Short-Term Electricity Price Forecasting

A. Arabali, E. Chalko, *Student Members IEEE*, M. Etezadi-Amoli, *LSM*, *IEEE*, M. S. Fadali, *SM IEEE*

***Abstract*— Price forecasting has become an important tool in the planning and operation of restructured power systems. This paper develops a new short-term electricity price forecasting scheme based on a state space model of the power market. A Gauss-Markov process is used to represent the stochastic dynamics of the electricity market. Kalman and H_∞ filters, two methods based on the state space model, are applied in order to estimate the electricity price and compare the quality of their state estimates. Our results show that performance measures for the H_∞ filter are generally superior to those for the standard Kalman filter.**

***Index Terms*—electricity price, Kalman filter, H_∞ filter, Gauss-Markov.**

I. Introduction

During the last two decades, with the introduction of restructuring, the price of electricity has become the focus of all agents in the power market. Consequently, electricity price forecasting plays an essential role in establishing appropriate economical operation. In the wholesale power market, electricity price forecasting is critical to both market participants and market operators. In spot markets, market participants (GenCos and DisCos) need price forecasts to decide their bidding strategies and maximize their profits. They use price forecasts to allocate assets, organize bilateral contracts, hedge risks, and plan facility investments. Market operators can also use electricity price forecasts to develop market power indices for monitoring the performance of market participants. Accurate prediction of electricity prices is a necessity in power markets for both market participants and operators. Electricity prices are impacted by factors such as transmission congestions, supply-side decision-making, and market power exercises. Therefore, the selection of input variables is critical to achieve high forecasting accuracy [1].

Different techniques for electricity price forecasting have been applied in the literature. These techniques are categorized into simulation methods, time series, intelligent systems, econometric methods, equilibrium analysis, and volatility analysis. Intelligent systems and time series are usually used for day-ahead price forecasting. Typical time series and intelligent system methods include auto-regressive moving average (ARMA), generalized auto-regressive conditional heteroskedastic (GARCH), feed-forward neural networks (FFNNs), recurrent neural networks (RNNs), adaptive network-based fuzzy inference system (ANFIS), and support vector machine (SVM) [2]. The prediction performance of most of these techniques is compared based on computer simulations in [1], [3] and [4]. These simulations show that considering the same inputs, a simple intelligent system, such as FFNN, performs more accurately and efficiently compared to time series, such as ARMA and GARCH, and complex intelligent systems, e.g., RNN and ANFIS [4].

A combination of time series methods and fuzzy inference system (FIS) for day-ahead electricity price forecasting was proposed in [1]. The FIS was adopted due to its clearness and interpretability and least-squares estimation (LSE) was used to estimate the parameters of the time series. The forecasting methodology was applied to the locational marginal price (LMP) based market environment [1]. Zonal loads and day-ahead constraints are included as inputs to represent impacts of demand-side factors and congested transmissions, respectively. The input variables were correlated variables including prices for the same hours on the previous day, prices one week earlier, demand loads of local zone, and area wide and transmission constraints while the output variables were the specified applications. The results of [1] showed that LSE provides the most accurate results, and FIS, which is also highly accurate, provides transparency and interpretability.

In [2] and [5], an approach was proposed for forecasting hourly prices in the day-ahead electricity market using a recursive neural network (RNN). Generally, RNN is a multi-step approach that includes one output node and recursively uses its prediction as input for subsequent forecasts. Thus, it is carried out recursively for twenty four steps to predict next 24 hour prices. A comparison of the prediction performance of the proposed RNN model to other methods used in literature was performed in [5]. Through simulation, [5] showed that the proposed RNN model can provide more efficient and accurate results for the Pennsylvania-New Jersey-Maryland (PJM) electricity market.

None of the methods in the literature define an appropriate scheme to use the state space of the power market for price forecasting. This paper proposes an algorithm to find the state space model of the power market. Using this algorithm and applying two estimation methods, the Kalman and H_∞ filters, a new scheme is developed to forecast the electricity price (LMP). Section II. *A* develops a state space model for the power market. Kalman and H_∞ filter equations are defined in Section II *B* and *C*, respectively. A new price forecasting scheme is proposed in Section II. *D*. Sections III and IV are devoted to the simulation results and conclusion, respectively.

II. Methodology

A. Modeling

A Gauss-Markov process can be used to fit many physical processes with reasonable accuracy. The process can be represented by shaping filter with unity Gaussian white noise input [6]. Figure 1 shows the general block diagram of a

The authors are with the EE Department at University of Nevada, Reno. (e-mail: aarabali@unr.edu, chalkoe@gmail.com; etezadi@unr.edu; fadali@unr.edu)

978-1-4799-1303-9/13/$31.00 ©2013 IEEE

Gauss Markov process. The transfer function of the shaping filter can be obtained from the autocorrelation or spectral density function by spectral decomposition.

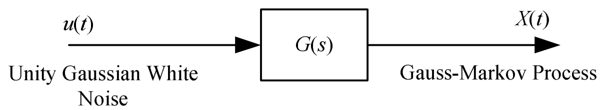

Fig.1. Gauss-Markov Process.

The auto-correlation function $R_X(\tau)$ of a random signal $X(t)$, which describes how the signal is correlated with itself at two different times (t_1 and t_2) is defined as:

$$R_X(t_1, t_2) = E\big(X(t_1)X(t_2)\big) \quad (1)$$

The autocorrelation for a Gauss-Markov process is [6]

$$R_X(\tau) = E\big(X(t)X(t+\tau)\big) = \sigma^2 e^{-\beta|\tau|} \quad (2)$$

where, σ is the standard deviation, τ is the time separation, and β is the reciprocal of the time constant.

The state space equations of a power market are not obtainable by physical modeling and must be derived based on the market data. The idea is we can find the state space equations of a system from the power spectral density function of the system given that the process is known. The power spectral density function (PSDF) shown in (3) is the Fourier transform of the autocorrelation function [6]:

$$S_X(j\omega) = \int_{-\infty}^{+\infty} R_X(\tau) e^{-j\omega\tau} d\tau \quad (3.a)$$

where ω and S are the frequency and power spectral density of $R_X(\tau)$, respectively. In practice, the discrete expression for the spectral density function is used:

$$S_X(j\omega) = \frac{1}{2\pi N} \sum_{n=1}^{N} R_X(\tau) e^{-j\omega n} \quad (3.b)$$

where N is the number of samples. The PSDF can be estimated by taking the Fourier transformer of the autocorrelation function after the latter is experimentally determined from data. It is usually better to directly estimate the PSDF using the *periodogram* of the sample signal [6]. The periodogram is given by:

$$\hat{S}_X(j\omega) = \frac{1}{N} \left| \sum_{k=1}^{N} X(k)\, e^{-j\omega k} \right|^2 \quad (3.c)$$

The periodogram can provide an estimate of the spectral density of a signal in units of power per radian per sample using (3.b). In Section II. *E* we show graphically that the power spectral density function of the power market can be approximated by a Gauss-Markov process model. Therefore, we assume that our system is Gauss-Markov. By taking the Laplace transformer of the autocorrelation function of the system (1), the spectral density function of a Gauss Markov process is:

$$S_{XX}(s) = \frac{a^2}{-s^2+b^2} = \left(\frac{a}{s+b}\right)\left(\frac{a}{-s+b}\right) \quad (4)$$

where, a and b are constants. The shaping filter is the causal part of the spectral density function with transfer function:

$$S_X(s) = \frac{a}{s+b} \quad (5)$$

B. The Kalman filter equations

The Kalman filter is a minimum variance estimator meaning that this filter provides the smallest possible standard deviation of the estimation error. The Kalman filter uses the state space equations of the system to provide an estimate of the state of the system.

Given the shaping filter (transfer function), it is easy to find the state space equations. For a system with the shaping filter of (4), we write the following state space equations:

$$\dot{x}(t) = -bx(t) + U(t) \quad (6)$$

$$y(t) = ax(t) + V(t) \quad (7)$$

where, U and V are the Gaussian white process noise and measurement noise, respectively, x is the state variable, and y is the measurement variable. After determining the state space equations of the system, the equations must be discretized so that we can use a discrete Kalman filter for state estimation [6], [7]. A simple computation of the discrete space equation and the corresponding noise covariance matrix is provided in [6]. The dynamic behavior of a stochastic system is described by the discretized state space model:

$$x_{k+1} = \varphi_k x_k + w_k \quad (8)$$

$$y_k = H_k x_k + v_k \quad (9)$$

where, x_k and y_k are the state and measurement of the system at time k and w_k and v_k are system and measurement noise vector. φ_k and H_k are discretized coefficients. The covariance matrices of w_k and v_k are as follows:

$$E(w_k w_i^T) = \begin{cases} Q_k;\, k = i \\ 0\,;\, k \neq i \end{cases} \quad (10)$$

$$E(v_k v_i^T) = \begin{cases} R_k;\; k = i \\ 0;\;\; k \neq i \end{cases} \quad (11)$$

We assume that the process noise and measurement noise are uncorrelated.

$$E(w_k v_i^T) = 0 \quad (12)$$

The estimation error is defined as:

$$e^- = x_k - \hat{x}_k^- \quad (13)$$

where, $\hat{x}_k^-$ is the minimum mean square error a priori state estimate, i.e. the estimate based on measurements up to time t_{k-1}before considering the effect of measurement at t_k.

The associated error covariance (P) is:

$$P = E(e^- e^{-T}) = E[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T] \quad (14)$$

We correct our a priori estimate based on the measurement y_k using the following equation:

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - H_k \hat{x}_k^-) \quad (15)$$

where, K_k is the Kalman gain. The Kalman gain is given by:

$$K_k = (P_k^- H_k^T)(H_k P_k^- H_k^T - R_k)^{-1} \quad (16)$$

The error covariance matrix for the corrected estimate is:

$$P_k = E(e_k e_k^T) = (I - K_k H_k) P_k^- \quad (17)$$

To complete the filter formulation we need a one-step predictor that yields the next a priori estimate:

$$\hat{x}_{k+1}^- = \varphi_k \hat{x}_k \quad (18)$$

where, $\hat{x}_{k+1}^-$ is the optimal estimate prior to incorporating the next measurement. The error covariance matrix associated with the a priori estimate is governed by the Riccati equation:

$$P_{k+1}^- = \phi_k [P_k^- - P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} H_k P_k^-] \phi_k^T + Q_k = \varphi_k P_k \varphi_k^T + Q_k \quad (19)$$

Figure 2 shows the flowchart of the discrete-time Kalman filter flowchart [6]. To initialize the filter, initial estimates $\hat{x}_0^-$ and P_0^- are used to calculate the Kalman gain (K_0). Using the new measurements (y_k) and Kalman gain, the state estimate $\hat{x}_k$ and its error covariance P_k are corrected based on (15) and (19). The corrected estimates are then used to calculate the projected estimate $\hat{x}_{k+1}^-$ and its error covariance P_{k+1}^-. The process is repeated for all subsequent measurements.

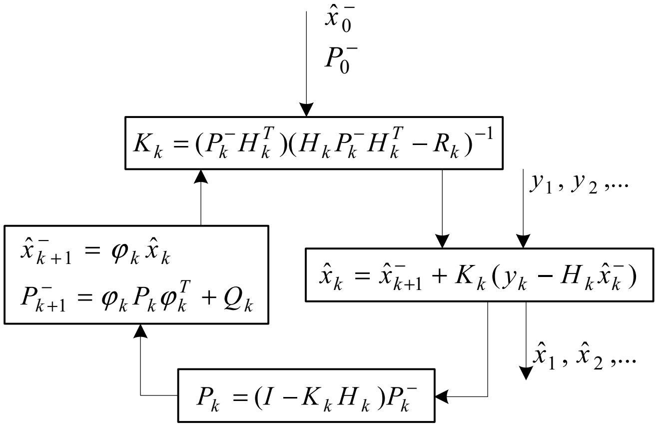

Fig.2. Flowchart of the Kalman filter.

C. The game theoretic H_∞ filter equations

There are some problems associated with the Kalman filter's assumptions in practical applications [8]. First, the expected value of the system noise, w_k, and measurement noise, v_k, must be zero. This problem can be addressed by subtracting and adding the mean value and introducing deterministic inputs but errors in the mean values can increase the estimation error. Moreover, the zero mean property must hold over the entire time history of the process as well as at each time step. Second, the standard deviation of the noise processes must be determined. The covariance matrices of the noise process matrices are used by the Kalman filter as design parameters [9]. For practical problems, determining the covariance matrices of the process noise or measurement noise from experimental data is not an easy task and the estimates may include significant errors [10]. The H_∞ (or min-max) filter makes no assumptions about the noise and minimizes the maximum estimation error instead of the covariance error as in [9] and [11]. The H_∞ filter minimizes the performance measure defined below:

$$\min_{\hat{x}} \max_{w,v} \{J\} \tag{20}$$

where, J is a performance measure for the estimator. This filter looks at the problem as a game theoretic problem in which the noise terms w and v are viewed as opponents who try to worsen the estimate. Because the characteristics of these noise terms are not known exactly, the objective of the H_∞ filter is to find an optimal state estimate that can minimize the worst possible effect of w and v on the estimation error. In other words, H_∞ tries to minimize the maximum estimation error. The function J is defined as follows [11]:

$$J = \frac{ave\|x_k - \hat{x}_k^-\|_Q}{ave\|w_k\|_W + ave\|v_k\|_V} \tag{21}$$

where, $\|x\|_Q = x^T Q x$ is a weighted norm of vector x, Q, W and V are matrices with compatible dimensions. The averages are taken over all time samples. By this minimization, an optimal solution may be found such that the distance between x and $\hat{x}$ is minimized. The matrices W, Q and V matrices used in the weighted norms in J are chosen by the designer to obtain desired trade-offs. For example, if the noise v is smaller than w, one way to choose these parameters is to make V smaller than W which can de-emphasize the importance of the noise v relative to w. Similarly, we should properly adjust Q if we want to improve the estimation accuracy of specific elements of the state vector x.

Note that larger noise sequences w and v cause larger estimation errors. Therefore, by making the noise very large, the "opponent" can simply make it far from the correct x. In order to prevent this, we include the average of the weighted norms of w and v in the denominator of J.

There are several formulations for the H_∞ filter. One set of H_∞ filter's equations which make J less than a desirable level M^{-1} (M is a large value chosen by the designer) has the predictor-corrector structure of the Kalman filter equations:

Corrector:

$$L_k = (I - QP_k^- - H_k^T V^{-1} H_k P_k^-)^{-1} \tag{22}$$

$$K_k = \varphi_k P_k L_k H_k^T V^{-1} \tag{23}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - H_k \hat{x}_k^-) \tag{24}$$

$$P_k = \varphi_k P_k^- L_k \varphi_k^T + W \tag{25}$$

Predictor:

$$\hat{x}_{k+1}^- = \varphi_k \hat{x}_k \tag{26}$$

$$P_{k+1}^- = E(e_{k+1}^- e_{k+1}^{-T}) = \varphi_k P_k^- \varphi_k^T + Q \tag{27}$$

where, I is the identity matrix and K_k is the H_∞ gain matrix. The same process as Kalman filter can be used to estimate $\hat{x}$.

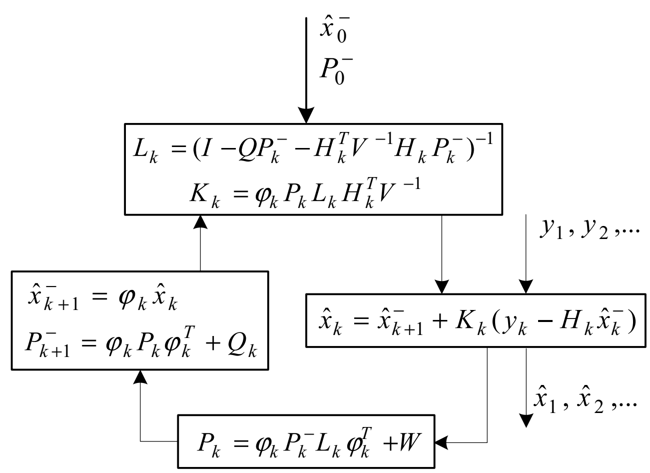

Fig.3. Flowchart of the H_∞ filter.

Figure 3 shows the flowchart of the H_∞ filter [8]. The figure shows that the H_∞ filter equations have the same form as those of the Kalman filter but with minor differences. The initial state estimate $\hat{x}_0^-$ should be the best estimate of x_0, and the initial value P_0^- should be set to give acceptable filter performance [8]. In general, P_0^- should be small if we are highly confident of our initial state estimate $\hat{x}_0^-$. By using H_∞ filter equations, we guarantee that $J < M^{-1}$ meaning that the ratio of the estimation error to the noise will always be less than M^{-1} for any noise terms w and v.

By choosing a very large value for M, we can guarantee that our estimation error is almost zero, since M^{-1} will be almost zero. The mathematical equations of the H_∞ are valid only if M is chosen such that the magnitudes of the eigenvalues of the P matrix are less than one. Therefore, if our choice of M is too large the H_∞ filtering problem may have no solution. In other words, we cannot find an estimator with an arbitrary small estimation error [8].

The main drawback of the H_∞ filter is that it requires more tuning to get acceptable performance. Several parameters such as V, W and Q must be well tuned in order to obtain a good estimate. Consequently, more effort is required to tune an H_∞ filter than a standard Kalman filter.

D. Day-ahead price forecasting techniques using Kalman filter and game theoretic H_∞ filter

The following algorithm provides a forecast of the electricity price using either the Kalman filter or the H_∞ filter:

a) Procedure

1. Load the electricity price data (LMP).
2. Calculate the spectral density function for the LMP data using Matlab command Periodogram.
3. Fit a curve to find the parameters of the shaping filter a and b, assuming a Gauss-Markov process.
4. Obtain a state-space model for the transfer function of the shaping filter.
5. Calculate the required parameters for the filter.
6. Use the filter (Figures 2 or 3) to estimate the price.
7. Calculate the performance measures to evaluate the performance of the filter.

E. Simulation and Results

Our simulations used historical data for the LMP from the California ISO website [12]. Note that for each set of LMP data, the estimated shaping filter coefficients a and b are different. The average of these parameters are used to build the state space model of the power market and the system noise is calculated by comparing the actual data and the data used for building the state space of the power market. For the simulation, the state variable and output are the LMPs of the power system, i.e. a is equal to one. Also, the measurement is simply the LMP. Figure 4 shows PSDF for the LMP data and the fitted curve. The spectral density function was obtained using Matlab's periodogram function and the LMP data. The figure shows that the PSDF of the power market can be approximated by a Gauss-Markov process.

The parameters of the shaping filter for the power market, a and b in (5), can be calculated using curve fitting. The difference between the fitted curve and the PSDF can be considered as the system noise. The figure confirms that the behavior of the power market generally follows that of a Gauss Markov process. Figures 5 and 6 show the predicted versus the actual value of LMP for a specific day and week, respectively. The figures show that the Kalman filter cannot provide a good estimate of the LMP after a sudden change in the market price.

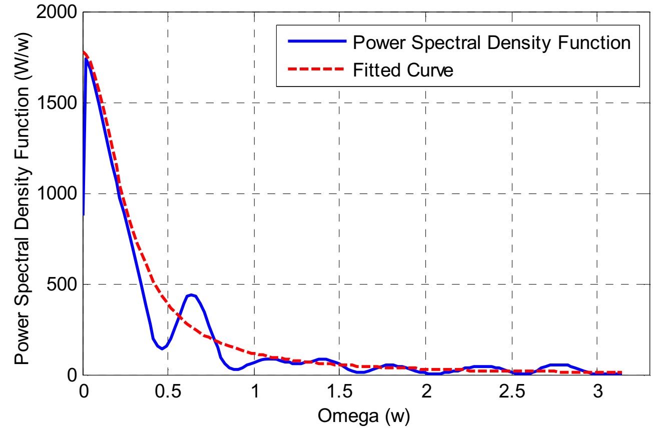

Fig.4. Power Spectral density function and fitted curve.

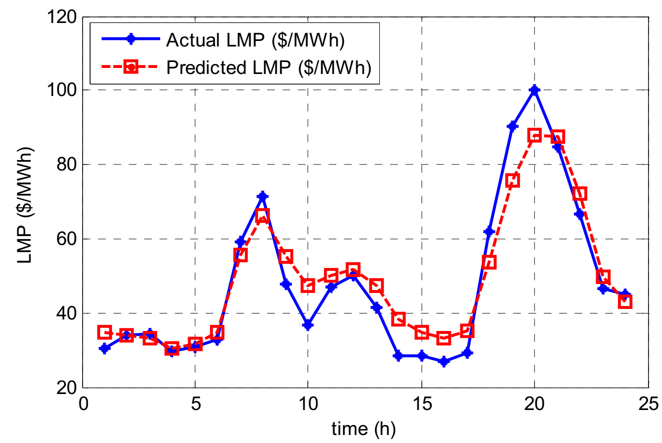

Fig.5. Actual vs. predicted LMP using Kalman filter for a specific day.

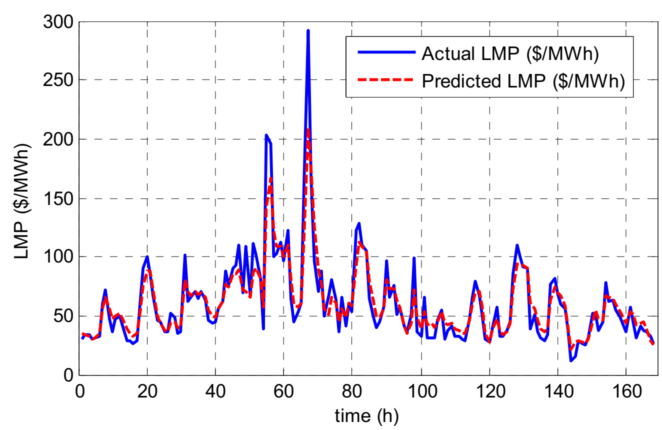

Fig.6. Actual vs. predicted LMP using Kalman filter for a specific week.

Figures 7 and 8 show the estimated LMP using the H_∞ for a specific day and week, respectively. A comparison of figures 5 and 7 shows that the estimates provided by the H_∞ filter are better than that of the Kalman filter, especially when there is a sudden change in electricity price.

F. Performance measurement

In order to evaluate the accuracy of these two methods, two performance measurement indexes are defined. The first index is Mean Absolute Error (*MAE*):

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\widehat{LMP}^i - LMP^i_{Actual}\right|}{LMP^i_{Actual}} \qquad (28)$$

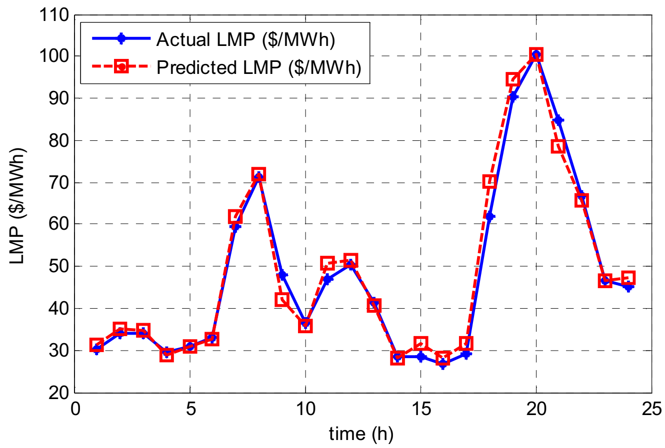

Fig.7. Actual vs. predicted LMP using H_∞ filter for a specific day.

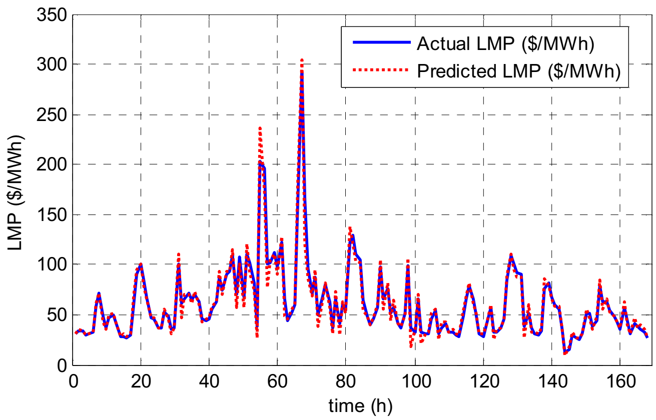

Fig.8. Actual vs. predicted LMP using H_∞ filter for a specific week.

The second index is Root Mean Square Error (*RMSE*):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\widehat{LMP}^i - LMP^i_{Actual})^2} \qquad (29)$$

where, N is number of samples, $\widehat{LMP}^i$ is the estimated LMP, and LMP^i_{Actual} is the actual LMP.

Table I shows the performance measure for both Kalman and H_∞ filter. The performance measure for the H_∞ filter is less than that of the Kalman filter, meaning that H_∞ filter is more accurate than Kalman filter. The Kalman filter does not perform well in the presence of model uncertainty because it uses noise covariance matrices that do not accurately reflect the characteristics of the noise. On the other hand, the H_∞ filter can tolerate modeling errors because it is optimized for the worst-case process and measurement noises.

Table I: Performance measure for the Kalman and $H\infty$ filter.

Method	*MAE (%)*	*RMSE*
Kalman filter (for a day)	0.1021	5.7787
H∞ filter (for a day)	0.0592	3.3505
Kalman filter (for a week)	0.1353	12.9476
H∞ filter (for a week)	0.0754	7.4677

III. Conclusions

This paper considered short term electricity price forecasting. It developed a new scheme for the electricity price forecast based on a state space model of the power market. A Gauss-Markov process is used to represent the stochastic dynamics of the electricity market system. Using the periodogram and curve fitting, the state space model of power market is developed. The Kalman filter and H_∞ filter are then used to estimate the electricity price. Performance measures indices are defined and calculated for both methods in order to evaluate the accuracy. Simulation results shows that, the H_∞ filter can forecast the price more precisely than the Kalman filter in the presence of significant model uncertainty. For future work, we will evaluate the effects of nonlinearity in the system described by PSDF on the estimation error using nonlinear filters such as the extended Kalman filter and the unscented Kalman filter.

IV. References

[1] G. L. Liu, C. Mattson, J. Lawarree, "Day-Ahead Electricity Price Forecasting in a Grid Environment," *IEEE Transactions on Power Delivery Systems*, Vol. 22, No. 1, Page(s): 266-274, Feb. 2007.

[2] N. Amjady, A. Daraeepour, F. Keynia, "Day-ahead electricity price forecasting by modified relief algorithm and hybrid neural network," *IET Gener. Transm. Distrib.*, vol. 4, issue. 3, Page(s): 432 – 444, Feb. 2010.

[3] J. Contreras, R. Espínola, F. J. Nogales, and A. J. Conejo "ARIMA Models to Predict Next-Day Electricity Prices," *IEEE Transactions on Power Systems*, Vol. 18, No. 3, Page(s): 1014-1020, Aug. 2003.

[4] T. Senjyu, P. Mandal, K. Uezato, and T. Funabashi, "Next day load curve forecasting using recurrent neural network structure," IEEE Proc.-Gener. Transm. Distrib., vol. 151, no. 3, pp. 388− 394, May 2004.

[5] P. Mandal, T. Senjyu, N. Urasaki, A. Yona, T. Funabashi, and A. K. Srivastava, "Price Forecasting for Day-Ahead Electricity Market Using Recursive Neural Network," Power Engineering Society General Meeting, page(s): 1 – 8, 2007.

[6] R. G. Brown, P. Y. C. Hwang, "Introduction to Random Signals and Applied Kalman Filter," John Wiley and Sons, Third Edition, 1997.

[7] R. S. Mangoubi, "Robust Estimation and Failure Detection," 1998.

[8] D. Simon, "A Game Theory Approach to Constrained Minimax State Estimation," Cleveland State University, April 6, 2005.

[9] Y. Hung and F. Yang, "Robust H∞ filtering with error variance constraints for discrete time-varying systems with uncertainty," Automatica 39(7), pp. 1185- 1194, July 2003.

[10] H. Choi, J. Lim, and D. Hong, "Game Theory Based H Infinity Filter Approach for Estimating the Number of Competing Terminals in IEEE 802.11 Networks," The 8th Annual IEEE Consumer Communications and Networking Conference, 2011.

[11] X. Shen and L. Deng, "Game Theory Approach to Discrete H_∞ Filter Design," *IEEE Transactions on Signal Processing*, Vol. 45, No. 4, April 1997.

[12] Available at: http://www.caiso.com/market/Pages/MarketMonitoring/Ma rket MonitoringReportsPresentations/Default.aspx.