NEMMCO", 1998. Available at [13].

[2] H. S. Hippert, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Transactions on Power Systems*, vol. 16, pp. 44-55, 2001.

[3] Sapeluk A, "Pool price forecasting: a neural network application," *UPEC 94 Conference paper*, vol. 2, pp. 840, 1994.

[4] Ramsay B, "A neural network for predicting system marginal price in the UK power pool," *University of Dundee UK*, 1995.

[5] Sansom-D and Saha-TK, "Neural Networks for Forecasting Electricity Pool Price in a Deregulated Electricity Supply Industry," *AUPEC'99 Darwin Australia Proceedings*, pp. 214, 1999.

[6] Sansom-D, Downs-T and Saha-TK, "Evaluation of support vector machine based forecasting tool in electricity price forecasting for Australian National electricity market," *AUPEC2002 proceedings*, 2002.

[7] C. Cortes, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273-97, 1995.

[8] Platt-JC, "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines," *Microsoft research report*, 1998.

[9] *http://www.kernel-machines.org/*.

[10] S. Rüping, "mySVM software," *http://www.kernel-machines.org/ under software -> mySVM*.

[11] E. Osuna, R. Freund, and F. Girosi, *Support vector machines: training and applications*: Massachusetts Institute of Technology, 1997.

[13] NEMMCO web site accessed Dec 2002. http://www.nemmco.com.au/

[14] Z. Xu and Z. Y. Dong, "Development of a Wavelet and SVM Model for Electricity Price Forecasting", Submitted to IEEE Trans on Power Systems,2002.

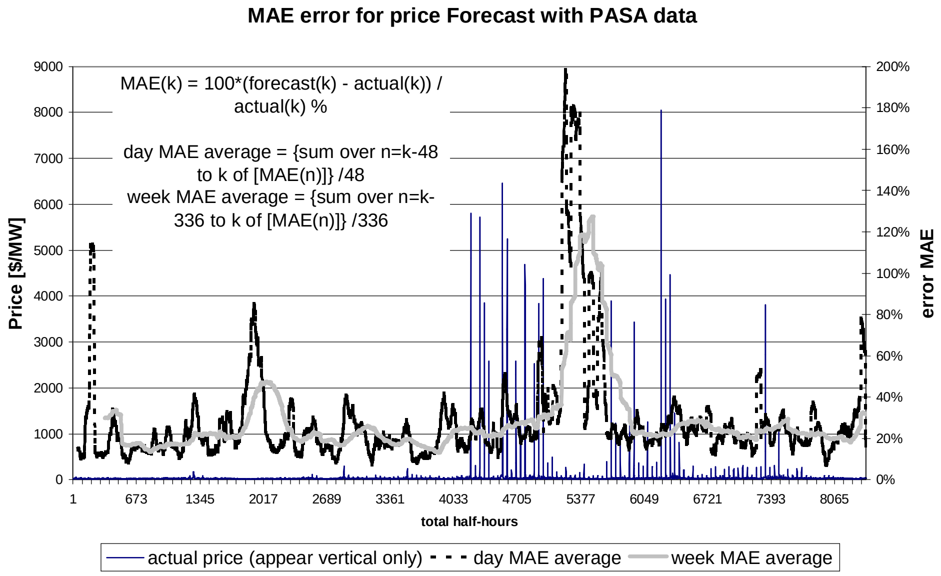

Figure 8 MAE error for Price Forecast with PASA data

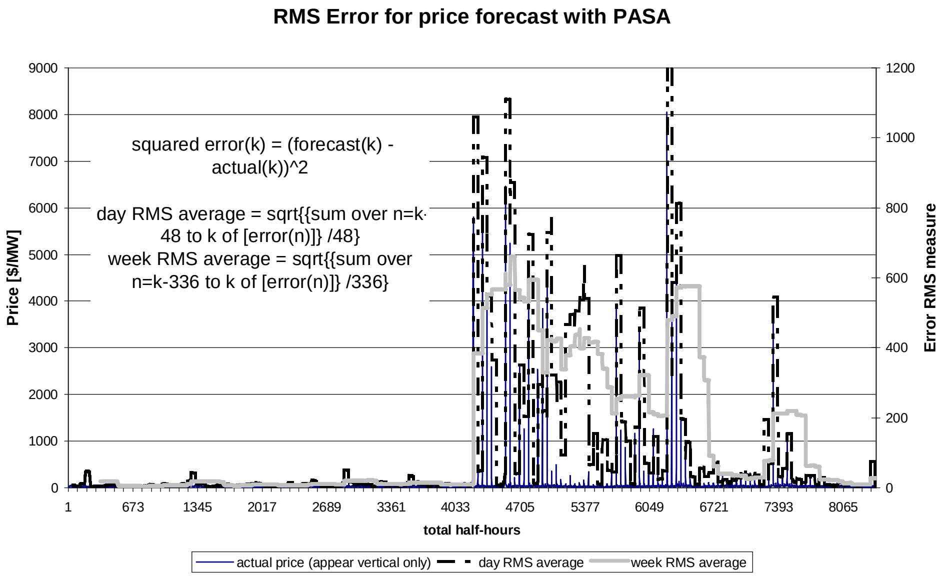

Figure 9 RMS Error for price forecast with PASA data