Forecasting Prices and Congestion for Transmission Grid Operation

Academic Team Members: Principal Investigators: Chen-Ching Liu (Professor of Electrical and Computer Engineering) and Leigh Tesfatsion (Professor of Economics, Mathematics, and Electrical and Computer Engineering). Research Assistants: Qun Zhou (ECpE Ph.D. Candidate, start date 8/18/07) and Nanpeng Yu (ECpE Ph.D. Candidate, start date 1/8/2008)

Abstract: This project is focused on the design of price forecasting tools for market operators and for market traders, taking into account the distinct purposes, data availability, and time horizons of the distinct users. Empirical price data from the MISO and RTE have been analyzed. A combined model is developed to forecast MISO day-ahead nodal prices. An ARIMA model is constructed to forecast RTE week-ahead daily system average price. The unsatisfactory result of the pure statistical model for the RTE price data calls for the adoption of a more structured approach. To facilitate the portfolio management process in power markets, a structured forecasting model incorporating demand scenario generation is proposed.

1. Introduction

Competitive electricity markets have been established in US and other countries. Development of price forecasting tools is critical to Market Operators (MOs) and Market Participants (MPs) in restructured wholesale power markets.

For market operators, forecasting tools can be used for identification of potential congestive conditions, detection of the exercise of market power, and scenario-conditioned planning. Apart from the above benefits, RTE can also apply forecasting tools for portfolio management. From a market participants' point of view, forecasting tools facilitate risk management, trading strategies design, and long-term investment planning.

To evaluate the usefulness of publicly available electricity market information in forecasting prices, two case studies are performed. The first case study forecasts MISO day-ahead nodal prices, using historical price and load data. The second case study uses RTE's historical price data to forecast the week-ahead daily system average price.

With only publicly available information in hand, applicable forecasting tools are restricted to statistical methods. Several researchers have contributed to electricity price forecasting using statistical models. Neural network approaches are studied in [1-3], while time series models are investigated in [4-7]. Having only historical price information, an ARIMA model is applied to fit the price data from RTE. With additional historical load data, a combined Artificial Neural Networks (ANN)/Time Series Models (TSM) model is developed to fit the price data from MISO.

It is highlighted in this project's interim report that, when developing forecasting tools for use in wholesale power market applications, five key questions should be kept in mind. These are: (1) for whom is the tool intended; (2) for what purpose is the tool intended; (3) what is to be forecasted; (4) what data are available to inform this forecast; and (5) for what time horizon is the forecast intended?

The portfolio management task is crucial to MPs and MOs in restructured electricity markets. To cope with the challenges of risk management in power markets, portfolio management tools and sophisticated forecasting tools that take us beyond 'just trust your instinct' should be utilized to our advantage. Hence this report chooses to investigate the development of forecasting tools that are intended for portfolio management. Portfolio management problems are formulated for two entities: (a) a typical generation company (GenCo) in the U.S.; and (b) the French market operator RTE. A structural model for electricity price forecasting that includes demand scenario generation is proposed.

The remainder of this report is organized as follows. Section 2 presents the combined ANN/TSM model for MISO day-ahead electricity price forecasting. In Section 3, empirical data analysis is performed on RTE's price data. An ARIMA model is tested for week-ahead system average price forecasting. A brief description of the developed scenario generation method is given in Section 4. Section 5 formulates the portfolio management problem for two entities and proposes a structural approach for electricity price forecasting. Section 6 summarizes project work to date. Future planned work is outlined in Section 7. Section 8 provides a listing of project-related publications and presentations.

2. Combined ANN/TSM Model for MISO Day-ahead Price Forecasting

2.1 Introduction

From Section 1, ANN and TSM are the two most used models in statistical price forecasting. ANN is considered as black-box forecasting which is suitable for complicated non-linear system, while TSM has an explicit form and models linear systems. In ANN forecasting, there is no requirement that residual terms should be a white noise process. However, if not, the indication is there is still room to refine the ANN model through additional extraction of information from the data. With this understanding, TSM is combined with ANN model resulting in white noise residual terms. The MISO case is used to validate this model. The results demonstrate the strength of the combined ANN/TSM model.

2.2 Problem formulation

In this section, the specific proposed model structure is presented. First, we explain why we decided to combine ANN with a time series model. Then, details of the proposed model are described. Finally, the performance evaluation criterion is given for the forecasting.

2.2.1 Intuition behind the proposed model

ANN is powerful in modeling non-linear systems. However, one of the drawbacks is that too many weights need to be adjusted during training of the ANN models. Although some techniques are used for improving the training performance, difficulties still remain. To be specific, it is common to use Mean Squared Errors (MSE) as the performance criterion for the constructed ANN model. When evaluating properties of the residual after ANN fitting, MSE should not be the only criterion; rather, the serial correlation of the residual terms should be taken into account.

The serial correlation of the residual terms after ANN fitting is rarely considered in the ANN electricity price forecasting literature. For a good model, the residual terms should satisfy the requirement of a white noise process: zero mean, uncorrelated and normal distribution. ANN training algorithm and performance do not guarantee this requirement. For an ANN model, if the MSE of the residual terms is not small and a white noise process is not obtained, it can be considered as under-fitted, meaning that it is necessary to refine the model. If the MSE is small and the white noise requirement is satisfied, it is possible that the model is over-fitted. Due to the amount of available data, and the complexity of ANN model, an over-trained ANN model can fit the data. Difficulties can thus arise when trying to find a proper point between under-fitting and over-fitting.

Time series models provide a better way to solve this problem. One step in developing a time series model is called the diagnosis check, which is a check whether the residual terms form a white noise process. Time series models are expressed in a simple, explicit and linear regression form that makes them easy to interpret and adjust. The commonly applied parsimony criterion, which imposes a penalty in proportion to the number of model parameters to be estimated, is a useful tool for combining with ANN modeling for refining the residual terms to a white noise process.

Generally, ANN is a basic fitting model for non-linear functions, i.e., it is a tool for *coarse-tuning.* On the other hand, the simplicity and explicit form of time series models is good for *fine-tuning*. For the day-ahead electricity price forecasting, the combined model is described as:

$$P_t = ANN(P_{t-24}, P_{t-25}, ...) + N_t \tag{1}$$

$$N_t = TSM\,(N_{t-24}, N_{t-25}, ...) + \varepsilon_t \tag{2}$$

$$\varepsilon_t \sim N(0, \sigma^2) \tag{3}$$

In the above expression, P_t represents the electricity price at time t, the output vector of the ANN model. The input vectors are presented by P^{t-24}, P^{t-25}, …, and L^t , which are electricity prices prior to time t, and load data at time t. In the actual day-ahead market, it is only possible to know price data from one day or earlier, instead of one hour earlier. This is the reason why input price data start from t-24. N_t is the residual term after ANN fitting, which is called the medium residual in this report for convenience. It is not necessary to be a white noise process. In particular, it is not a white noise process in most cases. Consequently, N_t can be further refined by time series modeling (TSM) methods. After TSM refining, the resulting residual ε_t follows a white noise process.

Once the refined residual terms are deemed to approximate a white noise process, the combined model is ready for forecasting. Trained ANN serves as the basic non-linear predictor for the future electricity price, and TSM corrects the ANN predicted values by adjusting the medium residual terms. The framework is shown in Fig. 1.

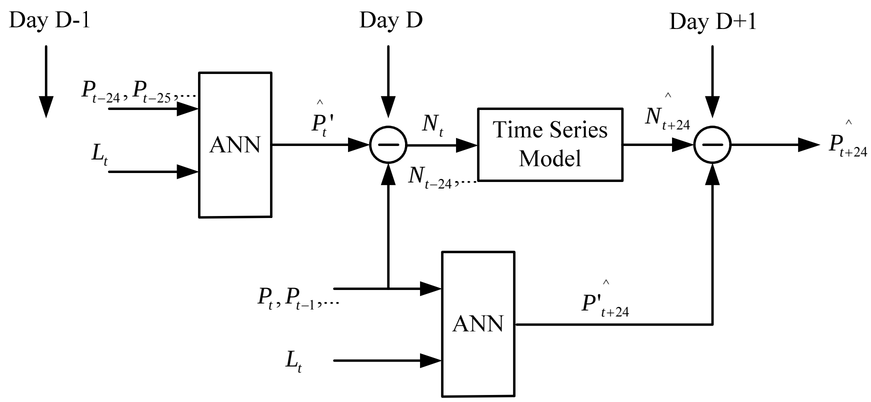


Fig. 1 Proposed framework for electricity price forecasting

In Fig. 1, the ANN model and TSM have been trained with historical price and load data. $\hat{P}_t^{'}$ and $\hat{P}_{t+24}^{'}$ are predicted prices through ANN model, $\hat{P}_{t+24}$ is the corrected price forecast. First the price value $\hat{P}_t$ on Day D is forecasted. On Day D, the actual price is known and the medium residual N_t is obtained sequentially. N_t together with past medium residual are the inputs to TSM for forecasting the medium residual $\hat{N}_{t+24}$ on Day D+1. Correcting $\hat{P}_{t+24}^{'}$ by adjusting the residual, one can obtain the final estimated price value for $\hat{P}_{t+24}$ on Day D+1.

2.2.2 Details of the proposed ANN/TSM model

a. Proposed ANN architecture

ANN applications have been proposed for power systems, e.g., load forecasting, unit commitment, price forecasting, etc. The most common architecture for ANN in price forecasting is multi-layer. This project proposes a multi-layer feed-forward backpropagation neural network, as illustrated in Fig 2. The network is composed of one input layer, one hidden layer and one output layer. As previously mentioned, the proposed method uses load and past price data as input vectors and day-ahead price data as the output vector. More factors can be included in the input vectors, such as fuel prices, weekday index, and weekend index. The output vector can be different according to the purpose of the forecasting, for instance, day-ahead or week-ahead electricity price data are usually taken as the output vector for short-term forecasting.

The role of ANN model in this combined approach is for coarse tuning and, therefore, the training process becomes simpler. However, it is crucial to avoid over-fitting of the ANN model. An over-fitted ANN model can have residual terms which approximate a white noise process even though the model itself is not appropriate for forecasting.

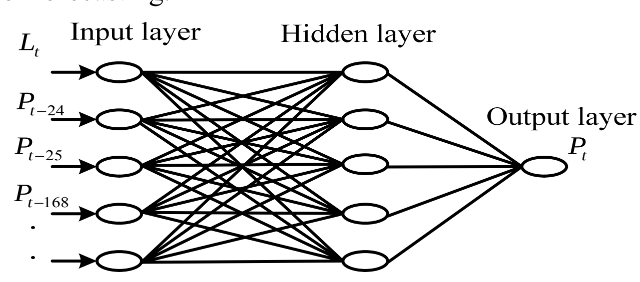


Fig.2 Proposed feed-forward backpropagation network architecture

b. Time series models

A variety of time series models have been used in electricity price forecasting. Since the purpose of time series models is for fine-tuning the medium residual, this project proposes two classical time series models, i.e., Autoregressive Moving Average (ARMA) and Generalized Autoregressive Conditional Heteroskedastic (GARCH). These two models are developed for different types of time series data, homoskedastic and heteroskedastic data, respectively. Time series data are said to exhibit homoskedasticity if the sample covariance matrix for the data is approximately constant over time, and heteroskedasticity if instead the sample covariance matrix displays persistent variation over time.

ARMA(p,q) is expressed as follows:

$$(1-\sum_{i=1}^{p}\phi_i B^i)y_t = c + (1+\sum_{i=1}^{q}\theta_i B^i)\varepsilon_t \quad (4)$$

$$\varepsilon_t \sim N(0,\sigma^2) \quad (5)$$

B is backshift operator, $B^i y_t = y_{t-i}$. c is constant. ϕ_i and θ_i are model parameters to be estimated by historical data. The residual after ARMA fitting is an uncorrelated independent white noise process.

The GARCH(P,Q) model is given as

$$(1-\sum_{i=1}^{P}\phi_i B^i)y_t = c + (1+\sum_{i=1}^{Q}\theta_i B^i)\varepsilon_t \quad (6)$$

P and Q are the order of GARCH model. It is noted that the residual after GARCH fitting is also a white noise process, but uncorrelated dependent white noise process. ε_t takes the form of

$$\varepsilon_t^{\,2} = \nu^2 h_t \quad (7)$$

$$\nu \sim N(0,1) \quad (8)$$

$$h_t = c + \sum_{i=1}^{P}\alpha_i h_{t-i} + \sum_{i=1}^{Q}\beta_i \varepsilon^2{}_{t-i} \quad (9)$$

Some tests can be implemented for judging heteroskedasticity and homoskedasticity, e.g., LM test [8]. However, this report will include both time series models and compare the performance.

c. Price forecasting procedure using the proposed ANN/TSM model

A flowchart is given in Fig. 3 for illustration of the procedure of the proposed forecasting approach. To check the requirement of white noise process for the residual terms (both medium and final), the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) are used to test for serial correlation in the residual terms.

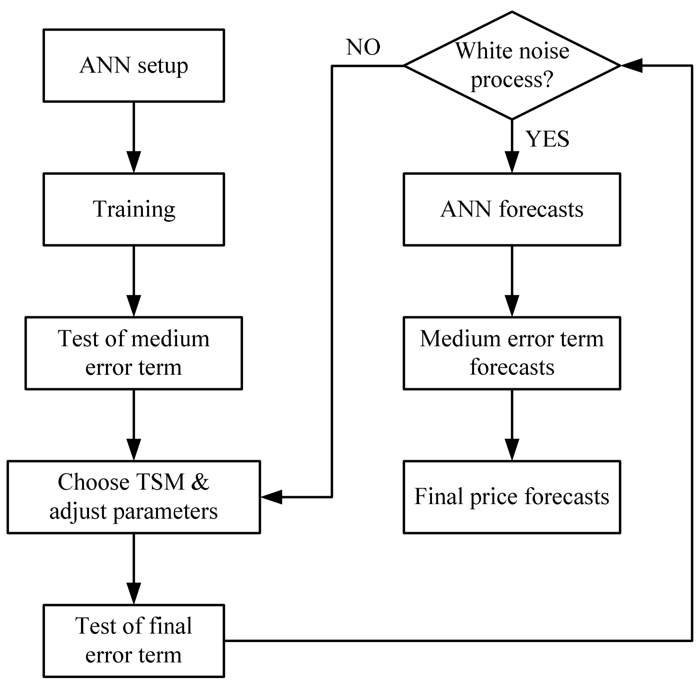


Fig. 3 Proposed forecasting procedure

2.2.3 Performance evaluation

Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) (proposed in [7]) are used to evaluate forecasting performance. The first measurement, RMSE, is the square root of the mean squared differences between true and forecasted prices. More precisely, RMSE is defined as follows:

$$RMSE = \sqrt{[1/N]*\sum_{i=1}^{N}(P_i^{true} - P_i^{frc})^2} \tag{10}$$

The second measurement, MAPE, is defined as follows

$$MAPE = [1/N]*\sum_{i=1}^{N}\left|\frac{P_i^{true} - P_i^{frc}}{\overline{P}}\right| \tag{11}$$

$$\overline{P} = \sum_{i=1}^{N} P_i / N \tag{12}$$

2.3 Case Study

2.3.1 Data and model introduction to the experiment

Data for MISO's day-ahead electricity market during 2008 are used for this case study. Four weeks in four seasons through the entire year have been selected to test the model. For a fair comparison, special attention is focused on the fourth week of March, June, September, December: namely, 03/23-03/29, 06/22-05/28, 09/21-09/27 and 12/21-12/27 respectively. The price data for the year and for the selected four weeks are plotted in Fig.4.

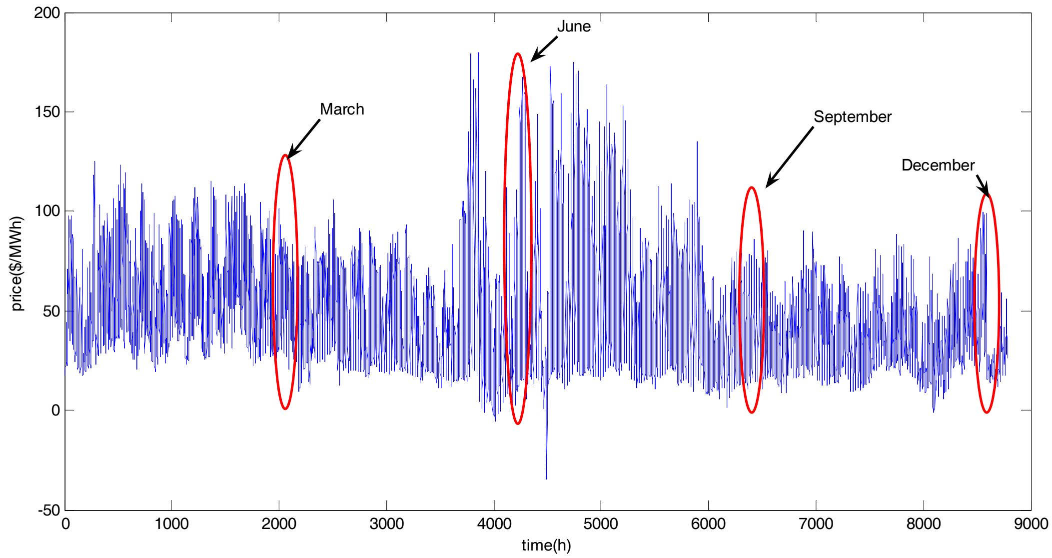


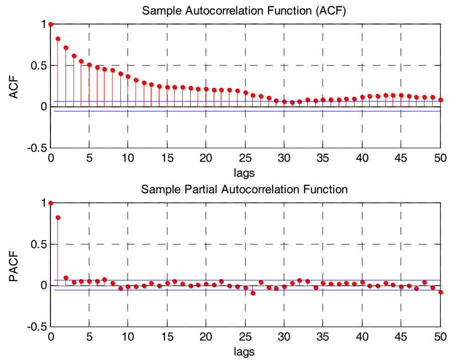
Fig. 4 MISO day-ahead price data in 2008

To construct ANN/TSM, historical price data from one week prior to the forecasted day and load data on that day are used as input vectors.

2.3.2 Efficiency of ANN/TSM by testing the residual

The goal of the combined model is to achieve white noise residual terms through ANN coarse-tuning and TSM fine-tuning. This report only includes the Spring week training data to compare the different impacts on the residual terms using ANN and ANN/TSM.

Fig. 5 and Fig. 6 are ACF and PACF plots for the residual terms before and after TSM remodeling. It can be observed that ANN combined with TSM achieves the goal of refining the complete model. Not only do the fitted residual terms become a white noise process, but also the forecasted results also improve, as shown in Fig. 7.



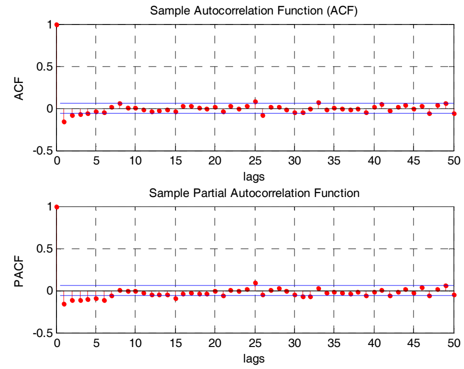


Fig.5. ACF & PACF after ANN fitting

Fig.6. ACF & PACF after ANN/TSM fitting

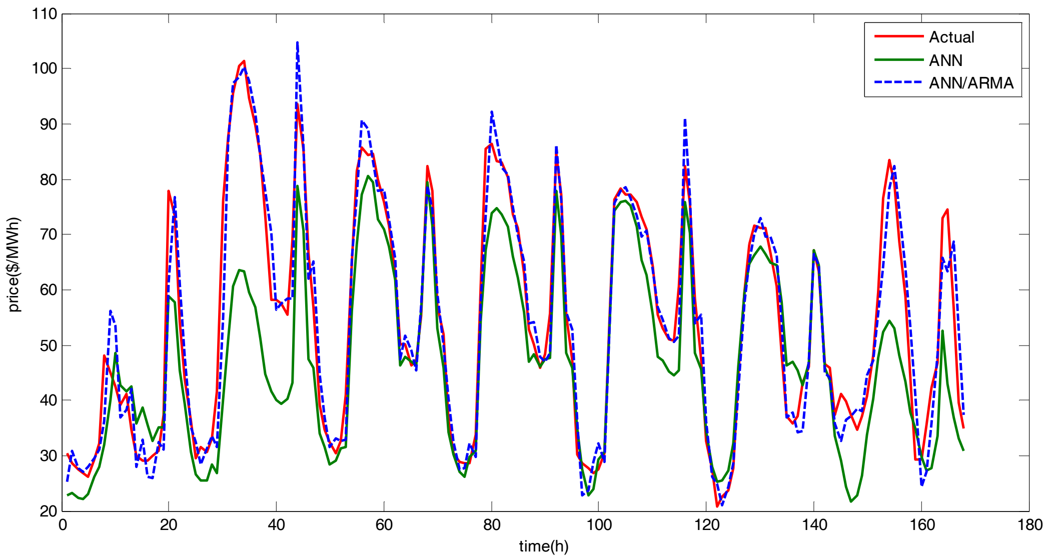


Fig.7 Actual and forecasted MISO day-head prices for the Spring week (03/23-03/29)

2.3.3 Results analysis

The forecasted results for Summer and Fall are shown here. For clarity, only forecasted results by ANN and ANN/ARMA are plotted. Fig. 8 illustrates the forecasted results for the Summer week (06/22-06/28). Compared to the Spring week, the prices for this Summer week are higher and more volatile. This can be observed from the price data for the entire year provided in Fig. 4. The higher volatility can explain the performance results in Tables I and II where it is shown that RMSE and MAPE for the selected Summer week are higher than those of the selected Spring week. The forecasted prices for the selected Fall week (08/10-08/16) are given in Fig. 9. It can be seen that these prices are lower and less volatile than the prices for the Summer week.

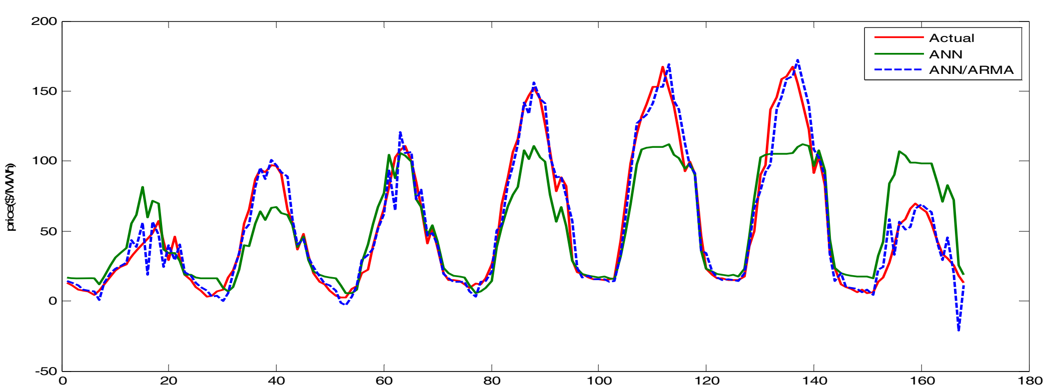


Fig. 8 Actual and forecasted MISO day-ahead prices for the Summer week (09/21-09/27)

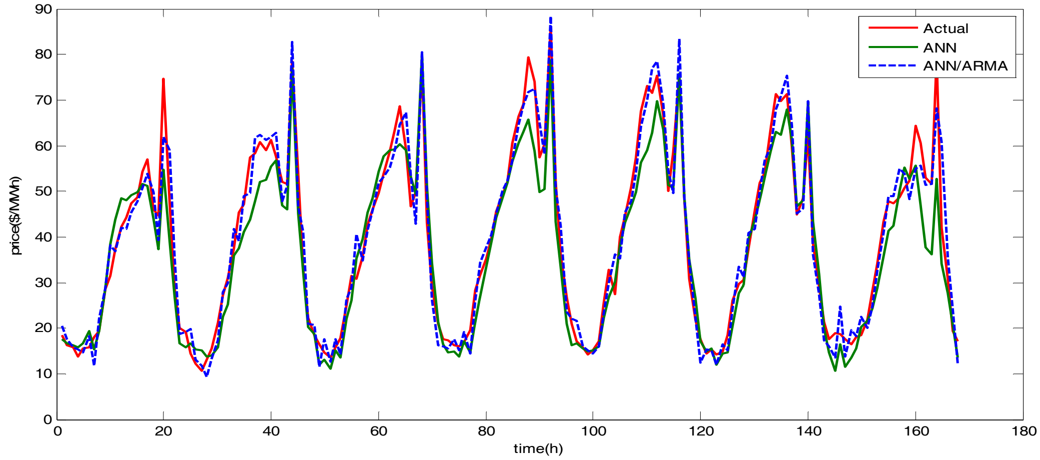


Fig. 9 Actual and forecasted MISO day-ahead prices for the Fall week (09/21-09/27)

Table I and II give the entire results. As seen in these tables, the combined ANN/TSM model generally outperforms either the ANN model or the TSM model alone. It is noted that, in this specific set of MISO data, ANN/ARMA performs better than ANN/GARCH. This indicates that the ANN model succeeds in filtering out the heteroskedasticity of the price data, and the ARMA model is more successful than the GARCH model in terms of ensuring white-noise residual terms.

TABLE I
COMPARISON OF DAY-AHEAD FORECASTING PERFORMANCE USING RMSE MEASUREMENT

RMSE	ARMA	ANN	ANN/ARMA	ANN/GARCH
Spring	13.17	12.24	5.20	5.26
Summer	30.66	22.41	9.91	11.06
Fall	15.12	5.88	4.31	5.41
Winter	14.17	11.96	6.58	6.63

TABLE II
COMPARISON OF DAY-AHEAD FORECASTING PERFORMANCE USING MAPE MEASUREMENT

MAPE	ARMA	ANN	ANN_ARMA	ANN_GARCH
Spring	0.21	0.16	0.07	0.07
Summer	0.51	0.30	0.12	0.16
Fall	0.33	0.11	0.08	0.11
Winter	0.34	0.26	0.13	0.12

The above figures and tables lead to a conclusion that the combined ANN/TSM model effectively extracts the information embedded in the residual terms, resulting in a white-noise process. The ANN/TSM results indicate that the model can forecast day-ahead prices efficiently and accurately.

3. Empirical Data Analysis and Week-ahead Price Forecasting for RTE

3.1 System Price Description

The sample period of France day-ahead electricity price data is from November 26th, 2001 to December 10th, 2008. The day-ahead data consist of 24 time series, one for each hour during a day. The 24 price time series are highly correlated pair-wise. The pair-wise linear correlation coefficients between any two hourly series during the sample period range from 0.267 (between hour 12 and 21) to 0.99 (between hour 14 and 15) with a mean value of 0.67. The pair-wise linear correlation coefficients between any two consecutive hourly series during the sample period range from 0.391 (between hour 20 and 21) to 0.99 (between hour 14 and 15) with a mean value of 0.885. The pair-wise correlation coefficients between any two hourly day-ahead France price series are significantly lower than the ones reported in [9] of the Nordic day-ahead price.

Both European Energy Exchange (EEX) and Powernext use the arithmetic average of all hourly prices (for base product) or of all peak hour prices (for peak product) as the reference prices in the casesettlement calculation. According to this practice, a new time series for this underlying variable is generated by calculating the arithmetic average of the 24 available data for each day. This average price will be referred to as the system price.

The daily system price and daily changes of the system price time series are given in Figure 10. A quick look at the data reveals several prominent features of the France daily system price time series: clustered volatility (time-varying volatility), rare but dramatic price spikes. The daily system price data has maximum and minimum values of 314.27 Euro and 0 Euro, with a mean of 40.48 Euro. The maximum is reached on Nov 15th 2007, whereas the minimum is reached on Feb 27th and March 6th of 2002.

Some descriptive statistics for the daily system price and other related times series are summarized in Table 1.

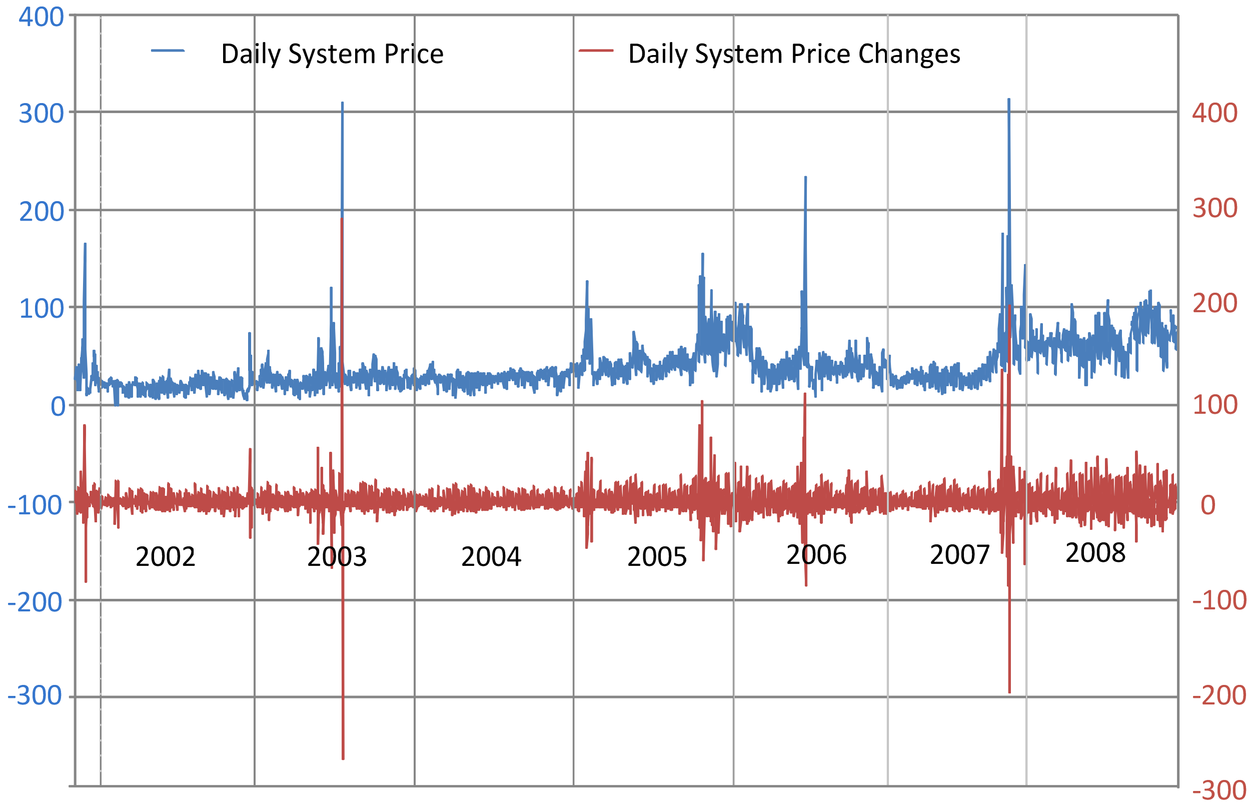


Figure 10. Daily System Price and Daily Changes of System Price (11-26-2001 to 12-10-2008).

Series	Number of Observations	Mean	Median	Maximum	Minimum	Standard Deviation	Skewness	Kurtosis
P_t	2572	40.4775	33.0752	314.2692	0	24.3321	2.5432	18.5673
$P_t - P_{t-1}$	2571	0.0203	-1.0404	291.0325	-264.7862	15.9139	1.4100	101.4020
$\ln(P_t+10)$	2572	3.8306	3.7629	5.7816	2.3026	0.4129	0.4494	3.3808
$\ln(P_t+10) - \ln(P_{t-1}+10)$	2571	0.0003	-0.0236	2.3905	-1.7515	0.2369	0.9303	10.9030

Table III. Descriptive Statistics for Daily System Price and Other Related Times Series

According to standard volatility measures, electricity prices are very volatile. The standard deviation of the daily changes in log-prices is 0.2369, which corresponds to an annualized volatility of 453%. As shown in Table III. Both the system price and log-price have positive skewness indicating heavy right tails. The sample kurtosis of system price is significantly higher than 3 (the kurtosis of a normal distribution). This statistic, combined with the positive skewness, reveals that high prices have a much higher probability of occurrence than that of a normal distribution with the same variance. The sample kurtosis for both changes in system price and changes in log-prices are higher than 3, indicating that large positive and negative jumps are relatively frequent.

The kolmogorov-Smirnov test rejects the null hypothesis that the system price comes from a standard normal distribution at a significance level of 0.01.

Series	Sample Autocorrelation of Lag							
	1	2	3	7	14	21	28	35
P_t	0.786	0.668	0.629	0.710	0.623	0.617	0.597	0.570
$P_t - P_{t-1}$	-0.226	-0.184	-0.014	0.28536	0.277	0.267	0.265	0.234
$\ln(P_t+10)$	0.835	0.722	0.680	0.812	0.739	0.733	0.717	0.696
$\ln(P_t+10)$ - $\ln(P_{t-1}+10)$	-0.158	-0.214	-0.071	0.512	0.496	0.483	0.480	0.465

Table IV. Sample Autocorrelation Function for the System Price.

As shown in Table IV, both the system price and log price are significantly positively correlated at lags that are multiples of seven. This implies a weekly correlation.

3.2 Week-ahead Daily Average Price Forecasting

An ARIMA (autoregressive-integrated moving average) model is a popular class of time series models due to its well established procedure, adaptive ability to represent a wide range of processes, and extendibility to multiple exogenous stochastic variables. The general iterative procedure to model a time series proposed by Box and Jenkins [10] is applied to model the system log-price by an ARIMA model:

Step 1) Identify and formulate a specific model based on an examination of the characteristics and statistics of the original time series.

Step 2) Estimate the parameters of the hypothesized model using available historical data.

Step 3) Perform diagnostic checking using the residuals of the fitted model. If the model is inconsistent with the assumptions, then go back to step 1 to modify the tentative model; otherwise use it for forecast testing.

Empirical knowledge shows clearly the weekly periodicity. It is also noticed that the autocorrelation at lag one is significant. Therefore, two potential differentiations (1-B) and (1-B) (1-B^7) are considered. The ACF of the series (1-B) (1-B^7) P_t dies out more quickly and hence one can proceed with this modified time series.

Three weeks have been selected to perform week-ahead system price forecasting and to validate the performance of the ARIMA models. To forecast each day's system price, the data from the previous 60 days are used to fit the ARIMA models. The general form of the model is

$$(1-\phi_1 B-\cdots\phi_p B^p)(1-B)(1-B^7)P_t=(1-\theta_1 B-\cdots\theta_q B^q)\varepsilon_t$$

$$\varepsilon_t \sim i.i.d.\ N(0,\sigma_\varepsilon^2)$$

As in Section 2.2.3, two indices, Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), are used to evaluate the performance of the price forecasts. These two indices are calculated separately for the two forecasting weeks, as follows:

$$RMSE=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{P}_i-P_i)^2}$$

$$MAPE=\frac{100}{N}\sum_{i=1}^{N}\frac{\left|\hat{P}_i-P_i\right|}{P_i}$$

where N is the number of sample prices, $\hat{P}_i, i=1,...N$, are forecasted prices, and $P_i, i=1,...N$, are the true (actual) prices.

Forecasted Period	RMSE	MAPE
1-25-2002 – 1-31-2002	2.070808	6.399645
3-26-2002 – 4-01-2002	6.553692	27.6827
5-25-2002 – 5-31-2002	2.263012	17.01064

Table V. Performance of ARIMA Model in Week-Ahead Daily Average Price Forecasting.

As shown in Table V, the MAPE of the last two weeks are relatively high which reveals that forecasting result of the ARIMA model without explanatory variables is not satisfactory. Actually, the system average price series does not demonstrate much predictability. The historical price itself does not contain sufficient information for forecasting. This can be illustrated by the unpredictable price spikes in the price series from Figure 10. In some year, no price spike appears, while in other years there could be price spikes in summer or in winter. There must be other critical information such as load data and fuel price data that could help explain the behavior of the price series better. Therefore, it is highly recommended that a more structured model be established for price forecasting.

4. Scenario Generation for Day-ahead Prices

This work has been published [1]. The homepage [12] for our forecasting project includes a link to an on-line pre-print of this paper, where a detailed discussion of our problem formulation and results can be found. In this section, the basic motivation and ideas underlying the paper are briefly discussed.

In recognition of the problem that restructured wholesale power markets are relatively young with a limited amount of price data available for testing and developing forecasting tools, this research proposes a scenario generation method for amplifying empirical wholesale power price data with simulated price scenarios. Scenarios capture typical patterns from the real-world data, and hence can capture realistic properties of this data. Moreover, scenarios also permit forecasters to amplify empirical data patterns that have been realized with data patterns that could be realized. Besides, precise predictions of every possible future scenario can provide a powerful tool for a company to take actions according to different future situations.

Ref.[1] proposes a two-stage approach for generating simulated price scenarios based on the available price data. The first stage consists of an Autoregressive Moving Average (ARMA) model for determining scenarios of cleared demands and scheduled generator outages (D&O), and a moment-matching method for reducing the number of D&O scenarios to a practical scale. In the second stage, polynomials are fitted between D&O and wholesale power prices in order to obtain price scenarios for a specified time frame.

Time series data from MISO are used as a test system to validate the proposed approach. The simulation results indicate that the proposed approach is able to generate price scenarios for distinct seasons with empirically realistic characteristics. The following is a simple illustrative example for scenario generation. Fig. 11 shows the electricity price scenarios in 2006 Spring, indicating regular pattern in Spring. It is observed in Figure 12 that some extremely high prices are included in scenarios by using the proposed approach. Electricity price scenarios in Summer is much more volatile than that in Spring. These two figures demonstrated that proposed approach can capture different price characteristics in Summer and Spring.

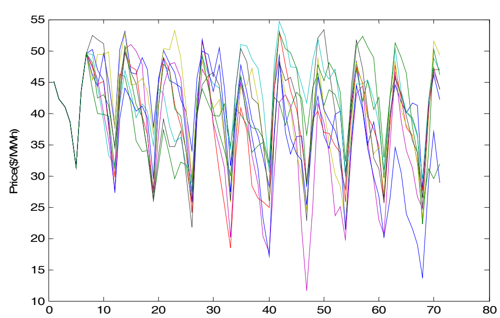


Fig 11 Wholesale power price scenarios in Spring

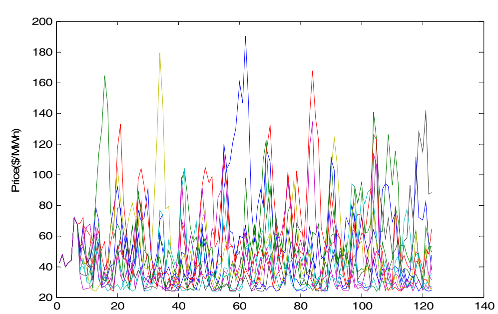


Fig. 12. Wholesale power price scenarios in Summe

This scenario generation method is incorporated in the proposed structuralized model in Section 5.

5. Developing Electricity Price Forecasting Tools for Portfolio Management

5.1 Overview

This section is concerned with the forecasting tools that are intended for portfolio management. Two potential users of this forecasting tool are a Generation Company (GenCo) participating in a U.S. wholesale electricity market and RTE. The portfolio management problems are formulated for each type of the users in the following.

The basic idea of portfolio management is to diversify the portfolio so that the risk is minimized given an expected profit or net earning. For example, a GenCo company would like to diversify its short-term trading portfolio through the day-ahead market, forward bilateral contracts market and FTR market to minimize its risk given an expected net earning requirement. There are five steps in the portfolio management process: establish clearly defined objectives; data gathering; evaluate all resource options; address the risks; and determine the optimal portfolio mix. The key to finding the optimal portfolio mix is to find the efficiency frontier. As can be seen in Figure 13, the points on the efficiency frontier represent the portfolio mixes that offer maximum expected profit at each given risk level. The risk manager will then pick a portfolio mix on the efficiency frontier that maximizes the company's utility expressed as a function of expected profit and risk. In practice, the portfolio manager in a GenCo should perform short-term trading according to this optimal portfolio. Of course, portfolio management should be an on-going process. An action plan should be prepared to assist in coping with and responding to the unexpected.

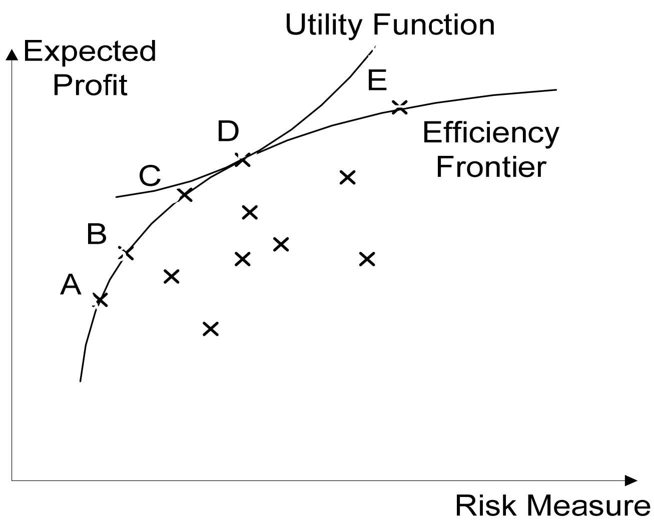


Figure 13. Example of an Efficiency Frontier

5.2 Portfolio Management for a GenCo

The goal of portfolio management for a GenCo is three-fold in the short to medium term: (a) To receive a stable cash flow from selling bulk power through multiple markets; (b) to mitigate risk originating from uncertain load, unexpected outages and fuel, electricity and emission allowance price volatilities; and (c) to generate a reasonably high return under a preferred level of risk. In the analysis below, it is supposed that the market instruments available to the GenCo are electricity forward (day-ahead) market, physical bilateral contract market and FTR secondary market.

The physical bilateral contract market operated by a marketer works as follows: market participant GenCo *i* can view all available physical bilateral contracts he/she can sign to sell power. The following information is posted with respect to the *j*th available physical bilateral contract.

- Bus – This is the buyers' location on the power grid (B_j)
- Maximum MW load in hour *h* - M_j^h
- Strike price in hour *h* – p_j^h (\$/MWh)

If upon viewing all the available physical bilateral contracts, GenCo *i* decides to accept a portion $\lambda_j \in [0,1]$ of the *j*th physical bilateral contract, then the offer will be automatically confirmed by the buyer who posted it. The incurred responsibilities and liabilities are as follows.

At hour h, if $LMP_{B_j}^h \geq p_j^h$, then GenCo i will pay the buyer, $(LMP_{B_j}^h - p_j^h) \times \lambda_j \times M_j^h$

However, if $LMP_{B_j}^h < p_j^h$, then the buyer will pay GenCo i, $(p_j^h - LMP_{B_j}^h) \times \lambda_j \times M_j^h$

where $LMP_{B_j}^h$ is the locational marginal price at bus B_j in hour h.

In addition, GenCo i needs to self-schedule the signed physical bilateral contract through an independent system operator (ISO). When participating in the day-ahead market, it should set the minimum generator output as $\lambda_j \times M_j^h$ at hour h.

The FTR secondary market operated by the ISO works as follows: market participant GenCo i can view all available FTRs to buy in the FTR secondary market. The following information is posted with respect to the kth available FTR.

- Source – This is the FTR's source (B_{sk})
- Sink – This is the FTR's sink (B_{ek})
- Maximum MW FTR at hour h - F_k^h
- Total price for the FTR with Maximum MW amount in each hour - P_k^F

If upon viewing all the available FTRs, GenCo i decides to accept a portion $\delta_k \in [0,1]$ of the kth posted FTR contract, then the offer will be automatically confirmed by the market participant who posted it. The corresponding amount of FTR will be transferred to GenCo i, and it will pay $\delta_k \times P_k^F$ for it.

It is assumed that the GenCo will not bid strategically into the day-ahead market. The supply offer that GenCo i will submit to the market operator is its marginal cost function which is given by $MC_i(q_i^h) = a + 2bq_i^h$, where q_i^h is the MW power dispatch in hour h. The total variable cost of GenCo i is given by $TVC_i(q_i^h) = aq_i^h + b(q_i^h)^2$. Suppose the total number of available physical bilateral contracts is J and the total number of available FTRs is K. Then the decision vector which represents a portfolio, is defined as **x**=(λ_1 , λ_2 , … , λ_J , δ_1 , δ_2 , … , δ_K). It will be chosen from a certain subset X of $\Re^n$. The planning horizon of the portfolio lasts from hour 1 to H. The random vector **y** in $\Re^m$ which represents uncertainties, is defined as the set of locational marginal prices from hour 1 to H at all the buses that are involved in the profit formulation.

Let $\pi_i(x, y)$ be the net earning of GenCo i's portfolio in the planning horizon associated with the decision vector **x** and random vector **y**. The profit formulation shown below contains three terms: the profit from the day-ahead market, the profit from the physical bilateral contracts, and the profit from the FTR contracts.

$$\pi_i(x,y) = \sum_{h=1}^{H}[LMP_{B_i}^h \times q_i^h - TVC_i(q_i^h)] \quad + \quad \sum_{h=1}^{H}\{\sum_{j=1}^{J}[(p_j^h - LMP_{B_j}^h) \times \lambda_j \times M_j^h]\}$$
$$+ \quad \sum_{h=1}^{H}\{\sum_{k=1}^{K}[(LMP_{B_{sk}}^h - LMP_{B_{ek}}^h) \times \delta_k \times F_k^h]\} - \sum_{k=1}^{K}\delta_k \times P_k^F$$

where $$q_i^h = q_i^h(LMP_{B_i}^h, \lambda_1, \ldots \lambda_J) = \begin{cases} q^{\max} & if \quad LMP_{B_i}^h > a + 2bq^{\max} \\ \dfrac{LMP_{B_i}^h - a}{2b} & if \quad 2b\sum_{j=1}^{J}(\lambda_i \times M_j^h) + a \leq LMP_{B_i}^h \leq a + 2bq^{\max} \\ q^{\min} & if \quad LMP_{B_i}^h < a + 2bq^{\min} \end{cases}$$

The generator's dispatch minimum is defined as $q^{\min} = \sum_{j=1}^{J}(\lambda_i \times M_j^h)$. The GenCo i's unit's dispatch is a function of the committed bilateral contracts level and the locational marginal price at its bus. The constraint on the decision vector **x** is that the total scheduled physical bilateral contract cannot exceed the upper capacity limit of the generation unit in any hour h.

$$\sum_{j=1}^{J}(\lambda_i \times M_j^h) \leq q^{\max}, \ \forall h$$

The objective of a GenCo is to determine the optimal portfolio mix that maximizes its expected profit for its preferred level of risk. Conditional Value-at-Risk (CVaR) which is considered to be a more consistent measure of risk than VaR (Value-at-Risk) [13] is used here as the measurement of portfolio risk. The portfolio optimization problem can be viewed as one of finding the efficient frontier of the portfolio. The points on the portfolio's efficient frontier could be found by the following two ways:

(1) $\min_{x \in X} \phi_\beta(\mathbf{x}, \mathbf{y}) \quad s.t. \quad E(\pi(\mathbf{x}, \mathbf{y})) \geq \rho$

(2) $\max_{x \in X} E(\pi(\mathbf{x}, \mathbf{y})) \quad s.t. \quad \phi_\beta(\mathbf{x}, \mathbf{y}) \leq \omega$

The very first step of tackling the optimization problem is to generate samples of the random vector **y**, which are the locational marginal prices (LMPs) at different buses in different hours. Note that the LMPs at different buses in different hours are correlated with each other. Two potential methods that could address this issue are multivariate time series model and structural model. In practice, the dimension of the random vector **y** is usually very high. It is not very practical to build a multivariate time series model of a high dimension, because the number of parameters will increase with the square of the dimension of random vector **y**. Therefore, a structural forecasting model in a virtual world is proposed for generating the samples of the random vector **y**, and testing the portfolio optimization method.

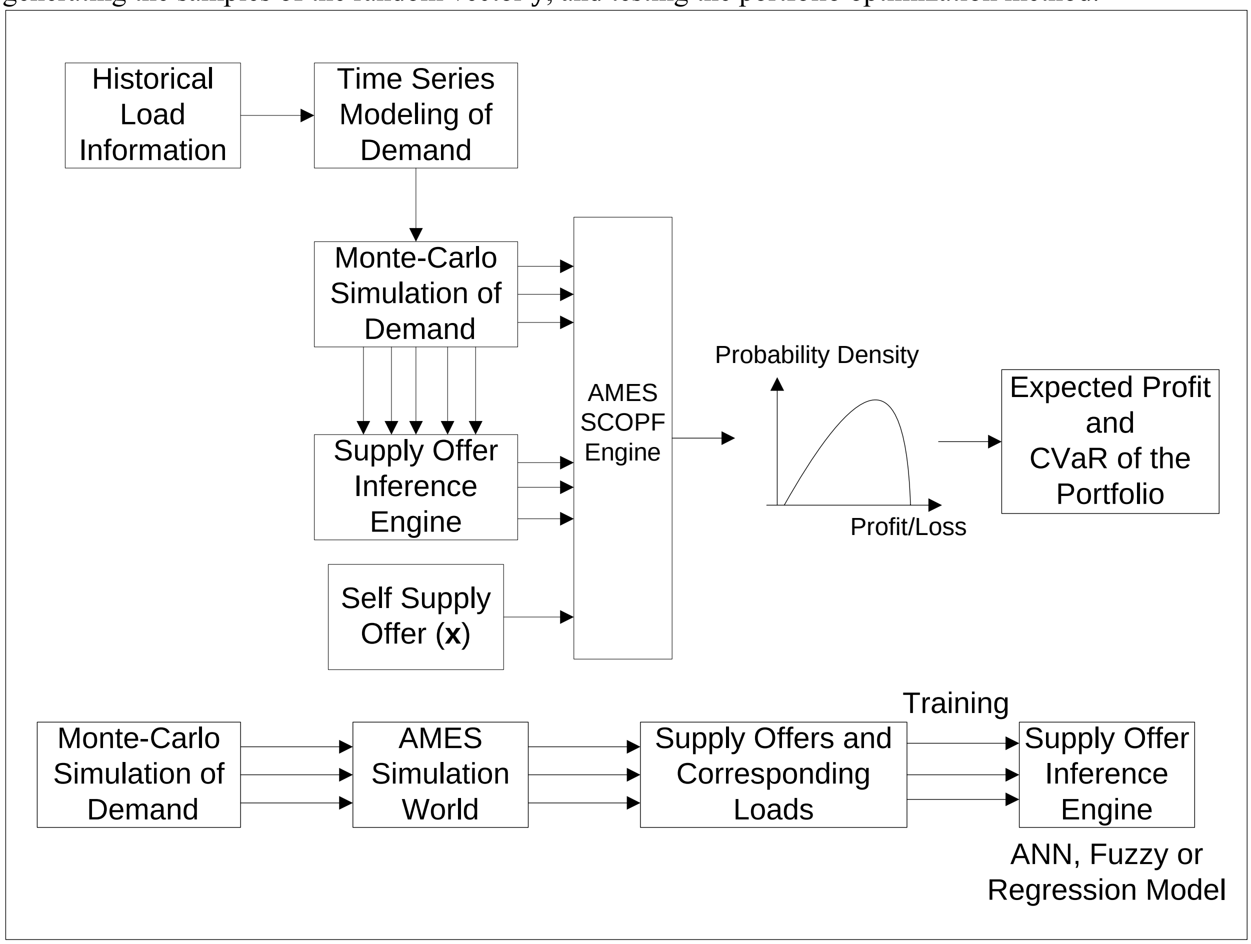


Figure 14. Structural Model for Determination of Electricity Price Samples and Testing of Portfolio Optimization.

As shown in Figure 14, the structural approach mimics the detailed operation of the electricity market. The underlying LMP data generating mechanism is modeled by the security constrained optimal power flow (SCOPF). The input to the SCOPF program includes transmission network data, load data, and generator's supply offer data. The historical load information is used to build a time series model. Through Monte-Carlo simulation, the samples of load are generated. The corresponding supply offers is generated by feeding the samples of load into the supply offer engine. The outputs of the SCOPF engine are the samples of LMPs that will be used to calculate the expected profit and CVaR of a particular portfolio. The optimal power flow developed in the open source agent-based test bed – AMES [14] (V2.0) will be used. The supply offer inference engine is trained based on the simulation results from the AMES virtual world. Load data that generated from the Monte-Carlo simulation is used as the inputs to the simulation. The GenCo agents empowered by learning methods would participate in the electricity market competition. Observations of the load information and corresponding supply offers by the GenCos will be used to train the supply offer inference engine.

Finally, the problem of determining the optimal portfolio mix is under the scope of global optimization problem. Several global optimization algorithms such as deterministic/stochastic branch and bound, simulated annealing and evolutionary algorithm are the candidate algorithms to solve the optimization problem.

5.3 Portfolio Management for RTE

RTE, as the system operator of the French electricity market, needs to purchase sufficient energy to cover the transmission system losses. The market instrument that is available to RTE includes the Future contracts from Powernext and European Energy Exchange. A benchmark portfolio is set up to evaluate the portfolio management strategy adopted by RTE. The benchmark is purchasing future contracts of different terms in small amount every day. RTE will be reimbursed according to the benchmark. RTE's portfolio management objective is to select the optimal mix of portfolio that minimizes the associated cost of the portfolio and minimizes the risks.

As discussed earlier, in order to perform portfolio optimization, meaningful electricity price forecast samples need to be generated. RTE's advantage over other market participants is that is has much more information about the underlying data-generating mechanism that determines electricity prices. Therefore, a structural model should presumably be preferred to a pure time series model from RTE's point of view. Currently, however, the only RTE information available to this project is historical electricity price information. Unfortunately, our research indicates that time series models that consider only this RTE electricity price data do not yield satisfactory forecast results. If the time series approach is to be taken, then other critical information such as historical load levels and fuel prices will need to be considered as well. By incorporating those explanatory variables into the time series model, the forecasting performance is expected to improve.

6. Conclusion

Combining the advantages of neural networks and time series models, an ANN/TSM model has been developed for the MISO day-ahead price forecasting. Results indicate that the combined model outperforms either ANN or time series models alone.

The statistical characteristics of average electricity prices for the French system have also been analyzed. The price series itself does not demonstrate much predictability. It is shown that the ARIMA model built with only the price information does not deliver satisfactory week-ahead forecast results. This evidence clearly suggests the desirability of pursuing a more structured modeling approach.

Portfolio management problems are therefore formulated for two types of entities: a typical U.S. GenCo, and the RTE. A structural price forecasting model tailored for these two portfolio management problems is then proposed.

7. Future Work

Task 1: Use structure-based sensitivity analysis to aid price and congestion forecasting

Currently, price forecasting by purely statistical models is progressing well in this project. However, the problem for purely statistical models is that the inner structural factors in power markets cannot be fully explored. Moreover, congestion and network topology are rarely included in statistical models, but they cannot be ignored due to their significant influence on electricity prices.

In future work we will make use of the AMES wholesale power market test bed developed by H. Li, J. Sun and L. Tesfatsion [15] to explore the impact of demand bids and supply offers on locational marginal prices. These structurally-derived price sensitivities will then be used to facilitate the forecasting of electricity prices and transmission line congestion. With the structural analysis of power markets, the forecasting is expected to become more informative and accurate. The results will be compared with purely statistical forecasting tools.

Task 2: Develop the proposed structural model for price forecasting

As elaborated in Section 5, a structural model for price forecasting will be developed in the next stage of the project specifically tailored to the portfolio management problems facing two types of entities: a typical U.S. GenCo; and the RTE . A portfolio optimization method will be developed to solve the combined forecasting/portfolio management problem proposed for each type of entity. Performance tests will be conducted to evaluate the effectiveness of the combined forecasting and portfolio management tools. For the RTE, this performance testing will require the receipt of additional types of critical data such as load levels and fuel prices.

8. Project Accomplishments

Journal and conference papers:

N. P. Yu, C.-C. Liu and J. Price, "Evaluation of Market Rules Using a Multi-agent System Method," submitted to *IEEE Transactions on Power Systems*.

H. Li and L. Tesfatsion, "The AMES Wholesale Power Market Test Bed: A Computational Laboratory for Research, Teaching, and Training", *Proceedings*, IEEE Power and Energy Society General Meeting, Calgary, Alberta, CA, July 26-30, 2009, to appear.

Q. Zhou, L. Tesfatsion, and C. C. Liu, "Scenario Generation for Price Forecasting in Restructured Wholesale Power Markets," *Proceedings*, IEEE Power Systems & Exposition Conference, Seattle, WA, March 15-18, 2009, to appear.

N. P. Yu, C.-C. Liu, and L. Tesfatsion, "Modeling of Suppliers' Learning Behaviors in an Electricity Market Environment," *International Journal of Engineering Intelligent Systems*, vol. 15, no. 2, pp. 115-121, 2007.

H. Liu, L. Tesfatsion, and A. A. Chowdhury, "Derivation of Locational Marginal Prices for Restructured Wholesale Power Markets", *Journal of Energy Markets*, vol. 1, no. 1, Spring 2009, 1-25.

N. P. Yu and C.-C. Liu, "Multi-agent System Applications in Power Systems," working book chapter to appear in Volume III: Advanced Techniques and Technologies: Facts and A.I. Part two: Artificial Intelligence Techniques.

Presentations:

L. Tesfatsion, "Auction Basics for Wholesale Power Markets: Objectives and Pricing Rules", to be presented at the IEEE Power and Energy Society General Meeting, Calgary, CA, July 26-30, 2009.

H. Liu, L. Tesfatsion, and A. A. Chowdhury, "Locational Marginal Pricing Basics for Restructured Wholesale Power Markets", to be presented at the IEEE Power and Energy Society General Meeting, Calgary, CA, July 26-30, 2009.

Q. Zhou, L. Tesfatsion, and C. C. Liu, "Scenario generation for price forecasting in restructured wholesale power markets," presented at the IEEE Power Systems Conference and Exposition, Seattle, WA, March 15-18, 2009.

9. References

[1] N. Amjady, "Day-ahead price forecasting of electricity markets by a new fuzzy neural network," *IEEE Transactions on Power Systems,* vol. 21, no. 2, pp. 887-896, May 2006

[2] Y. Y. Hong and C. Y. Hsiao, "Locational marginal price forecasting in deregulated electricity markets using artificial intelligence," *IEE Proceedings-Generation Transmission and Distribution*, vol. 149, no. 5, pp. 621-626, Sept.2002.

[3] Y. Y. Hong and C.-Y. Hsiao, "Locational marginal price forecasting in deregulated electric markets using a recurrent neural network," *Proc. Power Engineering Society Winter Meeting 2001*, pp. 539-544.

[4] G. Li, C. C. Liu, C. Mattson, and J. Lawarrée, "Day-Ahead electricity price forecasting in a grid environment," *IEEE Transactions on Power Syst.*, vol. 22, pp. 266-274, Feb. 2007.

[5] S. Deng, "Pricing electricity derivatives under alternative stochastic spot price models," *Proceedings of the 33rd Hawaii international conference on system sciences*, vol. 4, Jan. 2000.

[6] H. Zareipour, C. A. Cañizares, K. Bhattacharya, and J. Thomson, "Application of public-domain market information to forecast Ontario's wholesale electricity prices," *IEEE Transactions on Power Syst.*, vol. 21, pp. 1707-1717, Nov. 2006.

[7] M. Shahidehpour, H. Yamin, and Z. Li, Market Operations in Electric Power Systems: Forecasting, Scheduling, and Risk Management. New York: Institute of Electrical and Electronics Engineers, Wiley-Interscience, 2002.

[8] J. H.Zhao, Z. Y. Dong , Z. Xu, and K. P. Wong, “A statistical approach for interval forecasting of the electricity price ,” *IEEE Transactions on Power Syst.*, vol.23, pp. 267-276, May. 2008

[9] J. J. Lucia and E. S. Schwartz, “Electricity prices and power derivatives: evidence from the Nordic power Exchange,” *Review of Derivatives Research*, vol. 5, pp. 5-50, 2002.

[10] G. E. P. Box and G. H. Jenkins, *Forecasting and Control*. San Francisco: Holden Day, 1970.

[11] Q. Zhou, L. Tesfatsion, and C. C. Liu, "Scenario generation for price forecasting in restructured wholesale power markets," *Proc. IEEE Power Systems & Exposition Conference,* Seattle, WA, March 15-18, 2009, to appear.

[12] EPRC Forecasting homepage : http://www.econ.iastate.edu/tesfatsi/EPRCForecastGroup.htm

[13] R. T. Rockafellar and S. Uryasev, “Optimization of conditional value-at-risk,” *Journal of Risk*, vol. 2, no. 3, pp. 21-42, 2000.

[14] J. Sun and L. Tesfatsion, “An agent-based computational laboratory for wholesale power market design,” *Proceedings of the IEEE Power Engineering Systems General Meeting*, June, 2007.

[15] H. Li, J. Sun, and L. Tesfatsion, AMES Wholesale Power Market Test Bed, homepage www.econ.iastate.edu/tesfatsi/AMESMarketHome.htm