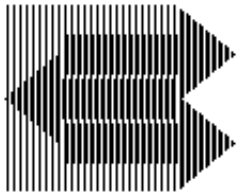

MIT EL 00-004

Energy Laboratory

Massachusetts Institute of Technology

Bid-based Stochastic Model for Electricity Prices: The Impact of Fundamental Drivers on Market Dynamics

November 2000

Bid-based Stochastic Model for Electricity Prices: The Impact of Fundamental Drivers on Market Dynamics

Petter Skantze, Andrej Gubina, and Marija Ilic

Energy Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139-4307

Energy Laboratory Publication # MIT_EL 00-004

November 2000

For further information please contact Marija Ilic at 617-253-4682 or via email at ilic@mit.edu

Abstract

The bid based model developed in this report is intended as a fundamental model for electricity price dynamics, to be used in a wide range of applications. The emphasis was placed on incorporating the unique characteristics of electricity prices, including seasonality on multiple time scales, lack of load elasticity, stochastic supply outages, strong mean reversion, and stochastic growth of load and supply. Principal component analysis is applied in the model in order to capture intra-day dynamics, while at the same time greatly reducing the computational complexity.

The model is calibrated on actual load and price data form the New England ISO. We also propose extensions of the model to deal with instances of multiple spot markets connected by transmission lines. Through simulations we illustrate how the model can be used to estimate the value of transmission rights in a two-market environment. It is also shown how the model can be used by a for-profit transmission provider in order to make optimal investment decisions in new transmission capacity. Finally, an extension of the model is proposed to simulate the interaction between technical innovation and long-term price dynamics in electricity markets.

Index

1 Problem of Interest

The restructuring of the electric utilities industry has forced industry participants to rethink their approach to a number of decision processes, including investment, speculation and risk management decisions in electricity markets. These problems all require knowledge of the future behavior of prices in the market. This has led to a push in the industry as well as in academia to develop viable models describing the stochastic behavior of electricity prices. Since the models are being applied to a wide variety of questions, there is no 'perfect' model. The model has to be evaluated in context of its application. Since a company may be required to coordinate its decision process in asset investments, risk management and speculation, there is however a distinct advantage to ensuring that the models that are used in each case are at least not conflicting in their estimates, but represent their users' best understanding of the marketplace.

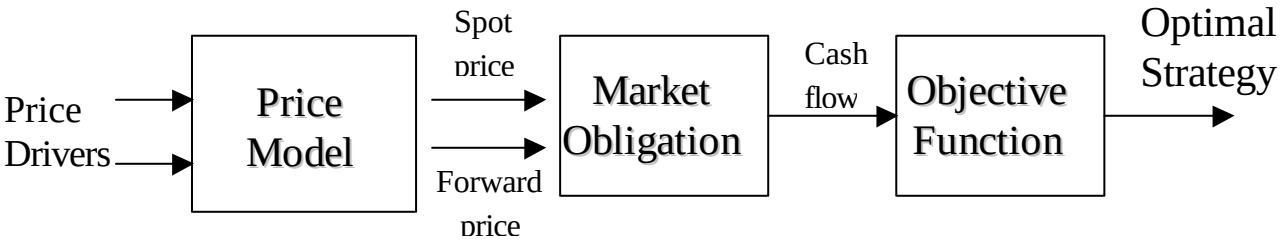

Fig. 1: Dependence of optimal decisions on the electricity price models

We begin by formulating the problem in terms of a generic set of future cash flows. The cash flows can be a result of an investment into physical assets, a contractual obligation, or an exchange-traded derivative. Future cash flows are a function of future market prices. Using the results of the price model in the cash flow function allows us to estimate the joint probability distribution of future cash flows, Fig. 1. This distribution, in turn, is fed into the users objective function, which becomes the criteria for optimal decision-making.

Depending on the objective of the user, a number of approaches for modeling price dynamics are available. In this section we separate these approaches into six broad categories. The approaches differ in complexity, detail and objective of the models.

1.1 Quantitative Modeling of Electricity Prices - [8],[21],[23],[24]

Objectives: To characterize the stochastic properties of commodity prices over time, specifically, to attempt to derive the variance and covariance of commodities prices. With this information the user is then able to price a broad category of financial derivatives, as well as to perform basic risk management functions.

Characteristics: The models used in quantitative modeling are usually generic in nature. The user attempts to find the lowest order model possible to accurately describe the stochastic properties of the commodity.

Advantages: Since the model is generic, the user does not require an in depth understanding of the economic or physical relationships involving the production and trading of the

commodity. Calibration schemes are also standardized and can be duplicated across multiple commodities.

Disadvantages: This category of models is calibrated using historical spot and forward market data, and when available, using implied volatilities from historical models. It requires the availability of a significant amount of price history data. In case of electricity, changes in the regulatory environment have made historic prices invalid for calibration purposes, leaving the user with an inadequate set of training data for the models.

1.2 Production (Cost) Based Modeling of Electricity Prices

Objectives: To model future electricity prices based on detailed models of the cost structure of individual products. This information is used to create a cost-based supply curve. Combined with estimates of future demand, this can be used to generate price estimates.

Advantages: Marginal cost information is generally available for all producers in a region. The creation of a supply function is therefore a relatively straightforward exercise. Furthermore, the cost can be linked to underlying fuel prices by using heat rate estimates on the unit. This allows the user to model the interaction of fuel and electricity prices.

Disadvantages: Cost based modeling ignores the strategic bidding practices of market participants. The effect of market power is likely to raise prices above cost-based levels. The cost-based models can therefore rarely be calibrated to correspond to actual observed prices in the market.

1.3 Economic Equilibrium Models of Electricity Prices [26],[27],[28],[29]

Characteristics: As means of incorporating strategic bidding into cost-based models, theories such as Cournot pricing are applied to the generation stack. At a given load level one can then solve for an equilibrium markup of bids above cost based levels. This markup will generally increase as a function of market concentration.

Advantages: By applying game theory type models it is possible to explain why prices rise above cost-based levels. This approach is useful in predicting expected price levels in markets with no price history, but known supply costs and market concentration.

Disadvantages: These models produce equilibrium price levels. However, electricity markets are constantly evolving, driven by stochastic demand and supply, and therefore never settle to equilibrium levels. In applications such as risk management, understanding the dynamic behavior of prices is crucial. In this case economic equilibrium models offer little insight.

1.4 Agent-based Modeling of Electricity Prices [15],[16]

Characteristics: Agent-based models attempt to capture the strategic behavior of investors (agents) on the marketplace. To approximate the dynamics of the market, participants are separated into groups, each with their own objective function. Based on the objective function

and observation of current price levels, a decision rule is defined for each group. These rules can be highly nonlinear in nature. Finally the system is simulated under various inputs.

Advantages: In contrast to cost-based and equilibrium models, agent-based models address the effect of market power both on the overall price markup, and the inter-temporal dynamics of price. The variety of dynamic behavior, which can be captured with a relatively small number of strategies, is impressive. The approach, for example, allows the user to study the impact of factors, such as collusion, on the overall system price.

Disadvantages: While agent based modeling is useful for studying the qualitative behavior of markets, it is much more challenging to get relevant quantitative results. To do so, one would need a consistent method of calibrating the parameters of the decision processes based on historical data. This seems like an overwhelmingly difficult task.

1.5 Experimental Modeling of Electricity Prices [25]

Characteristics: In the experimental modeling approach, a group of people are gathered and assigned assets and obligations in the market place. They then simulate the behavior of the market by submitting bids, which are used to clear the market.

Advantages: The organizer of the experiment has full control over the parameters and can change factors such as market concentration or number of participants in order to observe the effects on the spot price.

Disadvantages: Experimental modeling is extremely difficult to map into a real marketplace. To get reliable results, one would need to convince actual marketers to participate in the process, and even then it is questionable if they would betray their actual trading strategies.

1.6 Fundamental Modeling of Electricity Prices [17]

Objectives: Determining the stochastic properties of commodities prices.

Characteristics: In the fundamental modeling approach, price dynamics are described by modeling the impact of important physical and economic factors on the commodity price. The model seeks to capture basic physical and economic relationships present in the production and trading of the commodity. By explicitly adding these constraints, one can increase the complexity of the model while decreasing the requirements on the available training data.

Advantages: By relaying the dynamics of the commodity price to the fundamental drivers, we gain a new set of training data. If the fundamental inputs are directly observable, we can use historical inputs in order to calibrate the model parameters. In the case of electricity this can be a crucial difference. Currently there is only 1-2 years of relevant electricity price history available (depending on location). However, if we choose temperature (a major determinant of electricity demand) as a fundamental driver, we have decades worth of historical measurements available.

Disadvantages: In creating the fundamental model we make specific assumptions about economic relationships in the marketplace. The price projections generated by the models are therefore very sensitive to violations of these assumptions. Thus there exists a significant modeling risk in the application of the fundamental approach.

2 Bid-based Stochastic Model for Electricity Prices

In this section we develop a Bid-based Stochastic Model (BSM) of the evolution of prices on electricity spot markets. We assume that the spot market operates as a double auction, similar to the rules of the California Power Exchange. The model can be modified to account for variations in the auction procedure.

We design the model to be applicable to hedging, speculation or investment decisions in electricity markets. As such, it focuses on quantifying the uncertainty of future price movements. We have used fundamental modeling approach, where the fundamental drivers are load and supply shifts. The model captures the most critical characteristics of demand (load) and supply as outlined below in electricity market.

2.1 Load Characteristics

1. **Load Elasticity**: We assume electricity demand to be completely inelastic (i.e. independent of market clearing price). This may appear to be a strong assumption, but in the current state of deregulation, few end users actually observe real time price movements.
2. **Seasonality**: Seasonality is a major driver for electricity demand. We observe seasonality over the daily, weekly, and yearly cycles.
3. **Mean reversion**: One can observe temporary spikes in electricity demand, often induced by extreme weather conditions. However, these spikes are not sustainable and demand reverts back to normal levels within a few days.
4. **Stochastic growth**: Growth in electricity demand is driven in part by trends in the overall economy. The growth is therefore hard to predict over longer time horizons, and must be considered stochastic.

2.2 Supply Characteristics

1. **Supply Elasticity**: In contrast to load, electricity producers are price responsive. The supply characteristic is mainly a function of generation technology, as operating cost can vary significantly with the type of generator used. Market power and strategic bidding also have an impact on the shape of the supply bid function.
2. **Stochastic Availability of Generation**: Due to unexpected equipment failure or because of planned maintenance, generators are taken offline from time to time. The effect of such sudden jumps in the availability of supply on the market-clearing price can be significant.
3. **Uncertain fuel cost**: Changes in the price of fuels such as oil and gas will affect the way generators bid into the market.
4. **Unit Commitment**: Nonlinear characteristics in the generator cost functions, such as startup costs and minimum run times, result in intra-day supply bid curve shifts.
5. **Import/Export**: Producers and consumers bidding into the market from outside its geographic limits can cause significant price shifts.

6. **Inter-Market**: Prices on related markets such as markets for capacity and ancillary services, represent opportunity costs for power suppliers. Hence there is a strong interaction between prices on these markets and price in the energy market.

2.3 Price as a function of Load and Supply

In our model we characterize spot price as a function of two variables; L representing load shifts, and b representing supply shifts. These variables can be interpreted as follows.

Load: We assume load bids are inelastic. Therefore L_k is selected to represent the market clearing volume of the exchange for hour k.

Supply: In contrast to load, supply bids have significant price elasticity. The elasticity (or the inverse of the slope of the supply curve) varies significantly with the clearing quantity. In general, supply will be highly elastic at low demand levels, and gradually become more inelastic as demand increases.

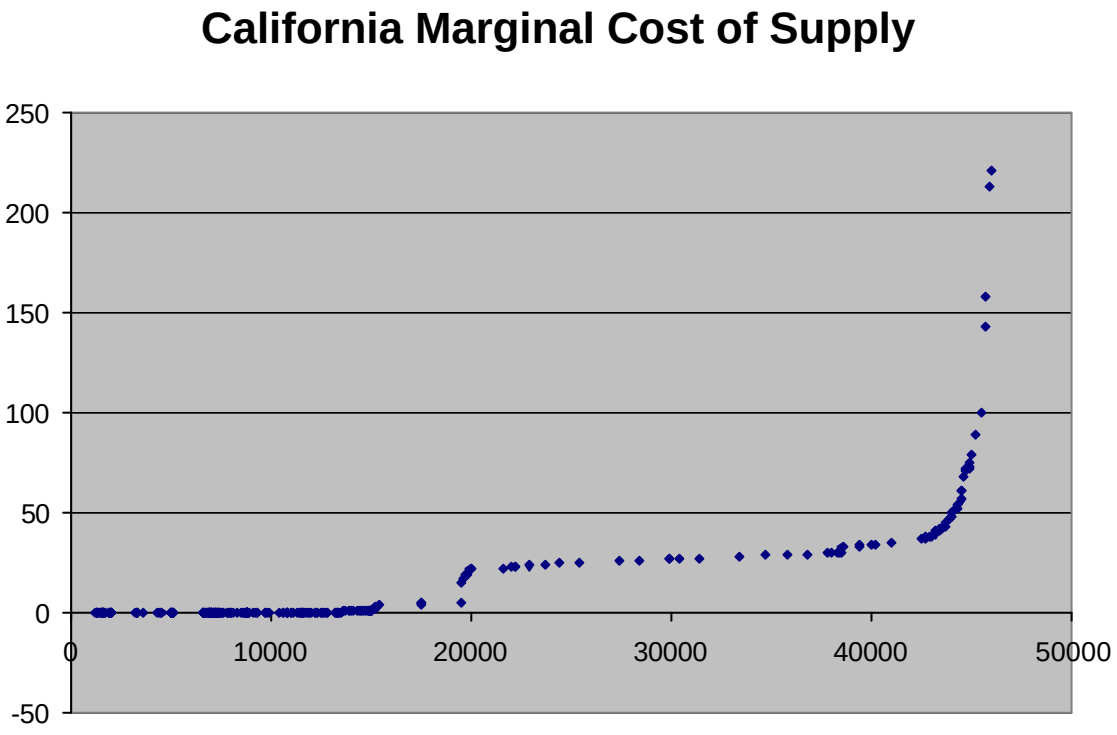

Fig. 2: Ascending-ordered marginal costs of generators for California

We can explain this characteristic of electricity supply in two ways. First we examine the cost structure of the underlying generators. Fig. 2 shows the 'stack' for California, created by ordering the generators from lowest to highest marginal cost. As seen in the figure the cost function is relatively flat for low demand levels, when the load is served mainly by hydro and nuclear plants. In the medium range we see a slight cost increase as efficient fossil plants are utilized. In the high demand range, inefficient peaking plants are dispatched, and the operating cost escalates significantly.

Another approach is to view the supply bids from a game-theoretic perspective. At low demand levels there is a high ratio of available generation capacity to electricity demand. Hence the market will be competitive and highly price responsive. As load approaches the total installed

capacity of the market, the few non-committed generators have a high degree of market power, and can withhold their capacity to push prices upward. In-depth analysis of market power and strategic bidding in power markets can be found in [15] and [16].

Fig. 3 shows a plot of the supply curves submitted to the California Power Exchange during a 24-hour period.

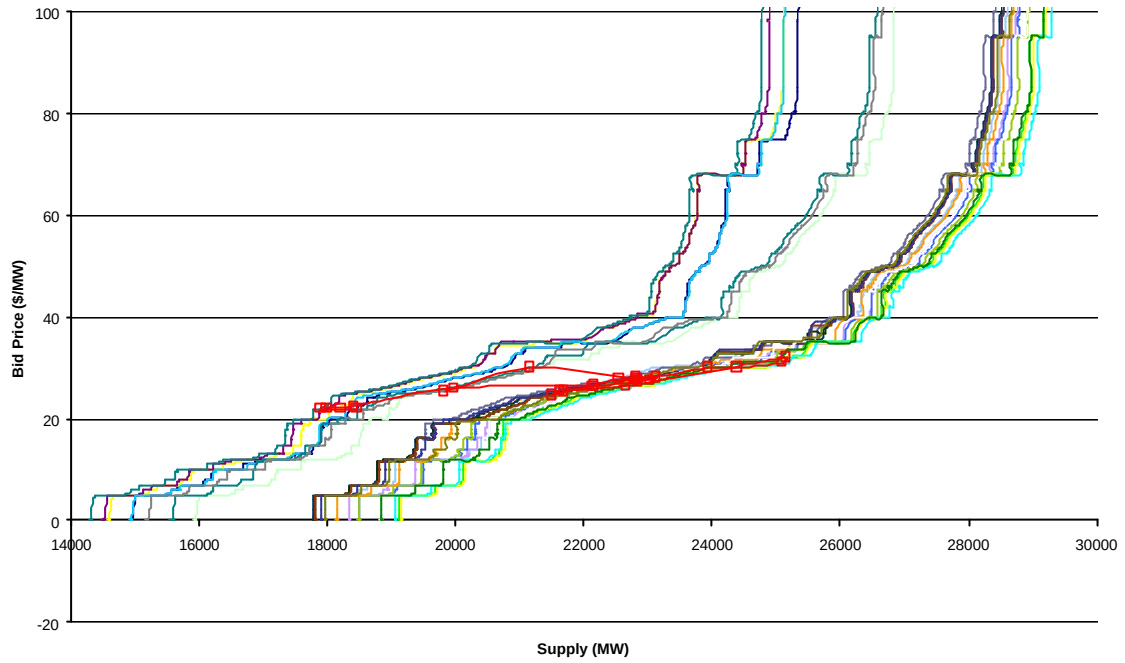

Fig. 3: Daily development of the supply curves and MCP

When comparing the cumulative bid curves submitted at different hours we find that the basic shape of the bid curve is preserved over time. This allows us to reduce the complexity of the supply model. We fix the shape of the bid curve and model its temporal shifts as a stochastic process. Specifically we chose an exponential function to approximate the shape. Price in hour k can then be written as,

$$P_k = e^{aq_k + b_k}$$

where a is a fixed parameter characterizing the slope of the bid curve, q_k is the market clearing quantity in hour k, and b_k denotes the position (or shift) of the curve. Next we add the constraint that demand bids are inelastic. The market clearing quantity q_k must then always be equal to the system load L_k. We can now write market clearing price in terms of our two fundamental drivers, load and supply:

$$P_k = e^{aL_k + b_k}$$

This approach reduces the complexity of the problem by constraining the number of free variables on the supply side. The downside of this assumption is that we risk misrepresenting the shape of the supply curve in certain regions. There are three major parts of the supply curve, lower, middle, and upper. The result of fitting each of them with an exponential is shown in Fig. 4. Since we are concerned with the price range, corresponding to the actual (market clearing) load at that hour, we have selected the second exponential (denoted as exp2), which best approximates the middle part of the bid curve.

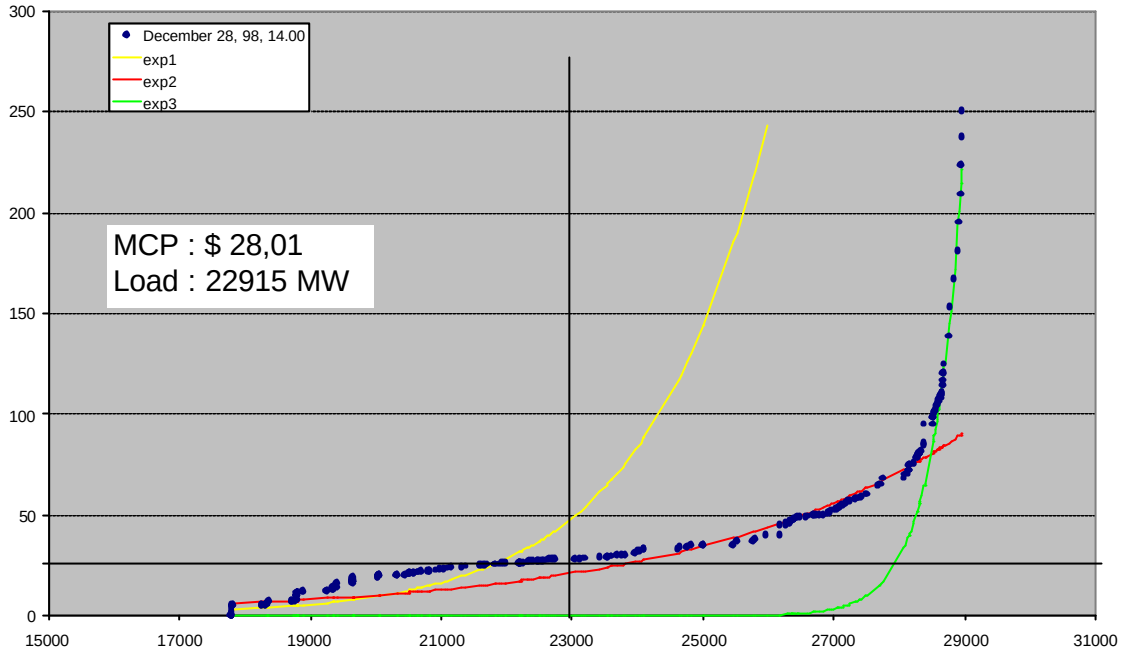

Fig. 4: Exponential fitting of the aggregate hourly supply curve

The next step is to postulate stochastic models for the evolution of the fundamental drivers. In order to keep track of the variables and parameters evolving at different time scales we use the following notation:

Subscript	Meaning	Superscript	Meaning
d	evolves at daily rate	L	Belongs to load process
m	evolves at monthly rate	b	Belongs to supply shift process
none	constants	δL	Applies to load mean process δ^L
		δb	Applies to supply mean process δ^b

The following section will outline the models used and the reason for choosing that specific form. In later sections we present step-by-step descriptions on how model parameters were calibrated based on historical market data.

2.4 Stochastic Load Model

We listed the four characteristics of electricity demand which we wanted to capture in our model; lack of price elasticity, seasonality, mean reversion, and stochastic growth. The elasticity assumption is already implicit in our formula for market clearing price P_k. The challenge is to incorporate the remaining criteria without making the model too complex for calibration and simulation purposes.

2.4.1 Modeling Demand Seasonality

The three types of seasonality in electricity demand are daily, weekly and yearly patterns. Fig. 5 shows demand in New England for a sample week in May, starting with Monday.

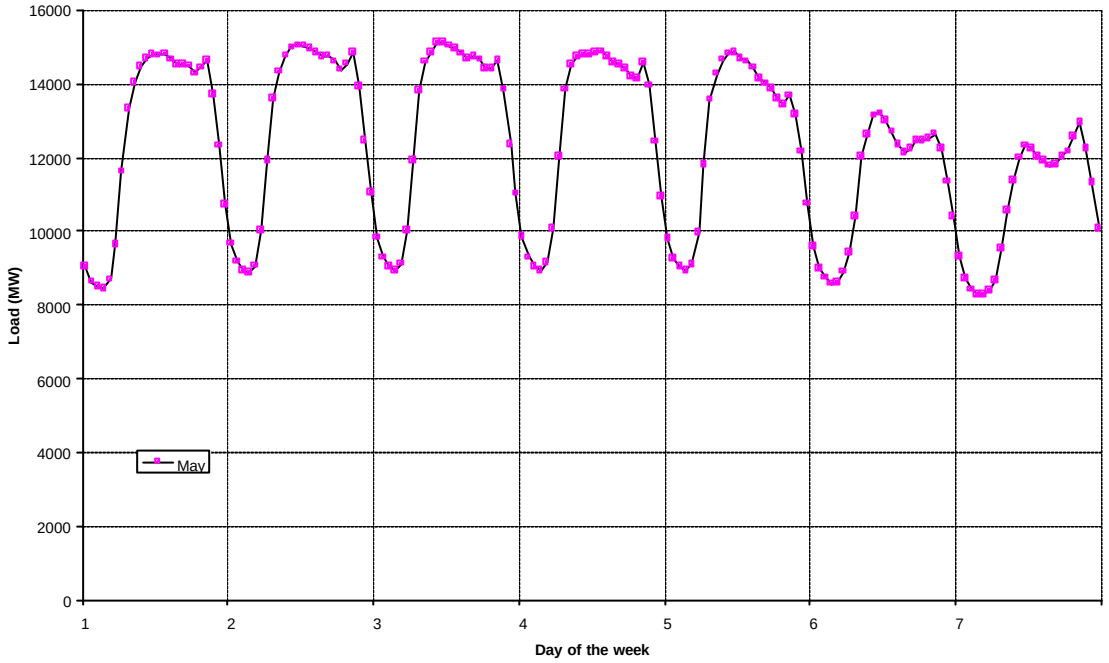

Fig. 5: Load diagram for a week in May, New England

We see that there is a regularly recurring pattern within the weekdays (daily seasonality) and that the weekend consumption pattern is significantly different (weekly seasonality). From here on we will simplify our task by eliminating the weekends and modeling only the weekday loads. This allows us to ignore the weekly seasonality. This simplification is taken directly from the forward markets, which trade weekdays and weekends as separate contracts.

Addressing the daily seasonality is more challenging. We have chosen to denote the daily load as a [24×1] vector $\mathbf{L}_d$, where each component represents an hourly load. This vector is defined as the sum of a deterministic and a stochastic component.

$$\mathbf{L}_d = \boldsymbol{\mu}_m^L + \mathbf{r}_d^L$$

The deterministic component $\mathrm{m_m}^{\mathrm{L}}$ is a [24×12] vector that represents the typical or average monthly load pattern for the day. This component evolves on a monthly time scale, since the typical load pattern for January is significantly different form the typical pattern for August. Fig. 6 shows a plot of average 24-hour load patterns for each month in New England.

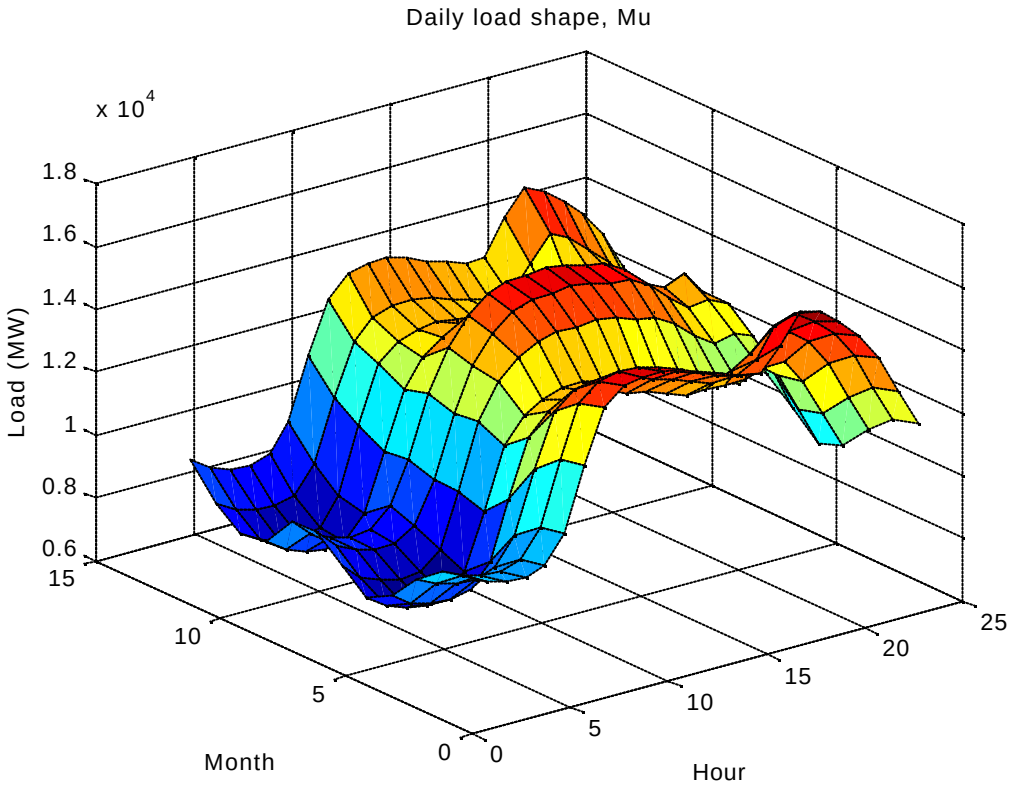

Fig. 6: Average monthly patterns of daily load, $\mathbb{m}_m^L$, New England

2.4.2 Modeling Load Uncertainty

The stochastic component **r** of the daily load pattern is needed to explain any deviation in actual observed load from the pattern given by $\mathbb{m}_m$. In order to achieve this, the vector **r** would have to contain 24 random variables. However in observing actual load patterns, one finds that there is a strong correlation between deviations in consecutive hours.

Intuitively, one could argue that if unusually hot weather causes demand to increase in hour 14, it is very likely to also cause higher demand in hours 15 and 16. To capture this mathematically, we applied principal component analysis (PCA) to the data, [10], [11] and [12].

Principal Component Analysis (PCA) is a method that enables us to describe a set of observations of n variables, which would normally require n dimensional representation, with a reduced set of j variables, $j \leq n$. In other words, PCA addresses the issue of how to characterize a probabilistic space of n dimensions using a reduced set of j basis functions.

Although some information will be lost in this process, PCA enables us to minimize this information loss by choosing the new basis as the best approximation by minimizing the variance of the error. In the original data set, groups of variables often move in the same direction, indicating that more than one variable is describing the same driver. Therefore, a group of variables can be replaced with a single new variable. At the same time, we retain the maximum information (variance) of original observations.

PCA generates a set of new variables, called Principal Components (PC). Each principal component is a linear combination of the original variables (the old basis). All PCs are

orthogonal to each other, forming a new orthogonal basis, so there is no redundant information. A rough explanation of the theory supporting principal components and their derivation is provided in Appendix A.

The output of the PCA algorithm is a set of principal components $\mathbf{v}^{Li}$ and associated weights w^i so that the best approximation of the load vector $\mathbf{L}_d$ in the new basis is given by:

$$\mathbf{L}_d = \boldsymbol{\mu}_m^L + \sum_{i=1}^{j} w_d^{Li} \mathbf{v}_m^{Li} .$$

In this report we will use a single principal component, a monthly [24×1] vector $\mathbf{v}_L{}^m$ to describe load behavior, reducing the load equation to:

$$\mathbf{L}_d = \boldsymbol{\mu}_m^L + w_d^L \mathbf{v}_m^L ,$$

where m_m and $\mathbf{v}_m$ are deterministic parameters and w_d is a daily stochastic process.

The choice of the number of principal components used is a tradeoff between accuracy and complexity. For short-term decision, making such as day-ahead bidding, a single PC may not provide a rich enough sample space. However, when applying the price process to hedging and valuation decisions over months or years, a small basis prevents the problem from blowing up in computational complexity.

Next we need to address how the stochastic component w_d evolves over time. The model we propose is a two factor mean reverting model:

$$e_{d+1}^L - e_d^L = -a^L e_d^L + s_m^L z_d^L$$
$$d_{d+1}^L - d_d^L = k^L + s^{Ld} z_d^{Ld} ,$$

where,

$$e_d^L = w_d^L - d_d^L .$$

2.4.3 Mean Reversion

We can interpret the states e_d^L and δ_d^L in terms of the temporal characteristics of load. The state e_d^L models short term deviations in load, such as those caused by sudden heat waves. These events are generally temporary, and load gradually reverts back to normal levels. The process for e_d^L is therefore chosen to be mean reverting. The parameter α determines the speed of reversion. Fig. 7 illustrates how the short-term spikes in load quickly revert to the long-term mean. For clarity, here the mean is being modeled as a monthly rather than a daily process. This time scale separation between the states is a method, which further simplifies the application of the model, and its advantages and disadvantages are described in the calibration section.

2.4.4 Stochastic Growth

As e^L reverts to zero, the weight w^L reverts to δ^L, or the "normal" load level. However since the power system is never at equilibrium, the normal load level is in itself a stochastic process. The δ^L process characterizes the stochastic growth of load over time. This growth could be positive or negative for any given period of time, and there can be significant uncertainty to the rate of growth, captured in the long-term volatility parameter $\sigma^{L\delta}$. The long-term growth of load in New England is illustrated in Fig. 8.

The effects of the structure of the stochastic process on the mean and variance of future load is explored in detail in the section on simulation.

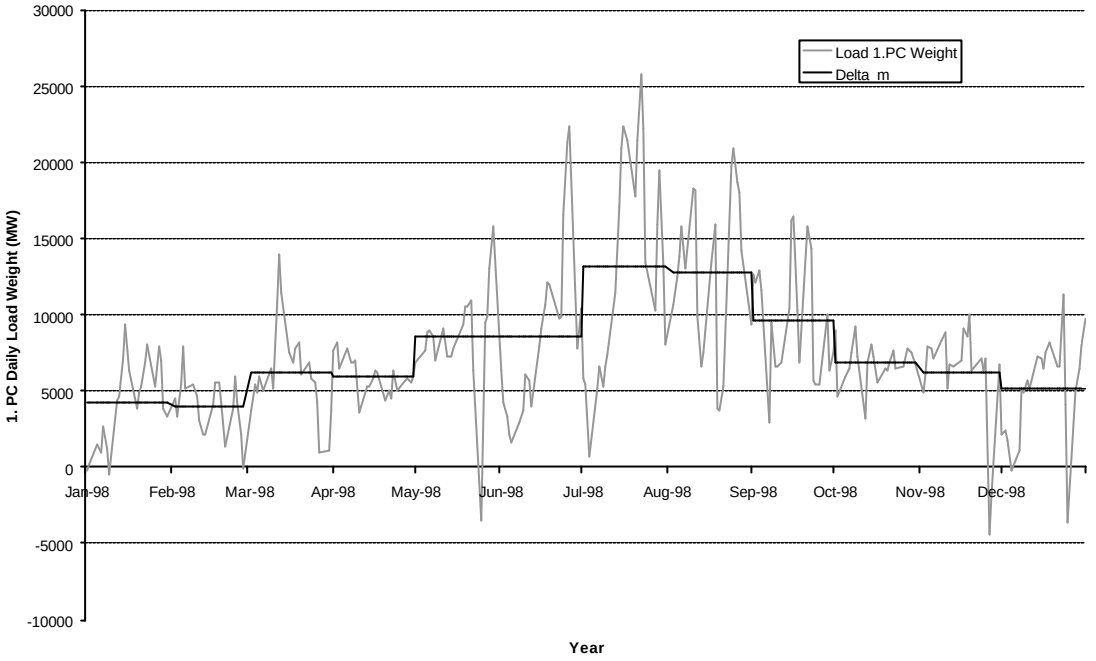

Fig. 7: Reversion of the load weights w_d^L to long-term mean $\delta_m{}^L$, year 1998

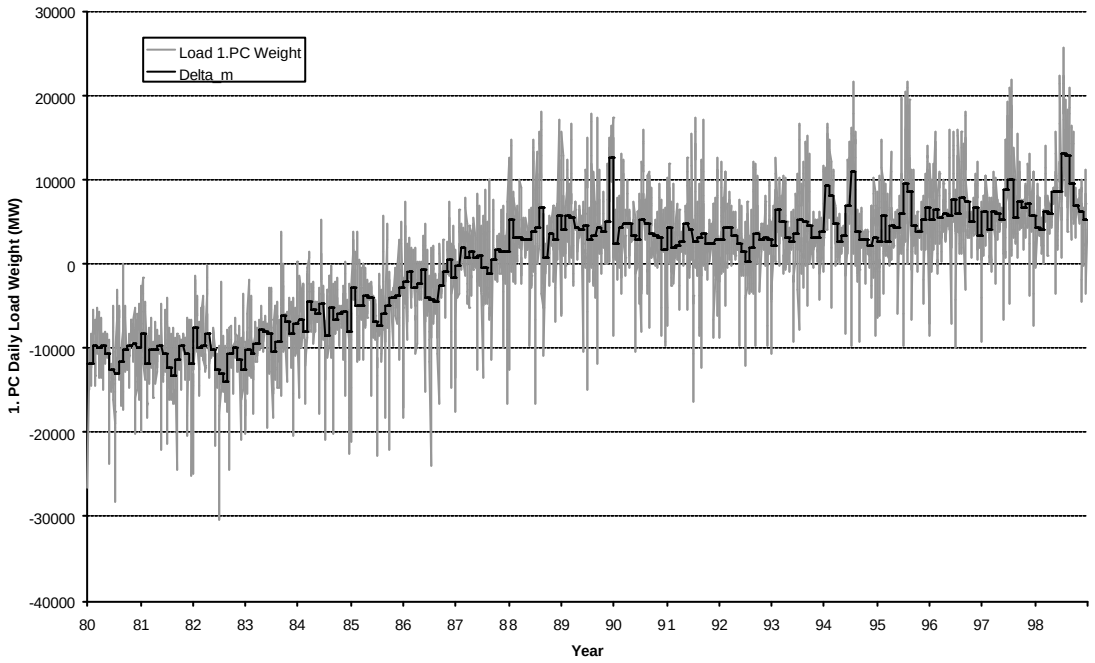

Fig. 8: Load weights w_d^L, and long-term mean $\delta_m{}^L$, New England

2.5 Stochastic Supply Process

Recall our underlying price model as a function of load and supply states L_k and b_k,

$$P_k = e^{aL_k + b_k}.$$

This implies that the aggregate supply bid curve is an exponential function of fixed shape (given by a), which shifts over time.

Let us consider the input drivers, which could cause the supply curve to shift:

1. **Fuel price**: An increase in fuel prices would force suppliers to increase their bids into the spot market in order to remain profitable. An increase in the fuel price would therefore be accompanied by a positive shift in b_k.
2. **Unit Outages and Scheduled Maintenance**: The withdrawal of a generation unit from the market, whether through an unexpected failure or a scheduled maintenance, causes a significant shifts in the supply bid function. The size and duration of these shifts, as well as the frequency of their occurrence, is technology dependent.
3. **Gaming and Strategic Bidding**: It has been shown that generators with significant market share may increase their profits by unexpectedly removing part of their generation assets from the market, forcing up price and increasing the payoff for the remaining units [20]. Such an event can be characterized by a positive shift in b_k.
4. **Unit Commitment Decisions**: While generators are often modeled as having well behaved quadratic cost functions, in reality there are significant non-standard costs and constraints associated with starting up and shutting down a generator. Translating such constraints into bids will cause generators, even though they may have no market power, to deviate from a marginal cost-bidding scheme.

We now attempt to translate the impact of these drivers into a stochastic process for the supply process. As with the load we characterize supply by a [24×1] daily vector $\mathbf{b}_d$ containing hourly supply levels. This daily vector is then decomposed into its deterministic and random components:

$$\mathbf{b}_d = \boldsymbol{\mu}_m^b + \mathbf{r}_d^b$$

2.5.1 Seasonality of Supply

Although less pronounced than the load, the supply process does exhibit seasonality over multiple time scales. The most pronounced are monthly and intra-day seasonality.

1. On a **monthly** time scale, we see the scheduling of maintenance. In a practice that has carried over from the regulated industry, units are regularly scheduled for maintenance during the off-peak seasons (mainly fall and spring), when demand spikes are unlikely. From the modeling perspective this creates a repeating twelve-month pattern of supply bid shifts.

   The fuel markets feeding the generators also experience seasonality on this time scale, mainly due to seasonal demand on oil and gas. Seasonal fuel prices therefore create a second pattern of supply shifts. The aggregate effect of these repeating yearly patterns is captured by the deterministic shifts in the monthly parameter m_m.
2. The second type of seasonality experienced in the supply process is **intra-day**, where we observe repeating 24-hour patterns of supply curve shifts. This type of seasonality is mainly contributed to unit commitment decisions done by the suppliers. The operator of the unit will estimate a day ahead of time the hours during which it will be profitable to run the unit, based on the startup/shutdown constraint of the generator. Once this decision is made he may

choose not to submit bids into the remaining hours, so as to not risk being scheduled and incurring a substantial startup cost. The result is a repeated pattern of shifts illustrated in Fig. 3. This behavior is captured by the daily shape of the vector m_m^b.

2.5.2 Modeling Supply Uncertainty

In modeling the random component of the daily supply vector we again apply principal component analysis, using a first order approximation (one PCA vector),

$$\mathbf{b}_d = \boldsymbol{\mu}_m^b + w_d^b \mathbf{v}_m^b .$$

The shape of the principal component, Fig. 17 and Fig. 18, is strongly related to the unit commitment decision of the generators.

The process defining the evolution of the weights is similar to that used for the load process:

$$e_{d+1}^b - e_d^b = -a^b e_d^b + s_m^b z_d^b$$
$$d_{d+1}^b - d_d^b = k^b + s^{bd} z_d^{bd} ,$$

where,

$$e_d^b = w_d^b - d_d^b .$$

The mean reverting component e_d^b reflects the transient characteristics of the supply process. This includes short-term fuel price spikes and short-term gaming. These effects are temporary and die out at a rate governed by α. The reversion of supply to its long-term mean is illustrated in Fig. 9:

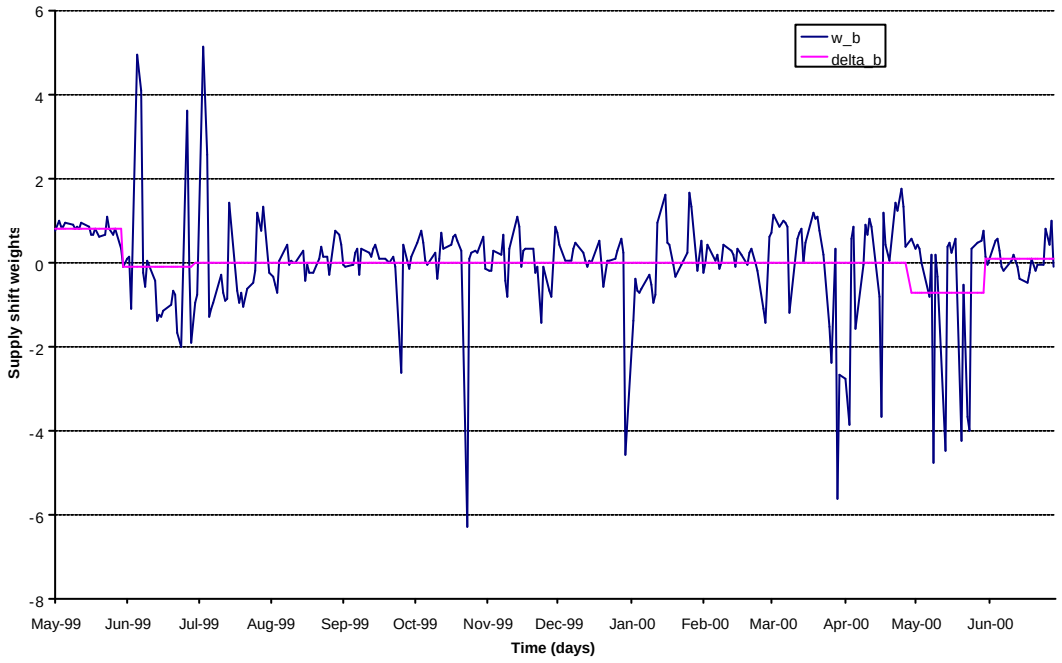

Fig. 9: Reversion of supply weights w_d^b to long-term mean δ_m^b

The non-reverting component δ_d^b models the long-term availability of generation. This will include any new installed or retired capacity on the market.

2.5.3 Modeling Unit Outages

So far our supply model has included smooth changes in the behavior of the supply bid curve, which can be characterized by an Ito process [9]. However, there exist a set of high impact, low probability events that cannot be approximated trough random walk type models. One such event is the unexpected failure of a major generator in the market. There are a number of unknowns associated with this event:

1. The probability of an outage in a given day.
2. The impact of the outage on market price.
3. The duration of the outage.

The answers to all three of these questions generally depend on the type of generation technology.

In our model we address these problems by adding a new factor to the supply process:

$$\mathbf{b}_d = \boldsymbol{\mu}_m^b + w_d^b \mathbf{v}_m^b + \sum_i p_d^i \, ?_m^i \, .$$

1. The probability of an outage occurring in a given day is modeled as a random incidence process, specifically a Bernoulli process. The probability of the outage occurring in a given day is independent of all other time intervals. This is denoted by the variable π^i, where $\pi^i = 0$ under normal conditions, and $\pi^i = 1$ when there is an outage in a plant of technology i.

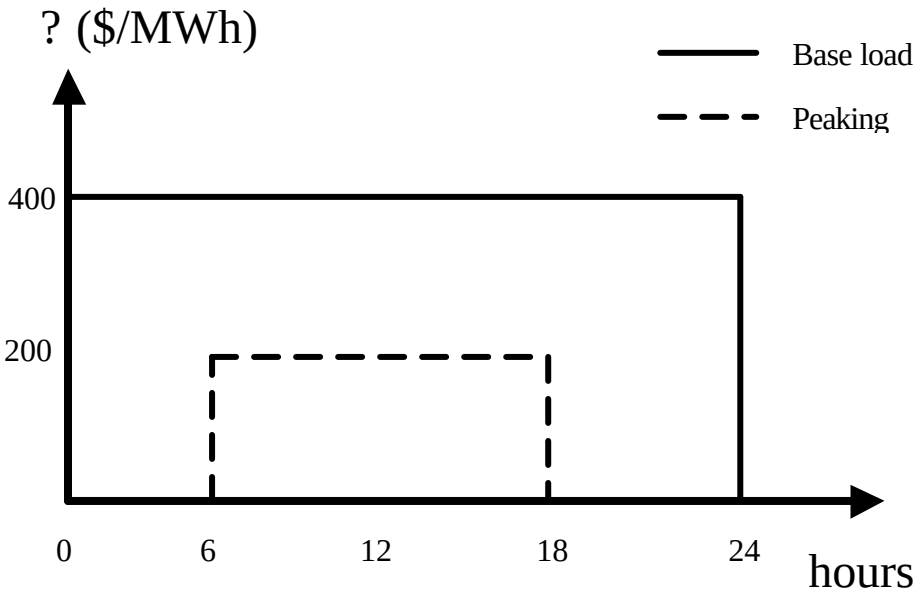

Fig. 10: Daily shape ψ for a 400 MW base load plant and a 200 MW peaking plant

2. The impact of the outage on market clearing price will depend on the capacity of the unit, and its characteristic operating schedule. An outage in a plant, which is scheduled to deliver at full capacity, results in a positive shift in $\mathbf{b}_d$, equal to the capacity of the plant. If however the plant was not scheduled to deliver (ie. bid in above market clearing price) then there is no effect on the price. The probability of a plant being selected to produce in a given hour generally depends on its cost structure, and therefore on its technology. We

incorporate this effect by assigning a [24×1] vector y_m^i to each technology i. The vector denotes the capacity of the unit as well as the likelihood of the unit being scheduled in a given hour. Fig. 10 denotes the daily shape of ψ for two types generation technologies, a 400 MW base load plant, and a 200 MW peaking plant.

3. The outage duration is modeled as a deterministic minimum outage time plus a stochastic Bernoulli component. By combining this process with the random arrival time of the outage (described in (1)), we can characterize the process for the state π_d^i as a Markov chain, as illustrated in Fig. 11. Here the numbers next to the arrows designate the probability of a state transition for a given day. The probability of going from normal operation to an outage for each day is given by λ_{out}. The probability of the unit returning on-line after the minimum outage period is given by λ_{in}. For the case shown the minimum outage time is four days.

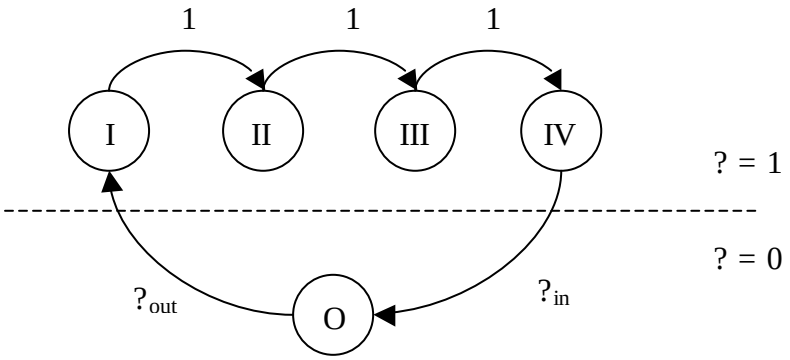

Fig. 11: Modeling of outage duration for the state π_d^i as a Markov chain

2.5.4 Modeling Scheduled Maintenance

Scheduled maintenance can be modeled in the same manner as unit outages. The only difference is that the π_d becomes a deterministic rather than a stochastic state variable.

2.6 Summary of the Bid-based Stochastic Price Model

The following is a compact summary of the mathematical model underlying the Bid-based Stochastic Model.

Spot Price Model:

Hourly price: $P_h = e^{aL_h + b_h}$

Daily 24-hour vector of prices: $\mathbf{P}_d = e^{a\mathbf{L}_d + \mathbf{b}_d}$

Load Model:

$\mathbf{L}_d = \boldsymbol{\mu}_m^L + w_d^L \mathbf{v}_m^L,$

$e_{d+1}^L - e_d^L = -a^L e_d^L + s_m^L z_d^L$

$d_{d+1}^L - d_d^L = k^L + s^{Ld} z_d^{Ld},$

where,

$e_d^L = w_d^L - d_d^L.$

Supply Model:

$\mathbf{b}_d = \boldsymbol{\mu}_m^b + w_d^b \mathbf{v}_m^b + \sum_i p_d^i y_m^i.$

$e_{d+1}^b - e_d^b = -a^b e_d^b + s_m^b z_d^b$

$d_{d+1}^b - d_d^b = k^b + s^{bd} z_d^{bd},$

where,

$e_d^b = w_d^b - d_d^b.$

and π_d is Markov process with parameters λ_{out} and λ_{in} as described in the previous section.

3 Calibration of the Bid-based Stochastic Model

Calibration of the Bid-based Stochastic Model (BSM) consisted of several steps. It used historical data on aggregate hourly market clearing load in the system and on market clearing price, as determined at the New England power system electricity market for the particular hour. The load history encompassed 18 years of hourly load data (1980-1998), whereas supply data (hourly market clearing price) were only available for the 14-month period between May 1998-June 2000. Both fundamental processes were independently calibrated using respective data.

Calibration of BSM is a two-part process, as shown in the flowchart on Fig. 12. In the first part, data is gathered from the sources, filtered and reformatted. This is mainly done by hand using spreadsheet program, or simple filtering programs designed in Matlab. Particular steps differ between load and supply process models, as the models play different roles in the BSM.

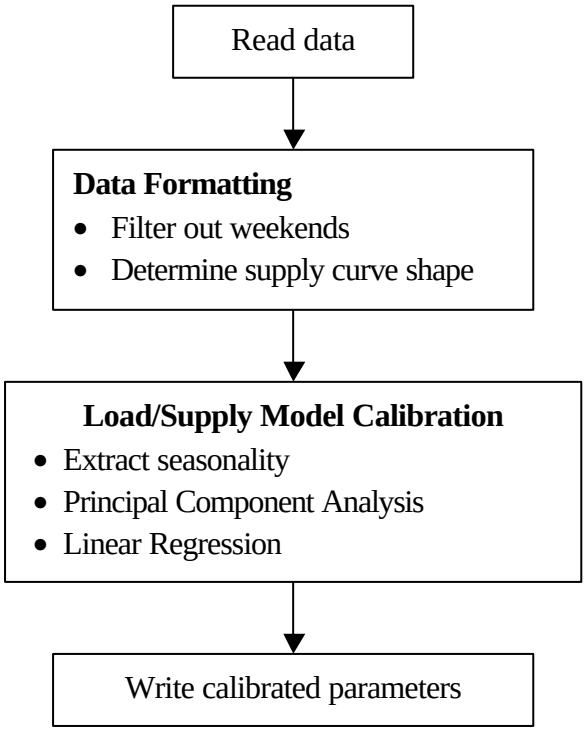

Fig. 12: Flowchart of the UVM calibration

The second part consisted of model calibration to data, prepared in the first part. It was to a great extent uniform in both models, although they were calibrated independently and separately. In this step, seasonality in load and supply models was taken into account, and the parameters of the both models were calculated using Principal Component Analysis and Linear Regression. The general flowchart is presented in Fig. 13.

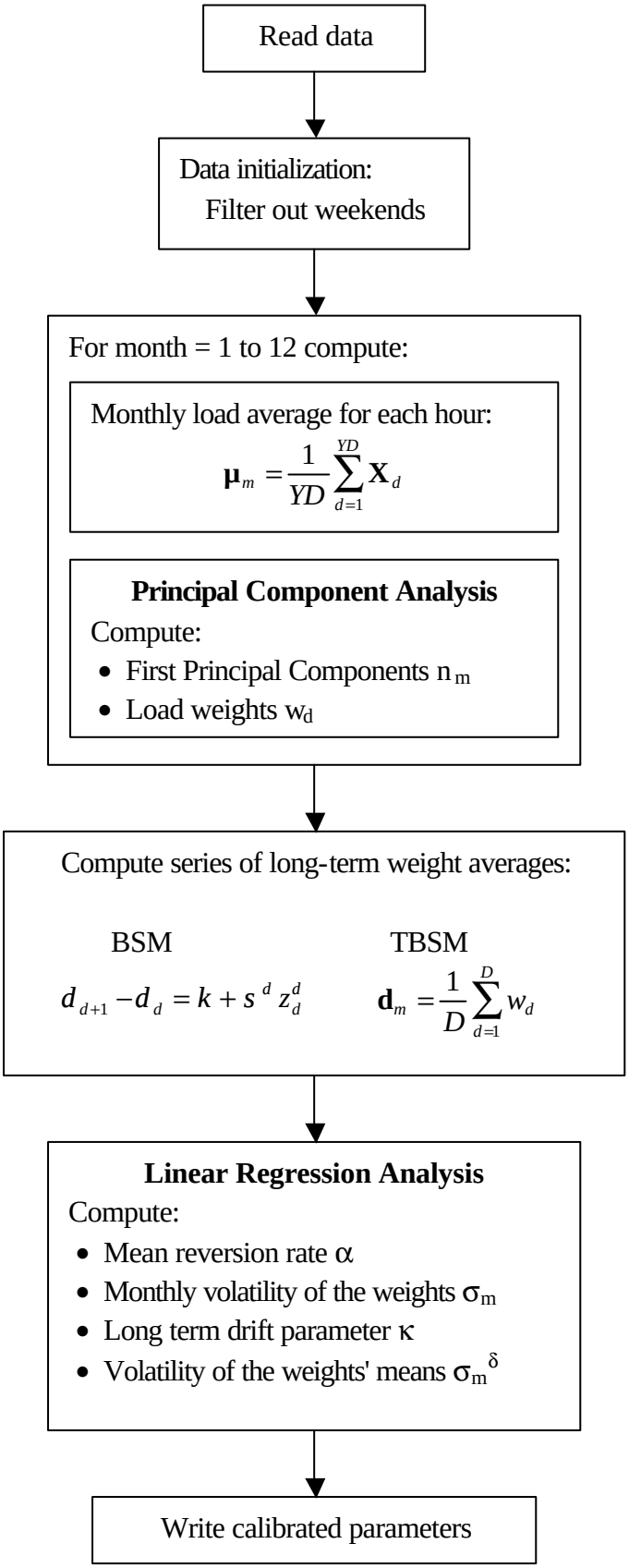

Fig. 13: Flowchart of the BSM calibration

3.1 Calibration and outages

The model of supply process in its original form facilitates description of outages. Although more flexible, it would make calibration harder since it would require actual data on generator outages along with price data. Another possibility would be to assume information on outages from the jumps present in supply bidcurves. Information on outages could then be inferred by measuring the skewness of probability density function from lognormal distribution of the bids.

For the purpose of calibration, the supply process $\mathbf{b}_d$ was in our case modeled in a simpler form without outages. It was therefore possible to postulate the evolution of the daily error for both fundamental processes, $\mathbf{L}_d$ and $\mathbf{b}_d$ in the following fashion:

$$e_{d+1} - e_d = -a\, e_d + s\, z_d$$

By substituting the w_d and δ_d parameters into the equation, we can get evolution of the daily weights:

$$e_{d+1} - e_d = w_{d+1} - d_{d+1} - w_d + d_d = a\left(d_d - w_d\right) + s\, z_d$$

The daily weights therefore emerge as a result of the mean reversion to the daily mean, long term drift of the mean κ, and the two independent stochastic processes with their respective volatilities σ and σ^{δ}.

$$w_{d+1} - w_d = a\left(d_d - w_d\right) + k + s^{d} z_d^{d} + s\, z_d$$
$$w_{d+1} = (1-a) w_d + a\, d_d + k + s\, z_d + s^{d} z_d^{d}$$

3.2 Application of PCA to BSM

3.2.1 Load side

We wanted to calibrate the Bid-based Stochastic Model using the New England load data, of which we had 18 years to our disposal. With a help of a Matlab-based computer program, we have calculated the matrices $\mathbf{v}$ and m. Using the PC Analysis, our goal was to reduce the order of $\mathbf{v}$ from $[12 \times n \times 24]$ to $[12 \times j \times 24]$, where n=24 and j = 1. Load was modeled as

$$\mathbf{L}_d = \boldsymbol{\mu}_m^L + \sum_{i=1}^{j} v_d^{Li}\, \boldsymbol{?}_m^{Li} \qquad (1)$$

The monthly average daily shape of load, $\mathrm{m_m}^{\mathrm{L}}$ for 24 hours and 12 months for New England is shown in Fig. 6 and the principal components in Fig. 14. Time series of daily load weights $\mathbf{w_d}^{\mathrm{L}}$ reverting to the long-term mean $\mathrm{d_m}^{\mathrm{L}}$ for New England for the years 1980-98 can be examined in Fig. 15.

As expected, the load data of a particular hour in single month were highly correlated, so by using only the first PC we were able to account for over 90% of the variance. The variance (in %) explained per month by the first PC is shown in Tab. 1.

Tab. 1: Variance of load explained by the first PC for different months

Month	1	2	3	4	5	6	7	8	9	10	11	12	Avg.
Var (%)	92.78	95.12	94.60	93.37	94.13	96.33	96.75	96.46	93.91	94.27	93.06	93.36	94.51

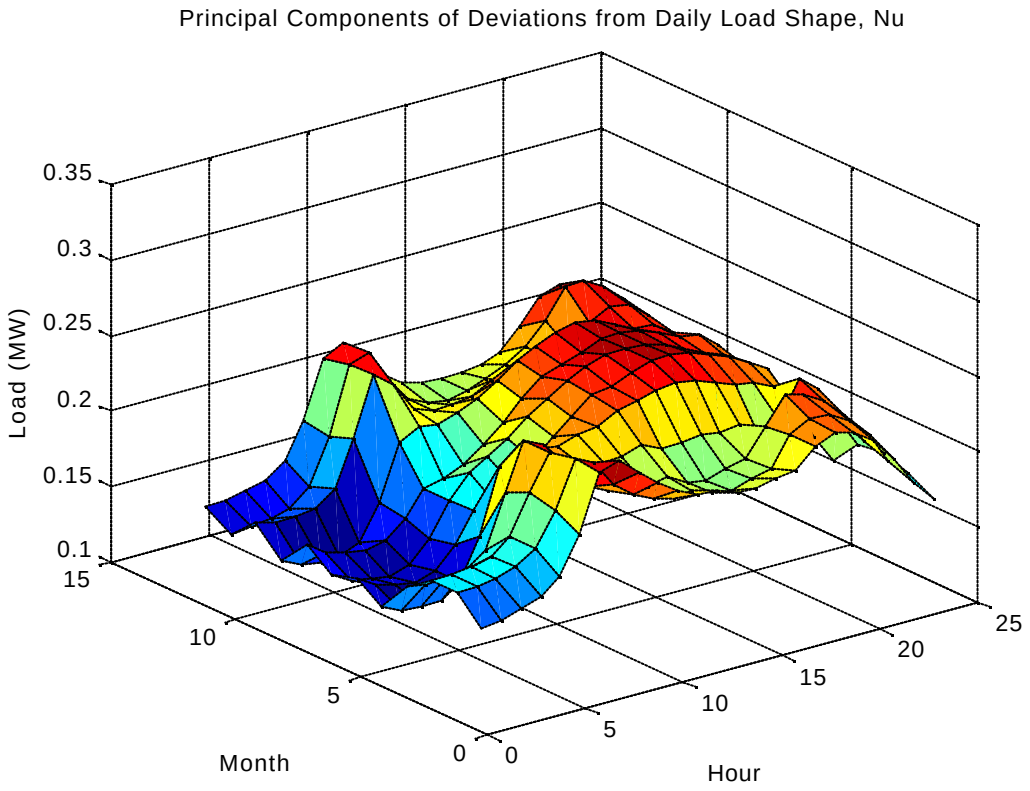

Fig. 14. Principal Components of load $\mathbf{v}_m^L$ for New England

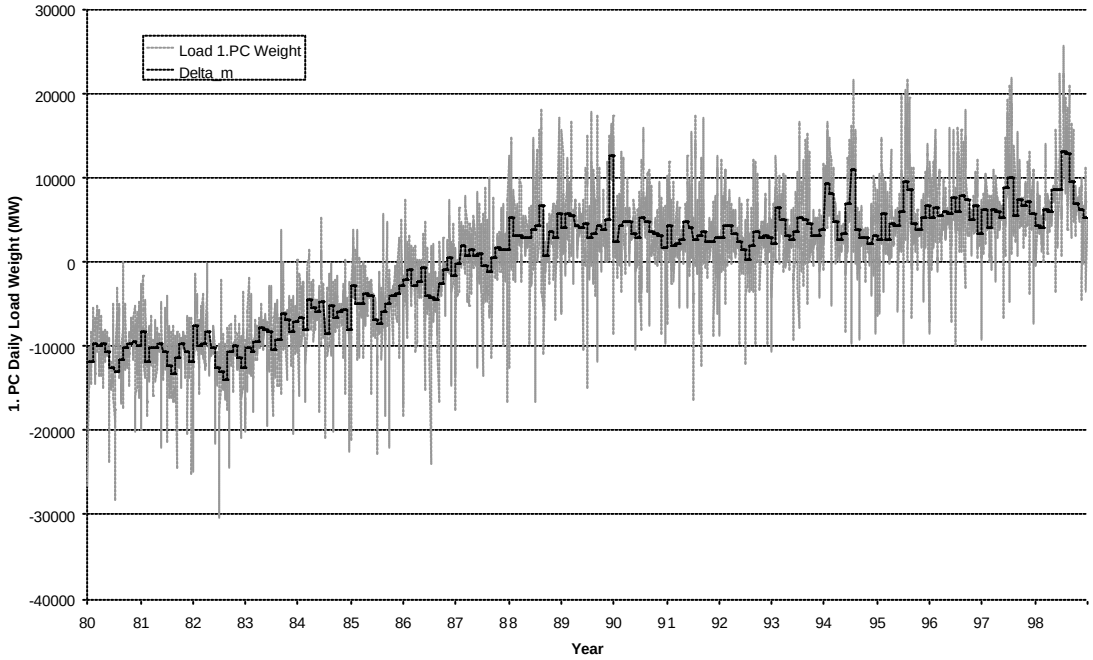

Fig. 15: Load weights of the 1. PC, w_d^L, and monthly average δ_m^L, New England

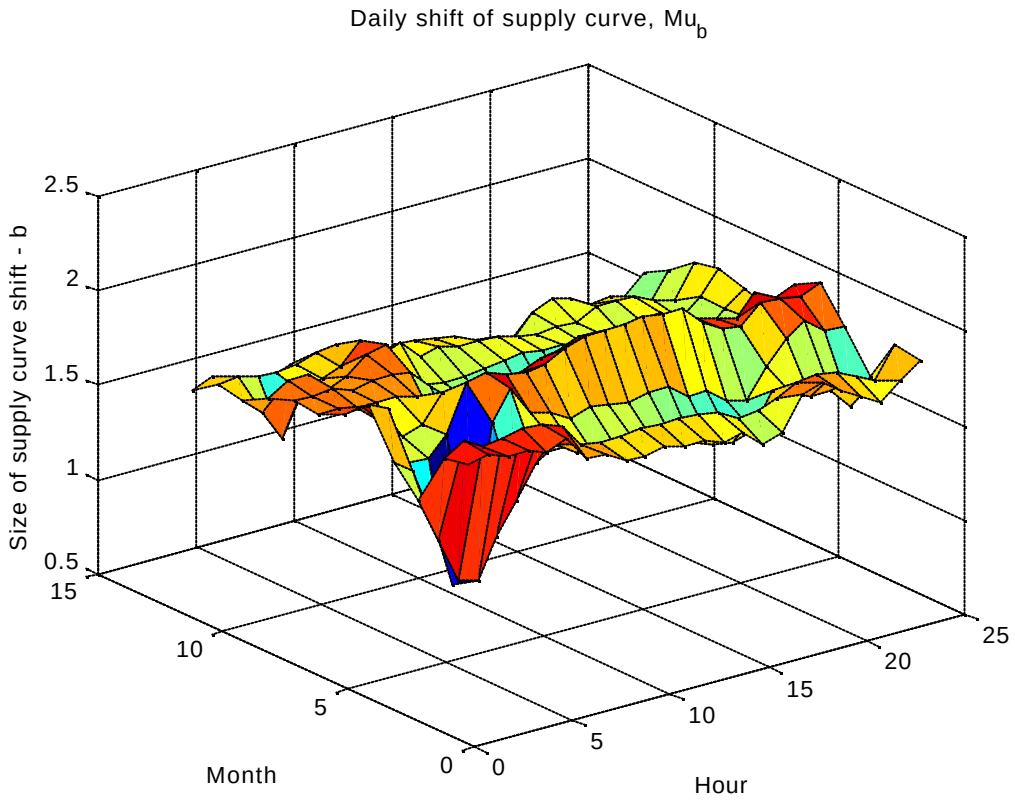

Fig. 16. Average monthly pattern of supply shift μ_m^b, New England

3.2.2 Supply side

At the supply side model, we wanted to model the evolution of the bidcurve shift factor b that appears in the price model equation.

$$\mathbf{P}_d = e^{\alpha \mathbf{L}_d + \mathbf{b}_d} \tag{2}$$

Similarly to our derivations for load, we wanted to apply PC decomposition to obtain the following expression for b

$$\mathbf{b}_d = \boldsymbol{\mu}_m^b + \sum_{i=1}^{j} w_d^{bi} \boldsymbol{?}_m^{bi}, \quad d = 1 \ldots n, \quad m = 1 \ldots 12 \tag{3}$$

The monthly average shape of hourly shift of price curve, μ_m^b for 24 hours and 12 months for New England is shown in Fig. 16. Time series of daily supply curve shift $\mathbf{b}_d$ and μ_m^b for New England for the years 1998-2000 can be examined in Fig. 15. The full set of principal components is shown in Fig. 17, and the interplay between the first two PC-s in Fig. 18.

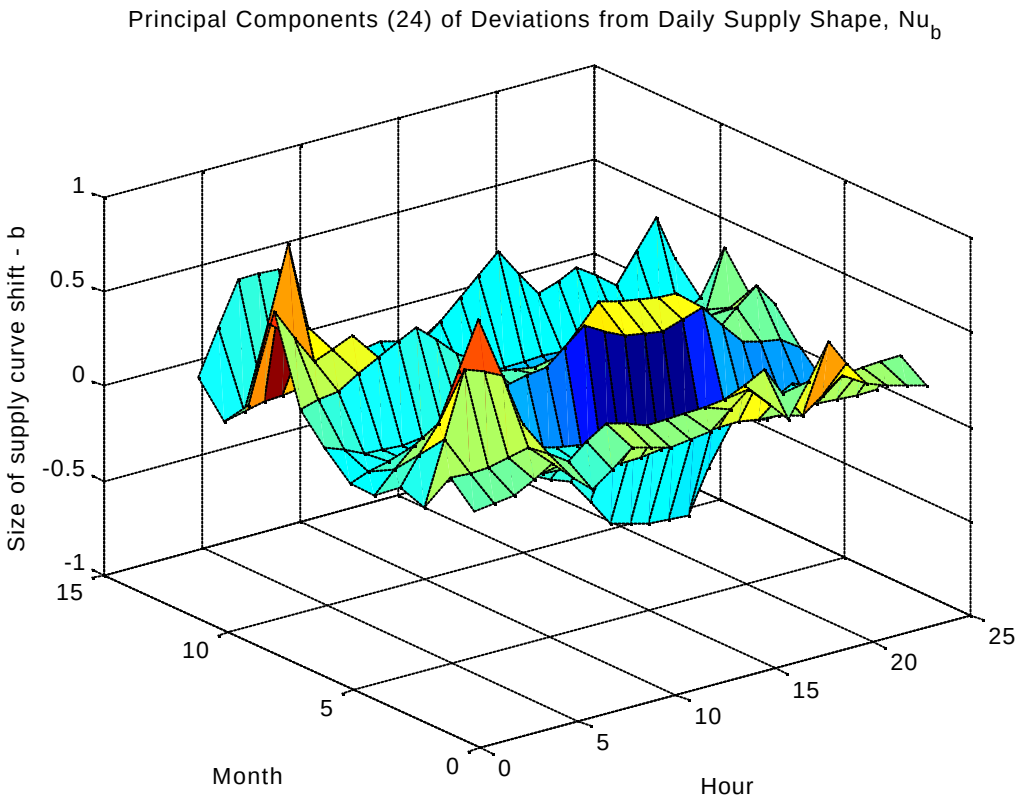

Fig. 17: Principal Components of supply shift $\mathbf{v}_m^b$, New England

To calibrate the supply side of the BSM using the New England supply data, the problem was that we only had 14 months of hourly **b** available. To use a full-size PC Analysis, the number of instances in data (in our case workdays in a month) should be at least equal the number of original variables, in our case the number of hours analyzed. Since on average there are only about 22 workdays in a month, we would require at least two instances of each month, raising the required number of months to 24.

To extend the available amount of data, three approaches were investigated.

1. Duplicating the missing months to obtain 24 months worth of data. Since data are result of two distinct stochastic processes, this would significantly alter data beyond usability, introducing a deterministic pattern.
2. Treating the entire year as composed of 12 equal months, thus introducing a single set of j principal components. As the PC analysis is used to model deviation from the monthly daily load pattern m_m, this approach would have adverse effects on the amount of information retained by the model. The investigated timescales for load pattern m and principal component matrix **v** is outlined in Tab. 2.

Tab. 2: Effects of different modeling timescale

	m	**v**	**Result**
1	yearly	yearly	Some of the effects cancel each other out
2	monthly	yearly	
3	monthly	monthly	OK

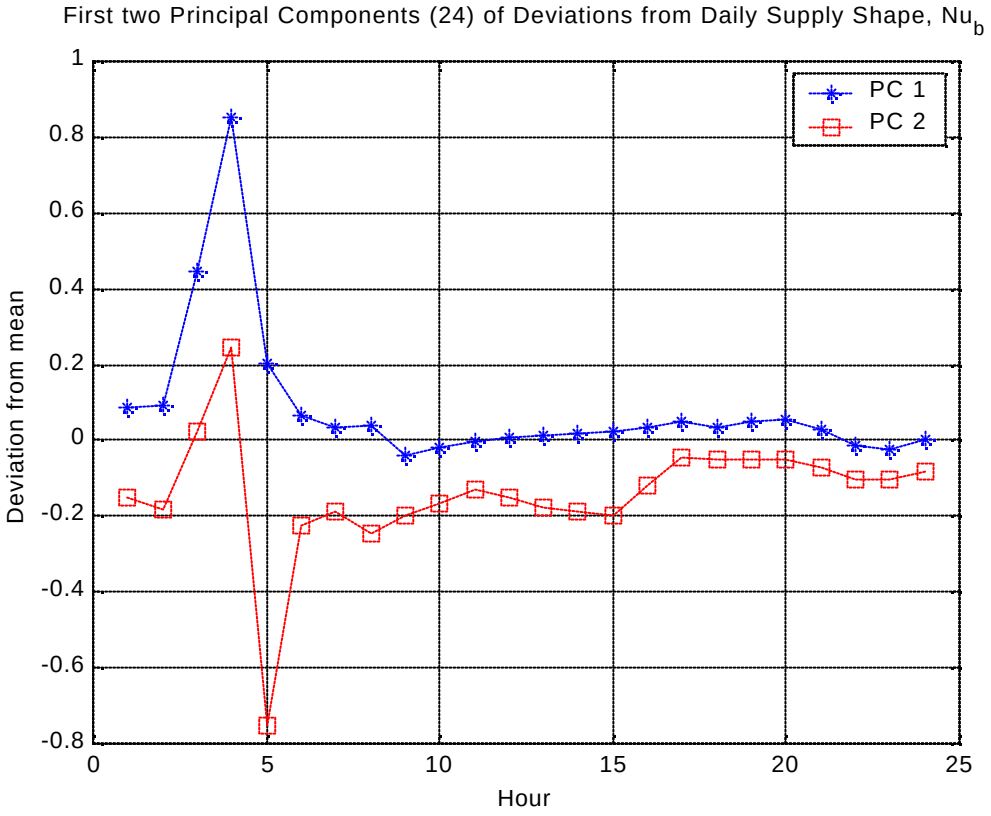

Fig. 18: First two PCs of deviation from m_m^b for 24 hours for December

3. Treating every month separately, but using reduced number of variables to calculate PCs. In our case, only 12 odd hours were used as original variables, reducing the order of PC matrix n to [12 × j × 12]. For the reduced order of the problem, data on single month (at least 20 days) were sufficient. After the matrix **v** was calculated and the number of retained PCs determined, $\mathbf{v}_m$ for the missing 12 even hours were interpolated. The interpolation of values of PCs is acceptable since price shifts between hours always occur continuously, i.e. there are no stochastic jumps between hours. Tab. 3 presents the amount of variance explained by the first five PCs.

Tab. 3 Variance of supply (in %) explained by the first four PCs for different months

Month / **PC**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Avg**
1	50.12	36.08	64.98	63.50	49.18	45.68	64.29	51.57	52.10	86.92	59.47	53.99	52.22
2	19.58	19.90	21.37	11.96	42.25	28.22	14.71	24.98	13.21	8.00	12.13	23.16	18.57
3	12.30	18.66	4.30	9.32	4.09	11.44	9.17	13.41	11.75	1.74	10.65	9.56	9.18
4	6.99	8.72	3.31	8.34	1.32	4.74	5.64	3.41	7.21	1.31	5.04	4.07	4.93

Here, the choice of only one PC is less obvious than in the load data. An average amount of variance explained by each of 12 original variables is 8.33 %, so according to guidelines

(Appendix A) we should have in some months considered using two or even three PCs. However, we have decided to use one PC since on average, the variance explained by it was 52.22%. Considering many assumptions we had made with regards to modeling of supply function, the error made by omitting other PC was not crucial and was therefore acceptable.

3.3 Estimation of the parameters of the BSM

The Bid-based Stochastic Model can be expressed in the state space as:

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \mathbf{Q?}_t$$
$$y_t = \mathbf{C}\mathbf{x}_t + \mathbf{D}\mathbf{u}_t + \mathbf{R}x$$

It can be shown that the model is controllable and observable. The parameters of the BSM could be jointly estimated using standard estimation techniques, such as either Extended Kalman Filter or the Maximum Likelihood Estimation method in conjunction with Iterative Kalman Filter, as outlined in the Appendix B.

The problem of joint estimation of system and noise parameters has been solved in the literature for simpler problems [5], [6], [7], [8]. However, there are significant differences between our problem and others. Some of the approaches were using a simpler two-factor model or assumed risk-neutrality, which does not hold true in electricity markets. Others were describing price directly without relying on underlying fundamental processes of load and supply, so they could use additional data on forward prices. Since in our case forward markets on load or supply shift do not exist, we could not use this kind of additional information.

The standard estimation techniques usually assume known covariance matrices of the stochastic processes in the model, i.e. process noise covariance matrix **Q** and measurement noise covariance matrix **R**. Alternatively, other techniques for estimation of noise covariances require complete knowledge of other system parameters. Since among the unknown BSM parameters there were also the stochastic process variances σ and σ^{δ}, the elements of the matrices **Q** and **R**, the standard estimation techniques failed to converge.

The parameters had to be estimated separately in several consecutive steps, in which the parameters were estimated independently. Since the load and the supply processes are described in a similar way in the model, their parameters α^{L}, κ^{L}, σ^{L}, $\sigma^{L\delta}$ and α^{b}, κ^{b}, σ^{b}, $\sigma^{b\delta}$ can be estimated separately and in the same way.

Estimation of the Bid-based Stochastic Model parameters can therefore be summarized in three successive phases.

1. **Long-term drift of the mean** k is estimated using linear least squares fit. After κ is determined, data is de-trended, i.e. the long-term drift is eliminated.
2. **Calculation of mean reversion factor**: Factor α, determining the mean reversion speed of the weight process can be estimated using linear regression over de-trended data.
3. **Estimation of process volatilities**: Using the estimated α, the remaining parameters of the model in state space form σ and σ^{δ} can be estimated using the adaptive Kalman Filter and the technique for identification of the variance-covariance matrices of the process and measurement noise **Q** and **R**, [3].

3.3.1 Calculation of mean reversion factor in the BSM

The evolution of weights can be calculated in the following way:

$$w_1 = (1-a)w_0 + a\,d_0 + s\;z_0 + k + s^{\,d} z_0^d$$

After assuming the initial values of ($\delta_0 = 0$) and the linear trend already eliminated from the data in the previous step ($\kappa = 0$), the following sequence of equations unfolds.

$$\begin{aligned} w_2 &= (1-a)w_1 + a\,d_1 + s\;z_1 + s^{\,d} z_1^d \\ &= (1-a)w_1 + a\,d_0 + s\;z_1 + a\,s^{\,d} z_0^d + s^{\,d} z_1^d \\ &= (1-a)w_1 + s\;z_1 + a\,s^{\,d} z_0^d + s^{\,d} z_1^d \\ w_{k+1} &= (1-a)w_k + s\;z_k + s^{\,d}\left(a\sum_{j=0}^{k-1} z_j^d + z_k^d\right) \end{aligned}$$

The last equation could be rewritten as

$$\begin{aligned} w_{k+1} &= (1-a)w_k + A_k \\ A_k &= s\;z_k + s^{\,d}\left(a\sum_{j=0}^{k-1} z_j^d + z_k^d\right) \end{aligned}$$

Since the part of the equation, denoted as A_k, is influenced solely by zero-mean processes z and z^δ, it is a zero-mean process.

$$A_k \approx N(0,1)$$

It is then possible to estimate α using Linear Regression over the time series vectors of weights $\mathbf{w}_d$. The vector $\mathbf{w}_{d+1}$ is shifted in time for a day ahead compared to $\mathbf{w}_d$. The operator $\Diamond$ denotes least-squares fit of the two vectors [13].

$$a = 1 - (\mathbf{w}_{d+1} \Diamond \mathbf{w}_d)$$

Although A_k is a zero-mean process, its variance is a cumulative sum of the variances of the underlying stochastic processes. The variance therefore grows rapidly with the number of samples considered in the regression procedure. Using the full set of samples (some 4500) rendered very unreliable estimates of α. It turned out that a more reliable estimate could be obtained using between 1500 and 50 samples.

The development of estimates of α^L for the load process is shown in Fig. 19 and for supply process, α^b, in Fig. 20. As the number of samples used in estimation decreases, the value of α converges to its actual value. The two curves show evolution of the estimate when using data from the beginning or from the end of the series.

A summary of the estimated parameters of the Bid-based Stochastic Model for both load and supply processes are shown in Tab. 4.

Tab. 4: Estimated parameters of the BSM

	Load process	**Supply process**
α	0.3	0.75
κ	4	-1.7e-3
a	1.13484e-4	

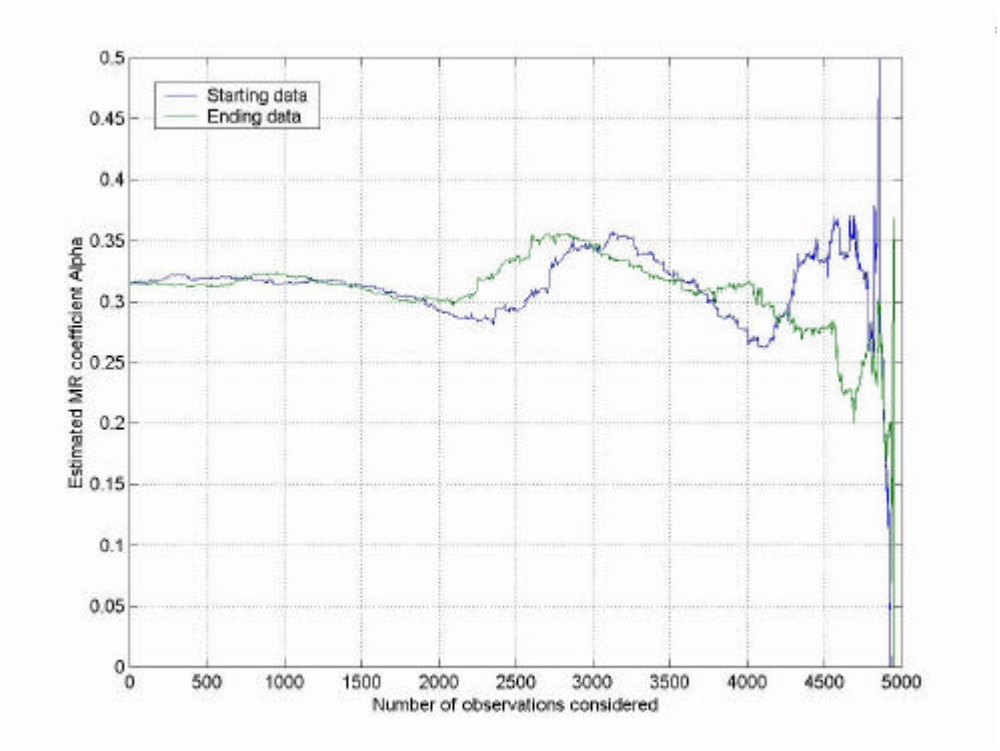

Fig. 19: Estimation of α^L as a function of number of samples considered

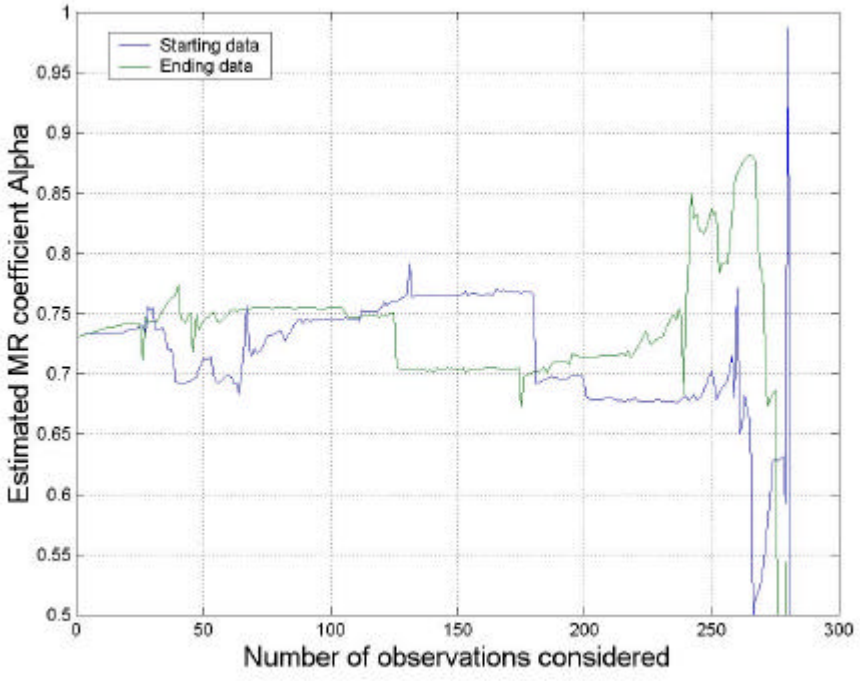

Fig. 20: Estimation of α^b as a function of number of samples considered

3.4 The Time-scale Separated Bid-based Stochastic Model

3.4.1 Introduction

The Time-scale Separated Bid-based Stochastic Model (TBSM) evolves from the daily Bid-based Stochastic Model as we introduce the assumption of time-scale separation between the fast, short-term and slow, long-term processes. The assumption significantly reduces complexity of the model and therefore the computational burden necessary in potential model applications.

The TBSM postulates time scale separated development of the daily weights in load and supply curve shifts. In this model, the daily weights w_d revert to the monthly mean δ_m, yet both states of the model evolve on a different time scales. We define a daily error of the weight e_d as the difference between daily principal component weights w_d and their long-term mean δ_m.

$$e_d = w_d - d_m$$

where

$$e_{d+1} - e_d = -a\, e_d + s_m\, z_d$$
$$d_{m+1} - d_m = k + s^d z_m^d$$

Typically, m >> d, with m denoting monthly and d denoting daily values. As the properties of both models differ substantially, the values of their parameters differ significantly as well, Tab. 4.

The assumption on time-scale separation was introduced to facilitate simpler calibration of the model. As a trade-off, several issues with the Time-scale Separated Bid-based Stochastic Model arise.

1. Time scale separation is not genuine: the time constant of the monthly process T_m is not sufficiently larger than the one of the daily process T_d to warrant the separation assumption.
2. The weight process does not fully revert to the mean within one month, so it is misleading to compute the long-term mean as an average of the weights over a month.

   $$d_m = \frac{1}{T_m} \sum_{t=T_{m-1}}^{T_m} w_{dt}$$

   Because of this, the long-term mean in the TBSM is larger than the actual mean, $\delta^{TBSM} > \delta^{BSM}$. Since the weights in the TBSM revert to a larger mean, the speed of mean reversion $\alpha^{TBSM} > \alpha^{BSM}$ is necessarily greater than the actual one.
3. The jumps in the δ^{TBSM} are critical and unacceptable for Unit Commitment under physical constraints.

3.4.2 Calibration of TBSM

Although load and supply processes describe different physical phenomena, the BSM postulates a similar structure for both of them. Calibration of both processes, although performed separately, therefore follows the same pattern.

From the time series of daily principal component weights w_d, the parameters of the model, which govern mean reversion of the w_d, α, and long-term drift of the mean δ_m, κ, needed to be determined. The following steps describe the procedure.

1. We construct a [D×1] time series vector of principal components weights w_d, **w**.

$$\mathbf{w} = [w_d], \quad d = 1..D$$

2. For every year and for every month within the current year, a mean of w_d, δ_m, was calculated. D is the total number of days in the data, while D_m is their number in the current month.

$$\begin{aligned} d_m &= \frac{1}{D_m}\sum_{d=1}^{Dm} w_d, \quad m = 1..12 \\ D &= \sum_{m=1}^{12} D_m \end{aligned} \tag{4}$$

The [12×1] vector of monthly means $\mathbf{d^*}$ could then be defined as

$$\mathbf{d^*} = [d_m], \quad m = 1..12$$

We could also define a [D×1] vector d, defined as a daily time series of δ_m.

$$\mathbf{d} = [d_{md}], \quad d = 1..D_m, \quad m = 1..12$$

3. The mean reversion of the daily weights w_d to the pertaining monthly mean δ_m is described as

$$w_{d+1} - w_d = a\ (d_m - w_d) + s_m z_d \tag{5}$$

where the change in weights is determined by the mean reversion part and the stochastic component, $\sigma_m z_d$. The stochastic process z_d is normally distributed with a zero mean and standard deviation of one.

$$z_d \approx N(0,1) \tag{6}$$

The mean of the stochastic process is zero and is not affected by the process volatility measure, σ_m. Coefficient α could therefore be determined using linear regression as the slope b of the "best fitting" regression line to satisfy the least-squares criterion [13].

$$y' = a + bx$$

$$b = \frac{\sum_{d=1}^{D}(x_d - \bar{x})(y_d - \bar{y})}{\sum_{d=1}^{D}(x_d - \bar{x})^2}$$

The regression is performed over the time series vectors of weights **w** and monthly means d. A shift of the vector $\mathbf{w}_d$ for a day ahead is denoted as $\mathbf{w}_{d+1}$.

$$\begin{aligned} \mathbf{x} &= [x_d] = (\mathbf{w}_{d+1} - \mathbf{w}_d) \\ \mathbf{y} &= [y_d] = (\mathbf{d}_m - \mathbf{w}_d) \\ a &= b \end{aligned} \tag{7}$$

4. Using α, a vector of estimated weights $\mathbf{w}'_{d+1}$ was obtained:

$$\mathbf{w}'_{d+1} = a\mathbf{d}_m + (1-a)\mathbf{w}_d \tag{8}$$

5. The difference between the estimated $\mathbf{w}'_{d+1}$ and $\mathbf{w}_{d+1}$ was the contribution of the stochastic component of the process, $\sigma_m z_d$. It was therefore possible to calculate the monthly volatility measure σ_m of the process by subtracting the estimated values of $\mathbf{w}'_{d+1}$ from the actual values $\mathbf{w}_{d+1}$ and calculating standard deviation of the parts of the time series vectors, belonging to a particular month:

$$\begin{aligned} s_m &= StDev\left(\mathbf{w}^m_{d+1} - \mathbf{w}'^m_{d+1}\right), \quad m = 1..12 \\ StDev(x) &= \sqrt{\frac{1}{D-1}\sum_{i=1}^{D}\left(x_i - \frac{1}{D}\sum_{j=1}^{D}x_j\right)^2} \end{aligned} \tag{9}$$

6. The parameters of the weight mean process δ_m have been determined using linear regression. The drift parameter κ has been calculated as a mean difference of time shifted time series of the monthly means $\mathbf{d}$.

$$\begin{aligned} \mathbf{?} &= \mathbf{d}_{k+1} - \mathbf{d}_k \\ \mathbf{?} &= [k_d], \quad \mathrm{d} = 1..\mathrm{D} \\ k &= ((\mathbf{d}_{k+1} - \mathbf{d}_k) \Diamond \mathbf{I}) \end{aligned}$$

7. With the help of κ, a [D×1] vector of estimated weight means $\mathbf{d}'_{d+1}$ was calculated:

$$\mathbf{d}'_{d+1} = k + \mathbf{d}_d \tag{10}$$

8. Similar to monthly volatility in $\mathbf{w}$, the volatility measure σ^δ of the mean process was calculated by subtracting the estimated values of $\mathbf{d}'_{d+1}$ from the actual values $\mathbf{d}_{d+1}$ and computing the standard deviation:

$$s^d = StDev\left(\mathbf{d}_{d+1} - \mathbf{d}'_{d+1}\right) \tag{11}$$

After applying the algorithm to both processes, the 1st principal component's weights of load w_d^L and of supply curve shift w_d^b, a set of parameters of BSM was obtained, presented in Tab. 5, Tab. 6 and Tab. 7.

Tab. 5: Calibrated parameters of the Load in TBSM

	α^L	κ^L	$\sigma_m{}^{\delta L}$
Load	0.0204105	75.4766	2023.60

Tab. 6: Monthly calibrated parameters of the Supply in TBSM

	a	α^b	κ^b	$\sigma^{\delta b}$
Supply	1.13484e-4	0.0318440	-0.0536667	0.413067

Tab. 7: Monthly calibrated parameters of the TBSM

Month	Load volatility measure $\sigma_m{}^L$	Supply volatility measure $\sigma_m{}^b$
1	876.755	0.200464
2	674.219	0.081697
3	477.651	0.350756
4	468.729	0.375337
5	831.078	0.384563
6	679.871	0.324813
7	1092.886	0.364438
8	748.013	0.104237
9	1030.650	0.177193
10	371.965	0.409969
11	779.300	0.134920
12	862.887	0.225309
Average	684.154	0.241054

A comparison of the calibrated parameters between the two versions of the model, BSM and TBSM, is shown in Tab. 8. The variances $\sigma_m{}^L$ and $\sigma_m{}^b$ in TBSM are 12-month averages.

Tab. 8: Comparison of estimated parameters in the TBSM and BSM

	Load process		**Supply process**	
	TBSM	**BSM**	**TBSM**	**BSM**
a	0.0204	0.3	0.0318	0.75
k	75	4	-0.0537	-1.7e-3
s	685		0.2411	
s^d	2023		0.4131	
a	1.13484e-4			

An important difference between the TBSM and BSM is also the speed of reversion to mean, α. In case of BSM load process, α^L is about ten times larger than in TBSM, whereas the factor is about 20 times bigger in case of supply process α^b. In the BSM, the long-term mean process δ evolves daily. Although the weight process $\mathbf{w}_d$ responds to stochastic influences in load and supply processes, it reverts to the long-term mean faster than in the case of TBSM where the mean evolves monthly.

4 Simulations

The BSM postulates the market-clearing price as an exponential of the two fundamental processes, load and supply.

$$P_k = e^{aL_k + b_k}$$

Both processes - supply and load - can be described in a similar way as a [24×1] vector **X** using generic formulation.

$$\mathbf{X}_d = \boldsymbol{\mu}_m + \sum_{i=1}^{j} w_d^i \, ?_m^i$$

$$w_{d+1} - w_d = a\left(d_d - w_d\right) + s_m z_d$$

$$d_{d+1} - d_d = k + s^d z_d^d$$

By calibrating the model to historical data of load and supply data, we obtain the values of parameters that determine the evolution of the load and supply processes on:

- **the monthly timescale**: the [24×12] matrix of average monthly 24-hour daily profile m_m and
- **within a day on the hourly timescale**: the [24×12] matrix of monthly principal components $\mathbf{v}_m$.

These parameters are independent of the type of model we use for calculation of daily weights.

At the same time, we obtain the parameters that govern the evolution of the daily weights w_d; mean reversion speed α, long term drift κ, a [24×1] vector of daily weight volatilities s_m and long-term mean δ_d volatility σ^{δ}.

Using simulation it is possible to investigate the properties of the two fundamental processes - load and supply, which drive the price in both models. The simulation also enables us to illustrate their influence on the price.

For the purpose of simulation and demonstration of the properties of the model only, the BSM volatility measures σ and σ^{δ} were approximated using the known parameters of the TBSM. Since these values don't represent true estimates, they were annotated as σ' and $\sigma^{\delta}{}'$.

- The daily weight process w_d evolves on the same timescale in both TBSM and BSM models, so their volatility measures should be roughly the same,

$$\sigma_{TBSM} = \sigma'_{BSM}.$$

- The long-term mean on the other hand develops much faster, and its volatility should therefore be much smaller, divided by a square root of the time constant factor. Assuming that there are about 25 working days in a month, the daily volatility should be about 5-times smaller,

$$s'_{DVM} = \frac{s_{TSVM}}{\sqrt{T}}, \quad T \cong 25 \text{ days}$$

The simulations were performed using either the Bid-based Stochastic Model or Time-scale Separated Volatility Model, calibrated to the short-term market-clearing price. The simulations investigated the impact of different parameters for both load and supply processes on the output of the model. The following properties were investigated:

1. Evolution of the average daily short-term market price (spot price) of electricity, as influenced by the average monthly 24-hour profiles, $\mathbb{m}_m^L$ and $\mathbb{m}_m^b$, and monthly principal components $\mathbf{v}_m^L$ and $\mathbf{v}_m^b$.
2. Expected value and Standard Deviation of the daily weights $\mathbf{w}^L$ and $\mathbf{w}^b$, driven by the daily process parameters, α, κ, s_m and σ^δ.
3. Development of daily averaged hourly price.

The results are briefly discussed and shown in the following sections.

The parameters used in simulation are given in Tab. 9. In addition to the calibrated ones, the parameters σ' and $\sigma^{\delta,}$ were approximated to demonstrate properties of the BSM and are presented in the shaded cells of the table.

Tab. 9: Parameters of the BSM, used in simulations

	Load process	**Supply process**
a	0.3	0.75
k	4	-1.7e-3
s'	685	0.2411
$s^{d,}$	400	0.825
a	1.13484e-4	

4.1.1 Deterministic price and monthly parameters

The monthly average spot price in the actual data to which the model was calibrated is shown in Fig. 21. There are relatively big differences among certain months, describing a year with unusually high summer prices. The summer of 1999 was very hot and the prices were higher than the historical levels. Since only a limited amount (14 months) of price data was available, the influence of a single month in calibration of supply process was stronger than in load process calibration, where almost 20 years of data was available and the influence of excessive months are less prominent.

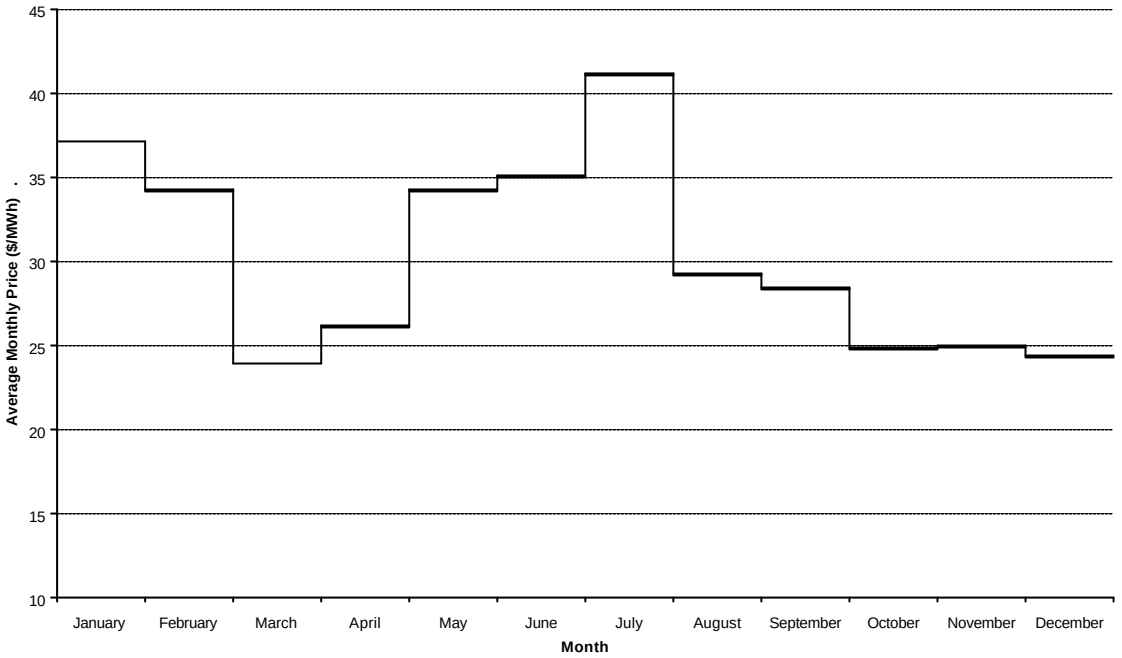

Fig. 21 Monthly average of the spot price, New England, May '99-June '00

In Fig. 22, simulated evolution of the average daily short-term spot price of electricity is shown, as influenced by the average monthly 24-hour profiles, $\mathbf{m}_m^L$ and $\mathbf{m}_m^b$, and monthly principal components $\mathbf{v}_m^L$ and $\mathbf{v}_m^b$. The price as generated by the model exhibits similar properties as the actual average monthly price in Fig. 21. The main difference could be observed during the summer months, where the influence of load process dampens the excessive shift in supply curve toward higher prices, as dictated by supply process.

4.1.2 Daily weight process properties

The stochastic properties of the model on the other hand can be illustrated without interference of monthly mean values by examining the daily weight processes $\mathbf{w}^L$ and $\mathbf{w}^b$. They are driven by four stochastic processes z_k^L, $z_k^{L\delta}$, z_k^b and $z_k^{b\delta}$ and governed by the daily process parameters, α, κ, s_m and σ^δ. Interplay of the short-term $\mathbf{w}$ process variances, s_m^L and s_m^b, and long-term d process variances s^L and s^b in the Bid-based Stochastic Model is schematically shown in Fig. 23. The short-term variances of the mean reverting process, which are bounded, dominate in the short run. As time progresses, the dominance of the long term variances prevails.

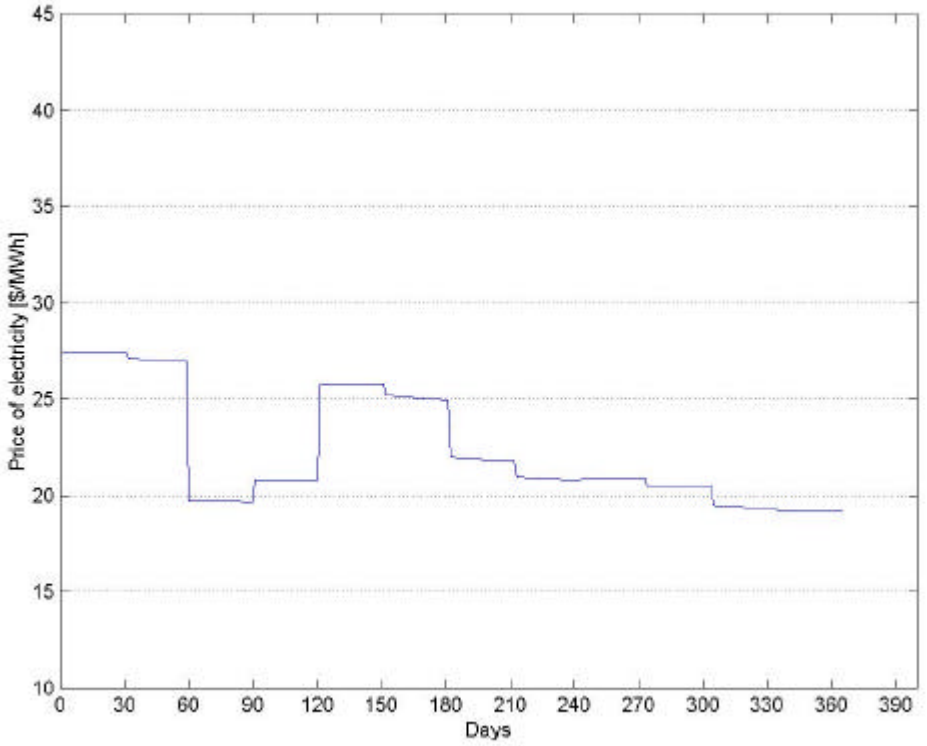

Fig. 22: Simulated average monthly price

Similar conclusions can be drawn from Fig. 24, where the evolution of the mean value of $\mathbf{w}^L$ and its volatility boundaries are shown. The standard deviation of the process is not uniform over the months, what is the consequence of interplay between two volatility measures, s_m^L and $\sigma^{L\delta}$. At the same time, the volatilities of electricity price differ from one month to another. During the periods of peak load, prices tend to be much more volatile than in spring or fall, which is reflected in the model output.

The mean grows steadily according to the long-term growth parameter κ^L. The weights were simulated for a two-year period with a 10.000 simulation runs.

The mean of the supply process weight, $\mathbf{w}^b$, and its volatility boundaries are shown in Fig. 25. The mean slowly drifts downwards, and the monthly shapes in standard deviation are more pronounced than in load process in Fig. 24.

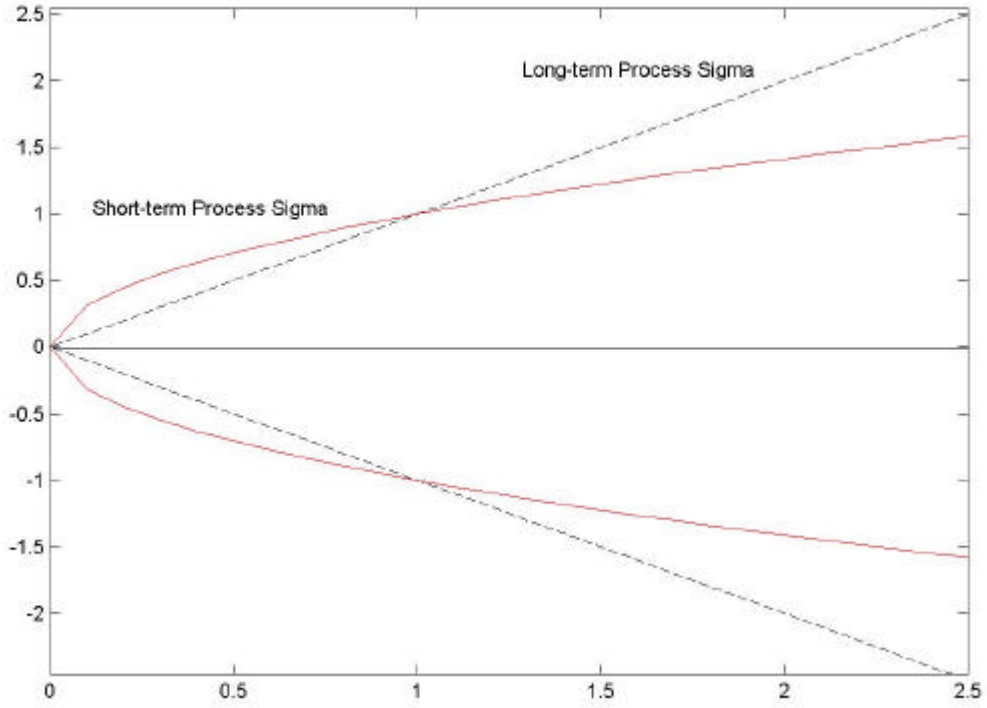

Fig. 23: Volatilities of short- and long-term processes, σ and σ^{δ}

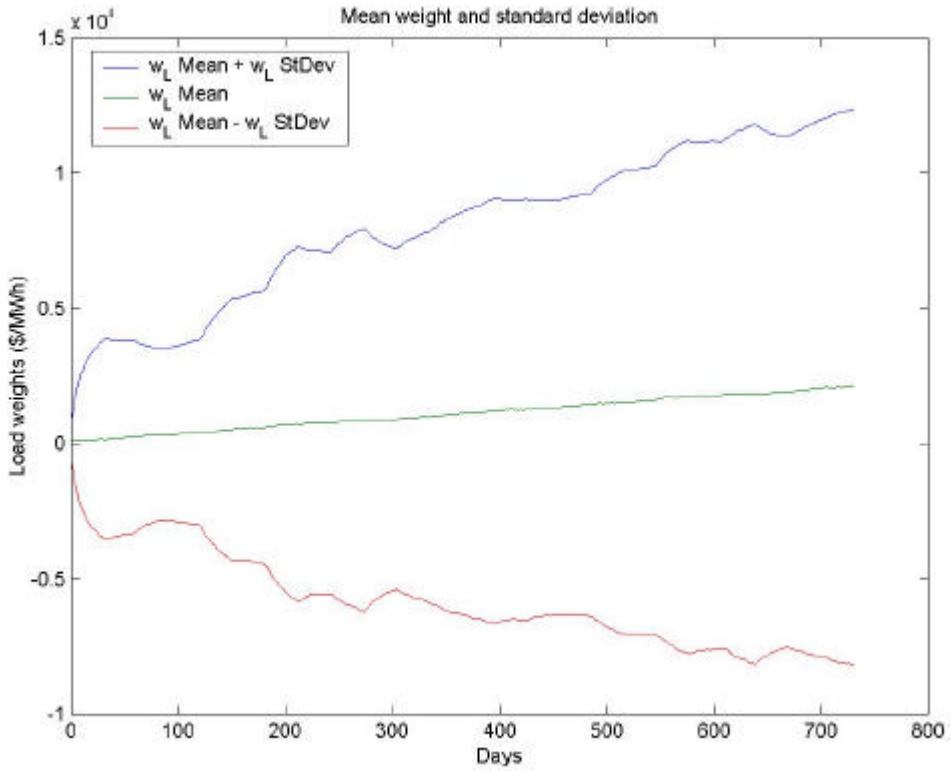

Fig. 24: Daily weights w_m^L : mean value and standard deviation

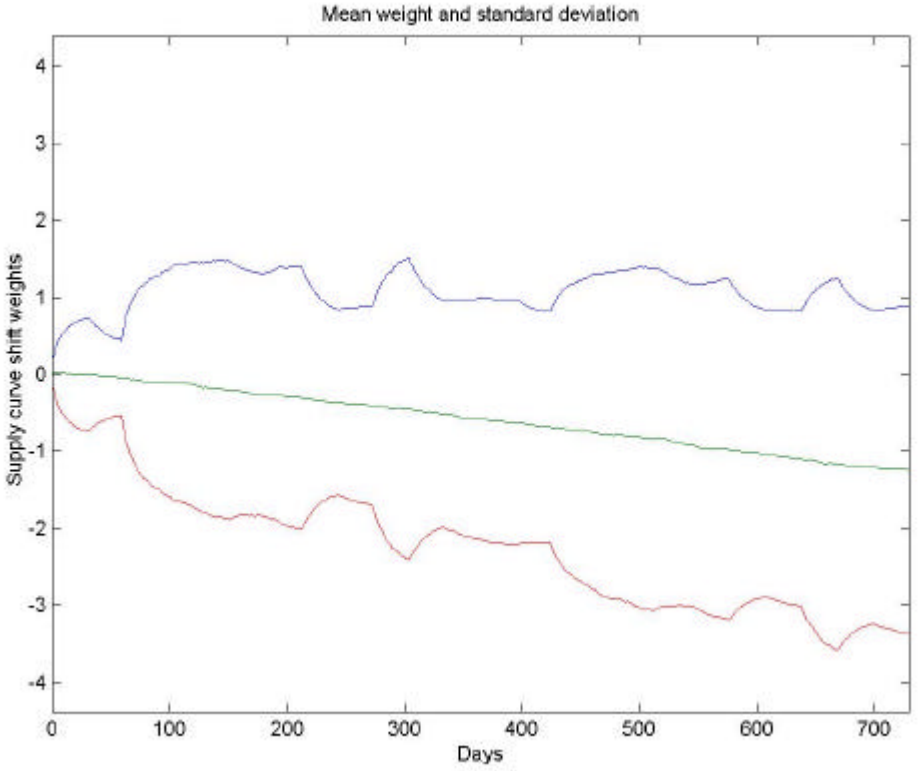

Fig. 25: Daily weights w_m^b : mean value and standard deviation

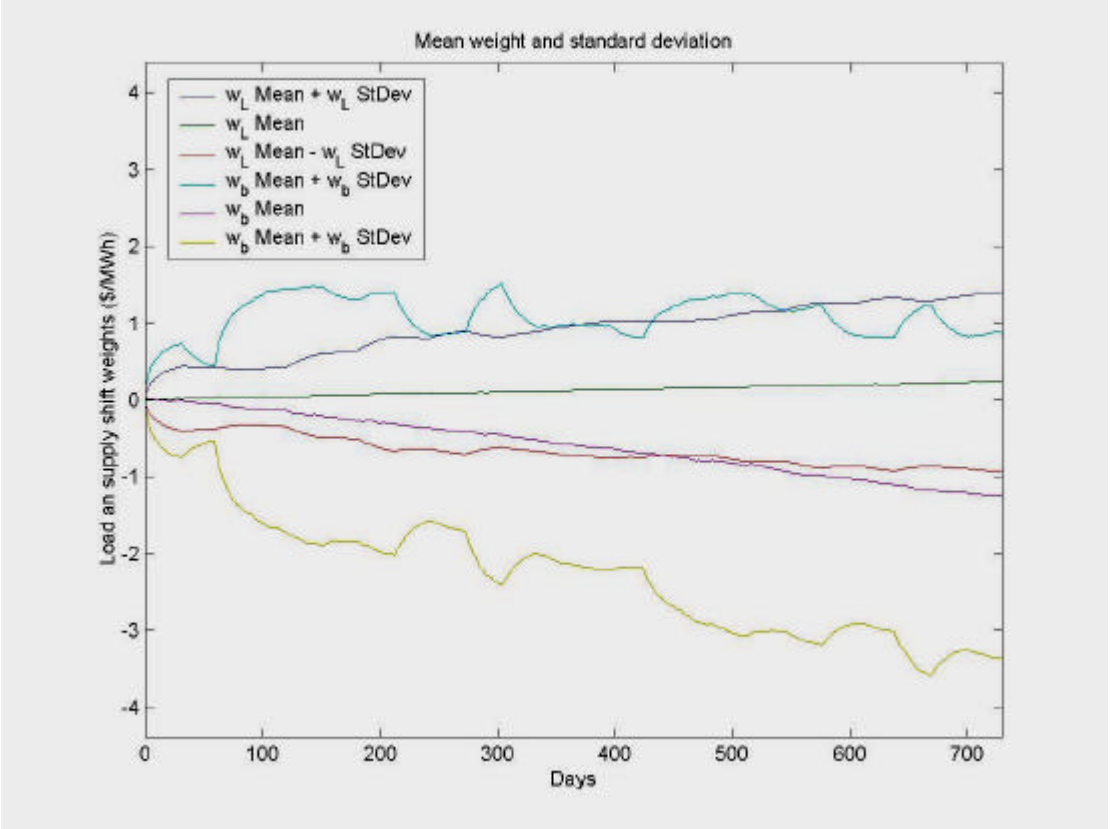

Fig. 26: Daily weights $a*w_m^L$ and w_m^b : mean value and standard deviation

The means and standard deviation boundaries for both processes are shown together in Fig. 26. Here, $\mathbf{w}^L$ is scaled with the factor of the exponential shape a. The volatility boundaries of the supply process $\mathbf{w}^b$ are much broader indicating the dominant source of uncertainty in the forecasted price of electricity is the volatility of the supply process.

4.1.3 Daily price using BSM

Using the BSM it is possible to generate hourly spot price $\mathbf{S}_T$ and its volatility. Because the intra-day dynamics that can be found both in hourly development of load and hourly clearing of market in supply, it is important to have the model that is able to capture the hourly price dynamics. On the other hand, it is sometimes also necessary to neglect the hourly dynamics and deal with daily prices, as it is the case in certain applications such as forward contracts.

In the Stochastic Model the price evolves as a sequence of daily, 24-hour vectors of prices.

$$\mathbf{P}_d = e^{a\mathbf{L}_d + \mathbf{b}_d}$$

$$\mathbf{P}_d = [P_{dh}], \quad h = 1..24$$

The daily price would therefore be calculated as a daily average of the vector $\mathbf{P}_d$.

$$\overline{P}_d = \frac{1}{24}\sum_{h=1}^{24} P_{dh}$$

In the simulation, we have examined the development of the average price $\overline{P}_d$ during the course of one year. The mean and the standard deviation of prices were calculated in 10.000 simulation runs.

Fig. 27 displays the mean value of electricity price and the volatility measure (in our case standard deviation) boundaries. Both prices and standard deviations exhibit strong monthly traits.

The standard deviations, when presented alone in Fig. 28 show corresponding monthly diversity but generally agree with each other and with the observations on volatility in daily weights.

4.1.4 Daily price using TBSM

The Time-Scale Separated Bid-based Stochastic Model properties have been investigated in the same way as the properties of the Bid-based Stochastic Model. Using the parameters from Tab. 4, the average daily price $\overline{P}_d$ was simulated for the period of one year with 10.000 runs.

The mean value of $\overline{P}_d$ is shown in Fig. 29 together with the standard deviation as volatility boundaries. The overall shape still reacts noticeably to changes in months, while the overall impression is that the volatility is somewhat larger than in BSM. The same conclusion could be drawn from analysis of standard deviation in Fig. 30.

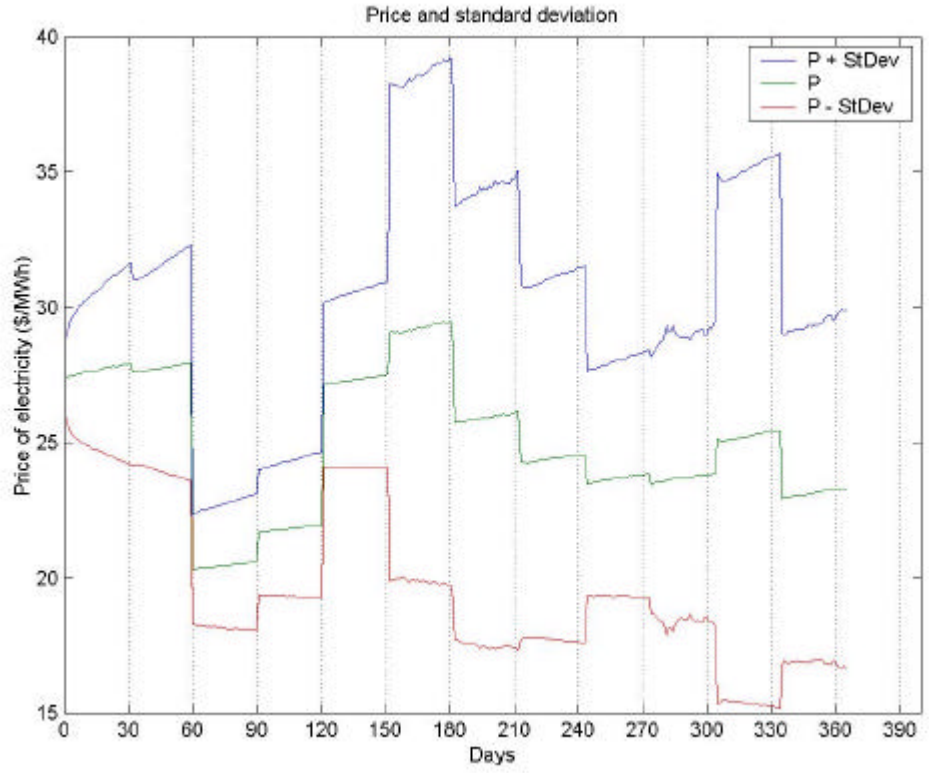

Fig. 27: Average price $\overline{P}_d$ mean and standard deviation

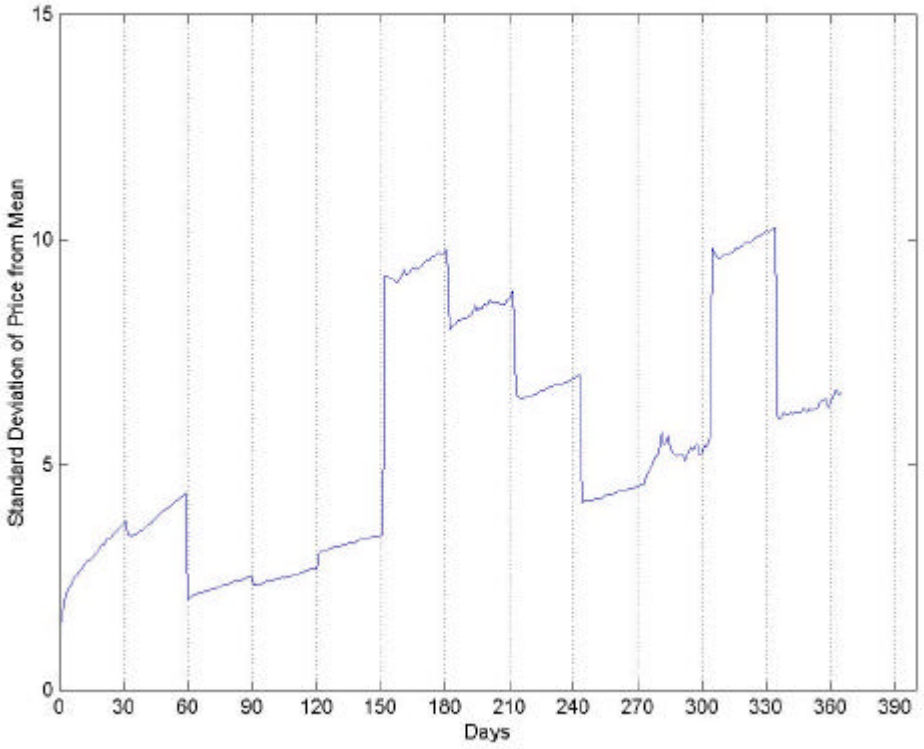

Fig. 28: Standard deviation of the average price $\overline{P}_d$

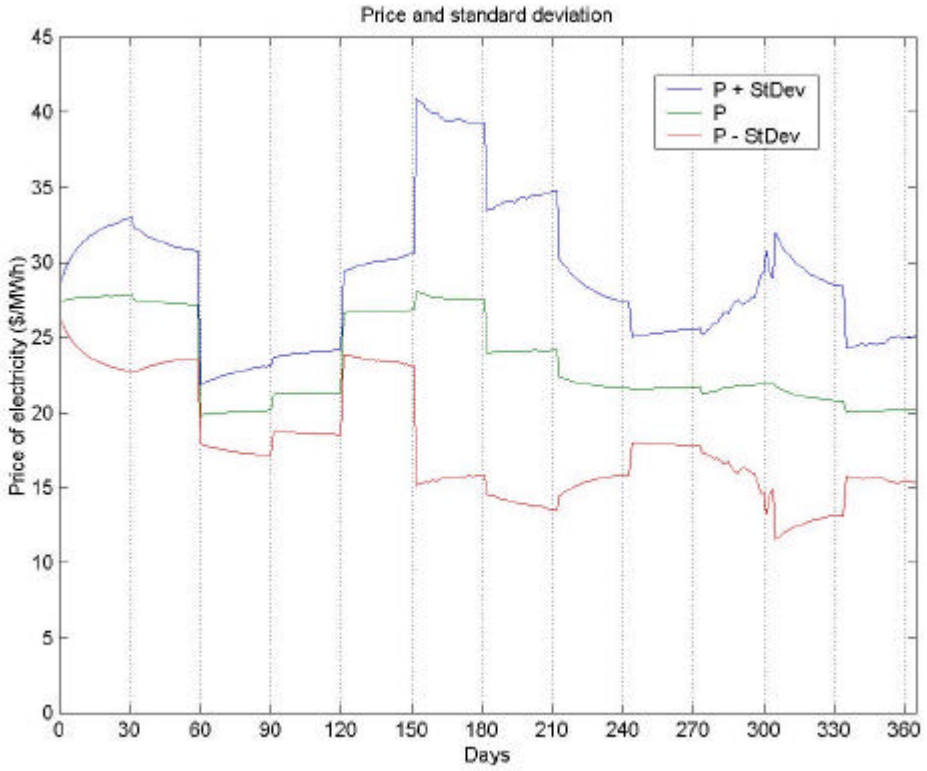

Fig. 29: Average price $\overline{P}_d$: mean and standard deviation

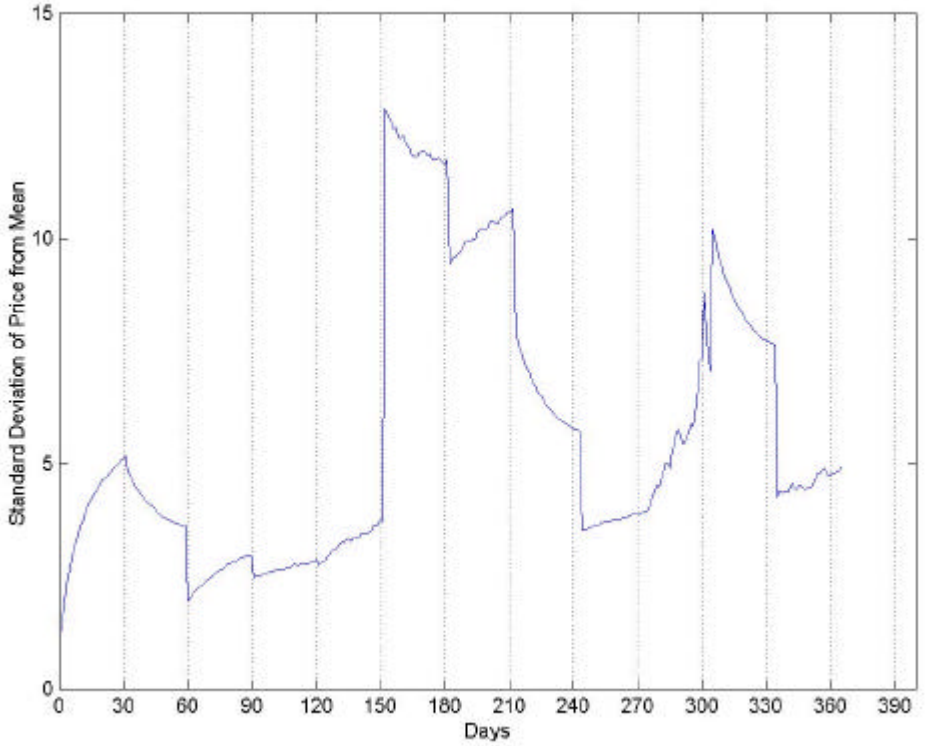

Fig. 30: Standard deviation of the average price $\overline{P}_d$

5 Applications

5.1 A multi-market model

5.1.1 Transmission line flow

We now extend the Bid-based Stochastic Model to simulate the behavior of prices in a multi-market scenario [17]. Specifically we consider the case of two spot markets connected by a tie-line of fixed maximum capacity F_d^{max}. The goal is to model the joint evolution of loads and electricity prices in the adjacent markets, and to show how the bid-based model can be used to estimate the value of a transmission right between the markets.

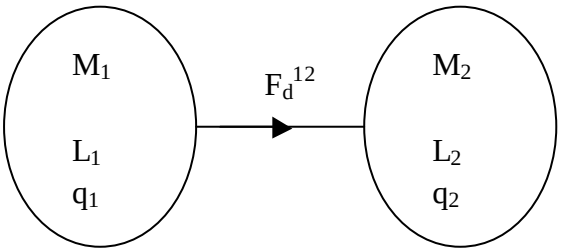

Fig. 31: Two markets, connected by transmission line

Scheduled transmission flows occur when there is cross-bidding between markets, that is, loads or suppliers in one market decide they are better off purchasing their power in the neighboring spot market. A positive flow from market i to market j can be caused by two types of actions:

1. Suppliers in market i decide to bid their power into market j. This causes b^j to increase and b^j to decrease.
2. Loads in market j decide to bid their demand into market i. This causes L^j to decrease and L^i to increase.

The net effect on price of the two actions is equivalent. Without loss of generality we decide to interpret all flows as the effect of load cross-bidding.

To incorporate this behavior into the model, we introduce a new variable q_d^i, representing the actual quantity bid into market i at time d. The variable L_d^i is interpreted as the native load of the market, that is the load, which is physically located inside the market's borders. Price in market i is a function of the total load and supply bid into this market,

$$P_d^i = e^{a^i q^i + b^i}, \quad i = 1..2$$

The relationship between q and L for the two-market example can be written as,

$$q_d^1 = L_d^1 + F_d^{12}$$

$$q_d^2 = L_d^2 - F_d^{12}$$

where F_d^{12} is the flow from market 1 to market 2, which can be positive or negative.

The power has to be balanced between the markets

$$q_d^1 + q_d^2 = L_d^1 + L_d^2$$

and the tie-line flow F_d^{12} is bounded by F_d^{max}.

$$\left|F_d^{12}\right| = \left|q_d^1 - L_d^1\right| = \left|L_d^2 - q_d^2\right| \le F^{\max}$$

Next we need to address the question of how much cross-bidding of load will occur in a given day. We assume that the load is rational, and that if there exists a price differential between the markets, load in the expensive market will submit bids into the cheaper market. The magnitude of the cross bidding is limited by the capacity of the transmission line. Thus load bids will keep shifting form the expensive to the cheap market until one of the following occurs:

1. The prices equalize, thus removing any incentive for further cross-bidding.
2. The transmission line becomes congested, preventing the loads from transferring the power back to their own location.

The first case corresponds to the following mathematical condition,

$$P_d^1 = P_d^2$$

$$a^1 q_d^1 + b_d^1 = a^2 q_d^2 + b_d^2$$

The flow necessary to reach price equality, as a function of native load and supply states, is given by,

$$\hat{F}_d^{12} = \frac{1}{a^1 + a^2}\left[\left(a^2 L_d^2 + b_d^2\right) - \left(a^1 L_d^1 + b_d^1\right)\right].$$

The transmission constraint F^{max} limits the flow both ways.

$$-F^{\max} \le F_d^{12} \le F^{\max}$$

The actual flow between the markets F_d^{12}, accounting for the limits, can therefore be written as,

$$F_d^{12} = \max\left\{\min\left(\hat{F}_d^{12}, F^{\max}\right), -F^{\max}\right\}.$$

Prices in two markets, P_d^1 and P_d^2, are equal always equal, until the transmission flow reaches the maximum capacity. At this point prices will diverge, and the dynamics of the two markets decouples.

5.1.2 Valuing a Transmission Right

In this section we estimate the value of a transmission right between market 1 and 2. The transmission right is interpreted as the right, but not the obligation, to transmit power from market 1 to market 2 in any day d. Furthermore we assume that the transmission right is firm, that is, it cannot be curtailed under any circumstances. The expected daily profit from owning the transmission right can therefore be expressed as:

$$C_d^{12} = E\left(\max_d\left(P_d^2 - P_d^1, 0\right)\right)$$

This is equivalent to the value of a spread option between the two markets.

5.1.3 Simulation of the Multi-Market Model

We attempt to estimate the value in ($/MWh) of owing a transmission right for one day, thirty days from today. To better illustrate the qualitative effects of moving to a multi-market environment, the following simplifications are made to the bid based price model:

1. We use a daily rather than an hourly model for price. This is achieved by substituting the vectors μ^L, μ^b, v^L, and v^b, by scalars.
2. We ignore the dynamics of the mean-process δ, forcing δ^L and δ^b to be zero.

$$d_d^1 = d_d^2 = 0$$
$$k^1 = k^2 = 0$$

The parameters for the supply process in both markets are identical. The mean-value (μ) of the load is set 7,000 MW higher in market 2 than in market one, creating a price differential. Finally, the maximum transmission capacity is set to 3,500 MW.

$$m_1 = 13000MW$$
$$m_2 = m_1 + 2F_d^{\max} = 20.000MW$$
$$F_d^{\max} = 3500MW$$

The model was then run 10.000 times for a 31-day period. The plots displayed below show the distribution of loads and prices on the 31st day.

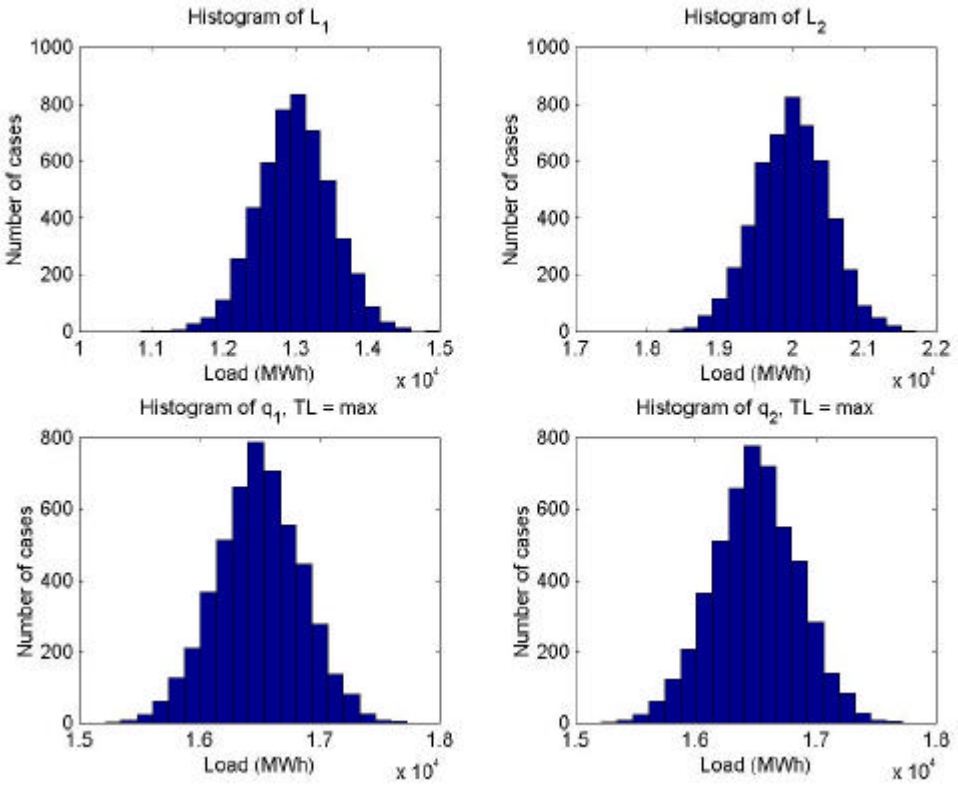

Fig. 32 Histograms of load in two markets, $F^{12} = 0$ and $F^{12} = \max$

5.1.4 Simulation Results

5.1.4.1 Simulation results for load

In Fig. 32, the histograms for the load in the two markets are presented. On the left side there are the histograms of market 1, whereas the market 2 histograms are on the right side. The load

histograms in respective markets are positioned far apart when transmission capacity is zero, and are identical, when transmission line is not congested.

5.1.4.2 Simulation results for prices

Histograms of prices of the two markets are shown in Fig. 33. Again, the two independent markets feature very different price distributions, while the distributions of uncongested prices become almost identical.

Gradual variation of transmission line capacity on two markets' prices is displayed in Fig. 34, where we can observe daily evolution of the prices. Uncorrelated at first (TL=0), the prices become more and more correlated as we increase the line capacity until they balance out. As expected, the turning point of the correlation, Fig. 36, is the mark 3500MW, which is half the difference in load between the markets.

5.1.4.3 Value of a spread option as a function of transm. line capacity

The histograms on Fig. 35 show the simulated distribution of spread option value, based on price difference between two markets. As the price in the two markets balances out, the value diminishes.

As the transmission line size increases, the markets become more and more correlated, at the same time diminishing price spread and the spread option C_{ij} value, Fig. 36.

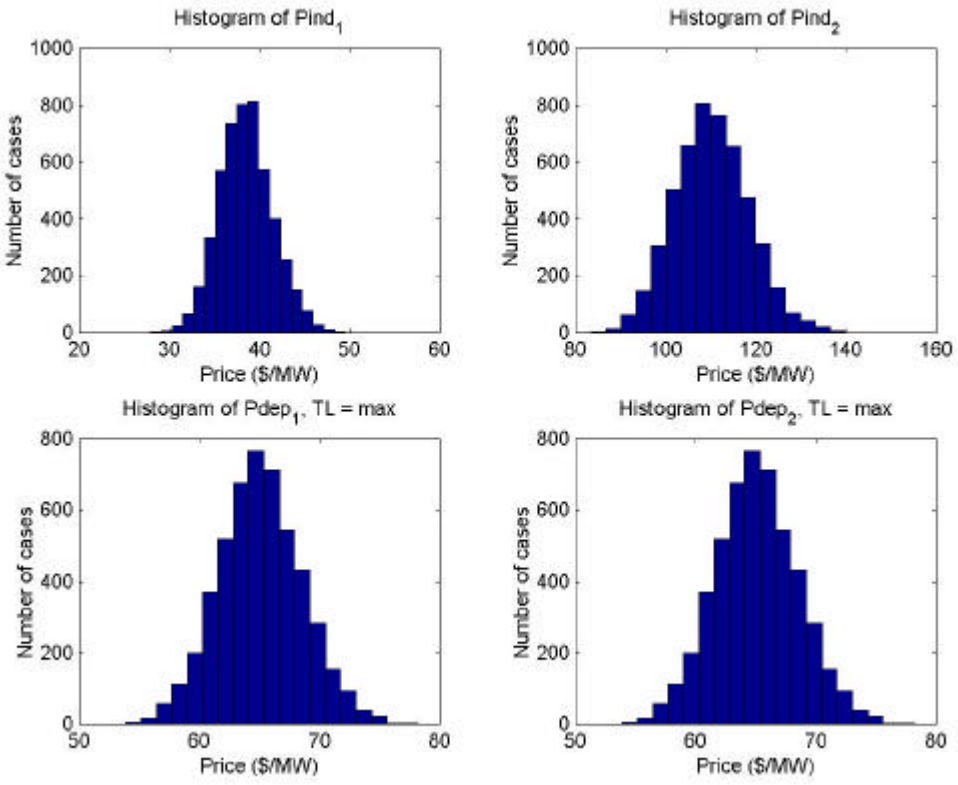

Fig. 33 Histograms of P_d^1 and P_d^2 with $F^{12} = 0$ and $F^{12} = max$

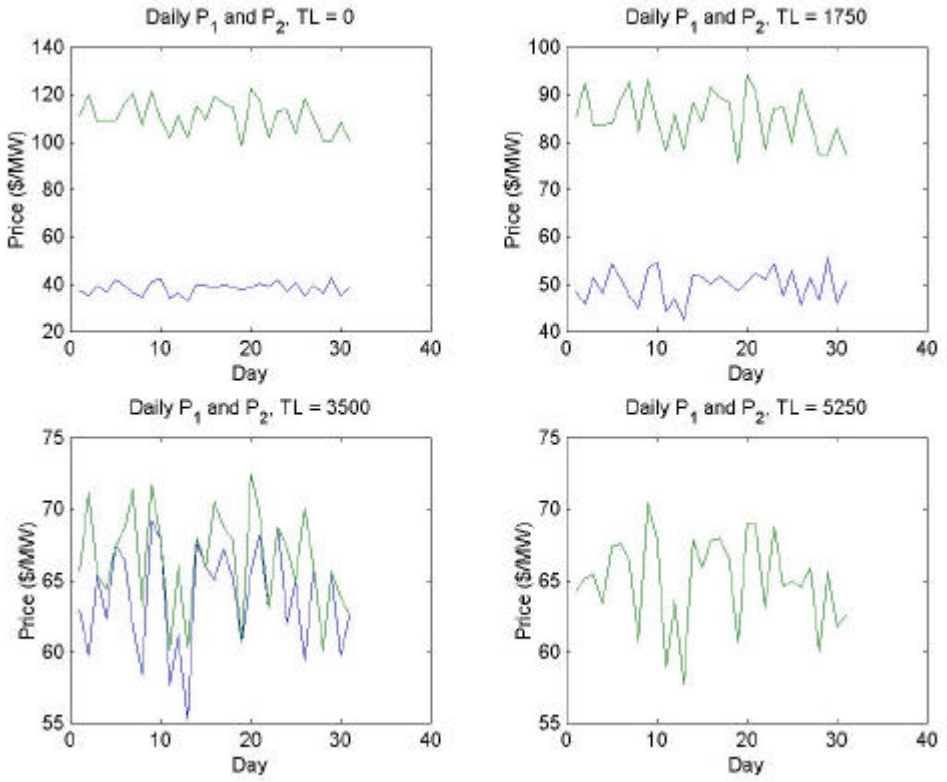

Fig. 34 Daily evolution of price in two markets with various TL capacities

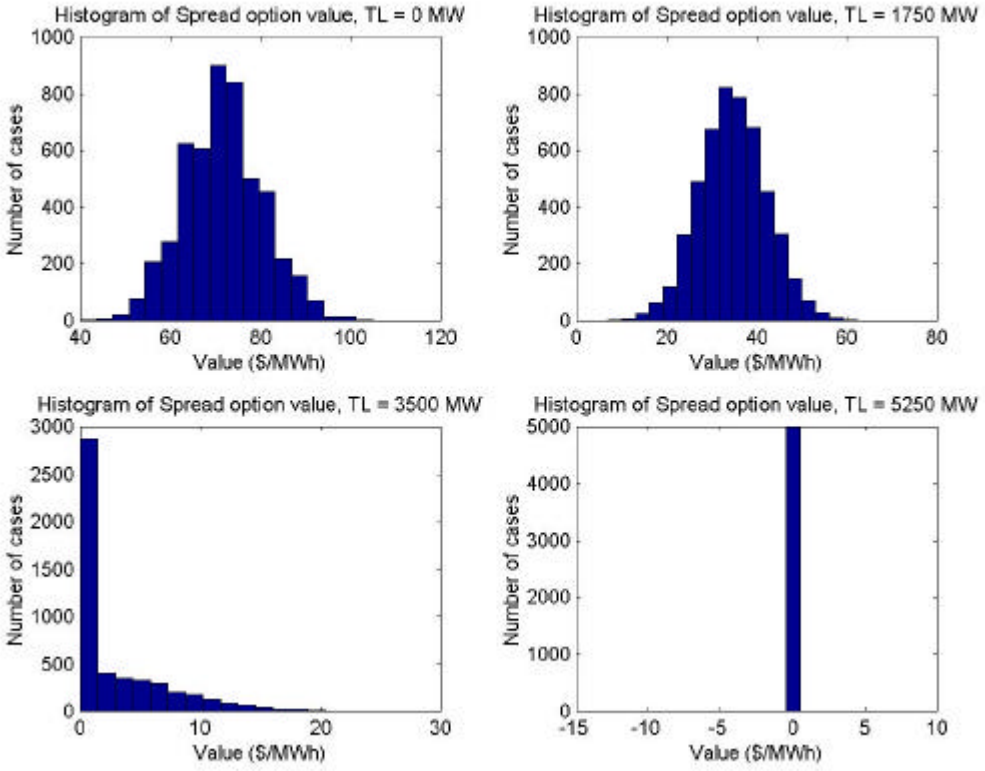

Fig. 35 Histograms of spread option value C_{12} with various TL capacities

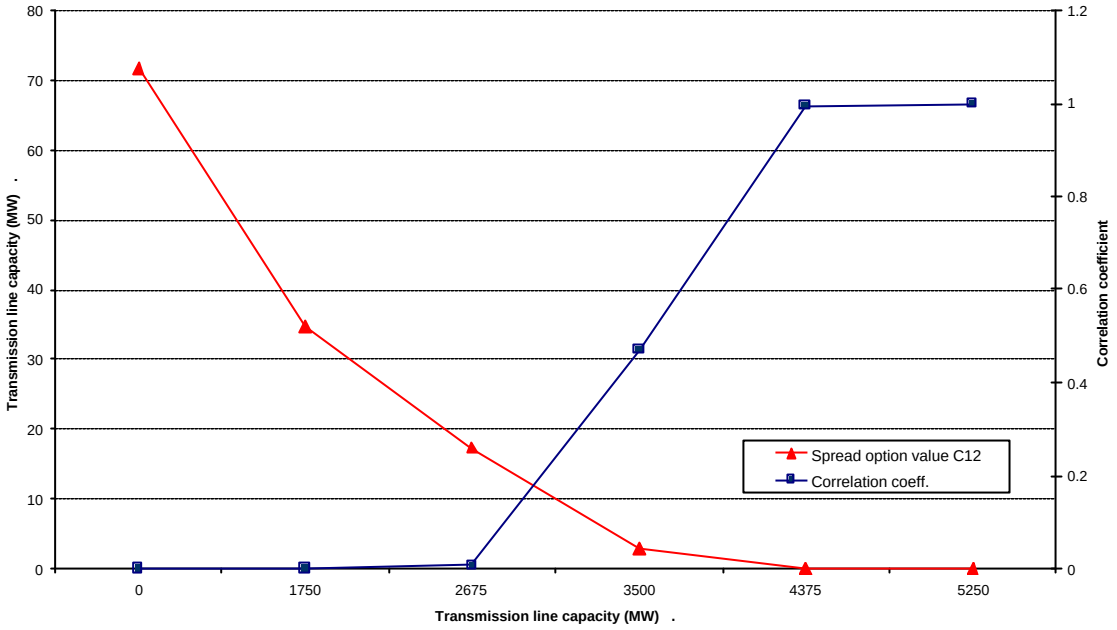

Fig. 36 Spread option value C_{12} and price correlation coefficient in various TL capacities

5.1.5 Interpretation of Simulation Results

From the histograms of prices in the previous section we can see how the capacity of the transmission lines affects the probability distribution of regional spot prices. Furthermore, the capacity of the line is strongly linked to the correlation of the two spot prices. This in turn affects the value of the spread option/transmission right. The higher the correlation between the market prices, the lower the probability that they will diverge significantly. Therefore the value of the transmission right is inversely proportional to the correlation of their prices, as shown in Fig. 36.

The advantage of the bid based model is that we make a quantitative link between the size of the transmission line in MWs, and the value of the transmission right in $/MWh. This has implications far beyond the simple valuation of transmission contracts. Consider the position of a for profit transmission provider who is contemplating whether to add a new transmission line between two markets. He needs to know whether he will be able to recover the fixed cost of investing in the line by selling transmission rights to market participants. By calibrating the bid-based model according to current price levels, and then adding the capacity of the new transmission line, the transmission owner can simulate future cash flows, and estimate the profitability of the investment.

5.2 Generation asset valuation with unit commitment constraints

The question of how to value generation assets is critical in a competitive power market. This problem is typically approached by defining the generator in terms of its efficiency (heat rate), in converting fuel to electricity. Based on this rating, the valuation is performed by modeling the

generator as a spread option between the price of the fuel used and the price of electricity. The payoff from such an option is given by

$$CF_k = \max\left\{P_k^e - C(P_k^f), 0\right\}$$

where P^e is the price of electricity and C is the cost marginal cost of production as a function of fuel price P^f.

This formulation however ignores several important constraints involved in the operation of the unit, such as start-up and shut down costs, minimum run time, and maximum ramp rate. These constraints have a significant effect on how units are bid into, and dispatched by, the spot market operator, and therefore on the owners' cash flow. By ignoring the unit commitment constraints, one is likely to undervalue plants with significant flexibility (such as micro turbines and fuel cells) while overvaluing large inflexible fossil plants.

The reason why the unit commitment problem is often ignored in the valuation of a power plant can be linked to computational complexity. In general, the operator of the unit has to solve a complex dynamic programming problem to arrive at the optimal unit commitment decision for the generator. This is a computationally intensive problem, which grows exponentially with the time horizon over which the optimization is carried out. While it is feasible to solve the unit commitment decision for a day ahead bidding problem, it is extremely challenging to extend this notion to a multi-year valuation problem.

By applying a principal component-based price model, we are able to define today's net profit from the generation asset as a function only of today's and yesterday's average spot price. By storing the mapping from the state of the spot price to the cash flow of the generator in a lookup table, we are able to simulate generator profits over multiyear periods with minimal computational complexity. The model also allows for stochastic fuel prices. This approach is described in detail in [20].

5.3 Modeling of the long term dynamics of the electricity prices

The success of deregulation cannot be measured based on short-term changes in electricity price levels. The reason many nations are going through the painful process of introducing competition into the electric utility industry is to induce technological innovation and create the right long-term incentives for future investments into the market. These two results are highly interdependent. The arrival of improved, cost efficient solutions to generate electricity leads to a surge of new investment as old, inefficient generators are pushed out of the market.

Because it decomposes price movements into supply and demand components, the bid based price model lends itself nicely to analysis of the interaction between technological innovation and investment. A possible extension of the model to deal with this approach is to change the form of the supply mean process δ^b. We postulate the form as

$$d^b_{k+1} - d^b_k = k^b (I^b_{k-T} - F_{k-T}) + s^{db} z^{db}_k .$$

In this formulation, I_k^b is an index representing the cost efficiency of the most attractive generation technology available at time k. F_k represents the average forward price over the next twelve months (currently forward prices are not available beyond this horizon). According to the formula, the growth rate of the supply bid curve (note that a negative shift in δ represents an

increase in the supply) is proportional to the difference between the cost of production index and the forward price. The rationale is that if the forward price goes far above the cost of production, the expected profit from owning a unit will be high, and there will be a rush to invest in new generation. Note that we have included a delay term T in the index as well as the forward price. This corresponds to the delay between the decision to invest in the unit, and the time that the unit actually comes on-line. The simulation and further analysis of this formulation is left for future work.

6 Conclusion

A wide range of literature has evolved on the modeling of electricity markets, and the associated price dynamics. In the introduction we listed a number of approaches; quantitative, cost based, economic equilibrium, agent based, and experimental, all with their own advantages and drawbacks. The modeling approach used is generally dependent on the type of problem addressed. A marketer may use one type of model for optimizing the short term bidding of assets, another for hedging in the forward markets, yet a third model for evaluating investment decisions. Ultimately, the user would like all of these models to reflect the best current estimate of the future. However, given the range of modeling approaches, it is often hard to check whether the models used are consistent with each other.

The bid-based model presented in this report is intended as a fundamental model for electricity price dynamics, and is to be used in a wide range of applications. The emphasis was placed on incorporating the unique characteristics of electricity prices, including seasonality on multiple time scales, lack of load elasticity, stochastic supply outages, strong mean reversion, and stochastic growth of load and supply.

The second emphasis was reducing the computational complexity of the model. This was achieved by applying techniques such as principal component analysis, which reduced the dimensionality of the model drastically, with a minimal loss in performance.

A timescale-separated version of the model was calibrated on real market data (New England). The lack of price data on the market makes the calibration on the supply side of the model tentative, but as more data becomes available, the parameter estimates will become firm. The scheme for calibrating the original version of the model is outlined, but its implementation is left to future research.

Multiple extensions of the model are possible to support a number of crucial decision problems in the market. We illustrated how a two-market version of the model can be used to estimate the value of transmission rights in a multi-market environment. This is applicable not only to market participants who wish to purchase these rights, but also has tremendous impact on the decision making of future for profit transmission owners. A second application of the model is illustrated in the valuation of generation assets with unit commitment constraints. Here the use of a principal component based model allows us to greatly reduce the computational complexity of the problem as illustrated in [20]. Finally, a new version of the model is proposed, to examine the long-term interaction between technological innovation and price trends.

The creation of a model which mimics the fundamental behavior of the market and its underlying physical and economic components lends itself to a number of further extensions, opening the door to future research possibilities.

7 Appendix A: Derivation of Principal Components

Using the Principal Component Analysis (PCA), it is possible to reduce the dimensionality of the problem by defining the new orthogonal basis. PCA generates a new orthogonal set of j variables, j ≤ n, where n is the number of the original variables in the observation set. They are called Principal Components (PC). The new variables are selected so that those describing the same driver can be replaced with a single new variable. Each principal component is a linear combination of the original variables. Because PCs are orthogonal to each other, there is no redundant information.

The first PC is a single axis in space. When each observation is projected on that axis, the resulting values form observations of a new variable, the variance of which is the maximum among all possible choices for the first axis. The second PC is orthogonal to the first, and the second variable's variance is again maximal among all possible choices for this axis. As more and more PCs are selected, they contain less and less variance. A simple two-dimensional example is shown of Fig. A.1, where vector **X** can be written as in (4), j = n = 2.

$$\mathbf{X} = a_1 y_1 + a_2 y_2 + \ldots + a_n y_n = b_1 n_1 + b_2 n_2 + \ldots + b_j n_j \tag{12}$$

The total number of PCs is usually equal to the number of original variables n. However, the first m PCs usually account for most of the variance in the original observations, j ≤ n. Sum of variances of the new variables equals sum of variances in the new variables.

$$\sum_{i=1}^{n} \mathrm{var}(n_i) = \sum_{i=1}^{n} \mathrm{var}(y_i) \tag{13}$$

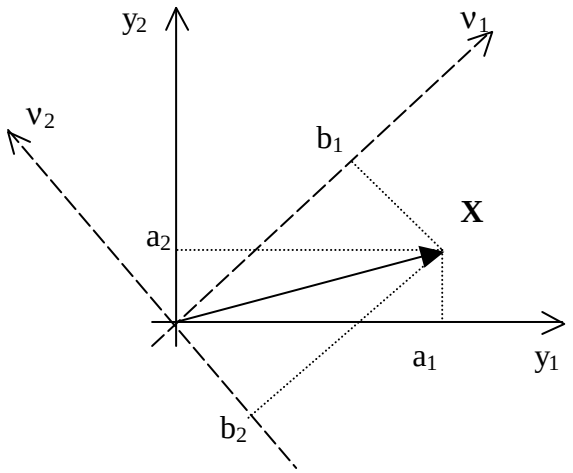

Fig. A.1: A two dimensional example of PC derivation for vector **X**

The iterative procedure of principal component derivation can be summarized in the following steps:

1. Find the largest PC: maximize the variance of $\mathbf{b}_1^T \mathbf{?}_1 = b_{11} n_1 + \ldots + b_{1n} n_n$:

$$\max \left[\mathrm{var}(\mathbf{b}_1^T \mathbf{x}) = \mathbf{b}_1^T \mathbf{C} \mathbf{b}_1 \right] \qquad \text{s.t.} \qquad \mathbf{b}_1^T \mathbf{b}_1 = \sum_{i=1}^{n} b_{1i}^2 = 1$$

where $\mathbf{b}_1$ is the vector of weights of the first principal component n_1 and $\mathbf{C}$ is the covariance matrix of $\mathbf{x}$. The condition of $\mathbf{b}^T\mathbf{b} = 1$ is necessary for the unique solution to exist; otherwise the weights could become arbitrarily big, leading to infinite variance.

$$\mathbf{b}_1 = [b_{11}, b_{12}, \ldots, b_{1n}]^T$$
$$\mathbf{?}_1 = [n_{11}, n_{12}, \ldots, n_{1n}]^T$$

2. Repeat the process for the subsequent PCs, until number of PCs = rank(**C**).

3. Determine, how many PCs are necessary to describe the process adequately. Form the reduced-order principal component [j × n] matrix n^*, where only the first m PCs are retained. A detailed description of the routine can be found in [11].

The eigenvalue λ_i, associated with ith PC corresponds to the equivalent number of variables this PC represents. A PC with an eigenvalue of $\lambda_i = 3.9$ describes as much variance as on average 3.9 original variables. By dividing the eigenvalue with the total number of PCs, j, we can obtain a total percentage on variance explained by each PC.

When all n PC have been determined, it is necessary to determine j, how many PCs are necessary to describe the data accurately enough. The three most common measures are:

1. Retain all PCs that represent more variance than original variables on average (its $\lambda_i < 1$).

2. Scree test. The incremental plot of variance accounted for by every PC is called scree plot. The number of points before leveling-off of the curve is the number of PCs retained.

3. Total variance of the data accounted for by the retained PCs. Some authors propose to retain as many PCs as to account for about 90% of the variance [12], while others propose less stringent criteria, depending on the reasons for performing the PCA [10].

8 Appendix B: Joint parameter estimation using MLE and KF

Unknown parameters of the stochastic model can be estimated using Maximum Likelihood Estimation (MLE) coupled with Kalman Filter (KF) state estimator. The iterative procedure estimates parameters of the model as to minimize the likelihood function and then computes the resulting system response using Kalman Filter. In the appendix we derive Kalman Filter and outline the MLE procedure.

8.1 Derivation of Kalman Filter

A Discrete Kalman Filter is a technique for estimation of states of a stochastic system [1]. It consists of a set of mathematical equations and provides an efficient recursive solution of the least-squares method. It addresses the problem of estimation of states **x** of a process, described by stochastic difference equations

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \mathbf{H}_t \mathbf{?}_t$$

where **A** is the system matrix, relating the system state $\mathbf{x}_t$ at time t to the next state $\mathbf{x}_{t+1}$ at time t+1 in the absence of the controlling input. **B** is the matrix relating the input $\mathbf{u}_t$ to the state $\mathbf{x}_t$. The measured noisy system output at time t y_t is

$$y_t = \mathbf{C}\mathbf{x}_t + e_t$$

The random variables h_t and ε_t represent process and measurement noise. They are assumed to be independent of each other and with normal probability distributions.

$$\mathbf{?}_t \approx N(0, \mathbf{Q})$$
$$e_t \approx N(0, \mathbf{R})$$

To write the BSM in the state space form, the system state $\mathbf{x}_t$, the process noise h_t, the input $\mathbf{u}_t$ and output y_t signals, the system matrices **A**, **B**, **C** and G, and noise covariance matrices **Q** and **R** take the following values.

$$\mathbf{x}_t = \begin{bmatrix} w_t \\ d_t \end{bmatrix} \qquad \mathbf{?}_k = \begin{bmatrix} z_k \\ z_k^d \end{bmatrix} \qquad y_t = \begin{bmatrix} w_t \end{bmatrix} \qquad \mathbf{u}_k = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 1-a, & a \\ 0, & 1 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} k \\ k \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} 1, 0 \end{bmatrix} \qquad \mathbf{G} = \begin{bmatrix} s\,, & s^d \\ 0, & s^d \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} 1, 0 \\ 0, 1 \end{bmatrix} \qquad \mathbf{R} = 1$$

8.1.1 Kalman Filter

Let's define $\hat{\mathbf{x}}_{t+1/t}$ as our a-priori estimate of state vector at step t+1, and $\tilde{\mathbf{x}}_{t+1/t+1}$ our a-posteriori estimate of state vector at t+1, given measurement $\mathbf{z}_{t+1}$. We can then define a-priori and a-posteriori estimate errors $e_{t+1/t}$ and $e_{t/t}$ as

$$\mathbf{e}_{t+1/t} = \mathbf{x}_t - \hat{\mathbf{x}}_{t+1/t}$$
$$\mathbf{e}_{t/t} = \mathbf{x}_t - \tilde{\mathbf{x}}_{t/t}$$

An estimate of the a-priori estimate error covariance is therefore $\mathbf{P}_{t+1/t}$, while of a-posteriori estimate error covariance being $\mathbf{P}_{t/t}$.

$$\mathbf{P}_{t+1/t} = E\left[\mathbf{e}_{t+1/t}\mathbf{e}_{t+1/t}^{T}\right]$$
$$\mathbf{P}_{t/t} = E\left[\mathbf{e}_{t/t}\mathbf{e}_{t/t}^{T}\right]$$

The algorithm of Kalman filter computes the equation that produces the optimal a-posteriori estimate $\tilde{\mathbf{x}}_{t+1/t+1}$ as a linear combination of the a-priori estimate $\hat{\mathbf{x}}_{t+1/t}$ and a weighted difference between an actual measurement $\mathbf{z}_t$ and predicted measurement $\mathbf{C}\hat{\mathbf{x}}_{t+1/t}$. When they agree completely, the residual $\bar{y}_{t+1/t}$ is zero.

$$\tilde{\mathbf{x}}_{t+1/t+1} = \hat{\mathbf{x}}_{t+1/t} + \mathbf{K}_{t+1}\left(y_{t+1} - \mathbf{C}\hat{\mathbf{x}}_{t+1/t}\right)$$

The factor $\mathbf{K}$ in the equation is called Kalman gain and is chosen in such a way as to minimize the a-posteriori covariance $\mathbf{P}_{t/t}$.

8.1.2 KF Algorithm

The Kalman filter algorithm is iterative procedure, that estimates process states as new measurements become available in each time step. Using initial estimates of system state $\mathbf{x}_{0/0}$ and a-posteriori error covariance $\mathbf{P}_{0/0}$, it computes the optimal a-posteriori estimate $\tilde{\mathbf{x}}_{t+1/t+1}$ and the pertaining Kalman gain K. The procedure is described below:

1. Select initial estimates: $\mathbf{x}_{0/0}$ $\mathbf{P}_{0/0}$
2. Compute time update (prediction) equations:

$\hat{\mathbf{x}}_{t+1/t} = \mathbf{A}\tilde{\mathbf{x}}_{t/t} + \mathbf{B}\mathbf{u}_t$	a-priori estimate of state vector $\mathbf{x}$
$\mathbf{P}_{t+1/t} = \mathbf{A}\mathbf{P}_{t/t}\mathbf{A}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T$	a-priori error covariance matrix

3. Compute measurement update (correction) equations:

$\mathbf{K}_{t+1} = \mathbf{P}_{t+1/t}\mathbf{C}_{t+1}^T\left[\mathbf{C}_{t+1}\mathbf{P}_{t+1/t}\mathbf{C}_{t+1}^T\right]^{-1}$	Kalman gain
$\bar{y}_{t+1/t} = y_{t+1} - \mathbf{C}\hat{\mathbf{x}}_{t+1/t}$	residual: measurement innovation
$\tilde{\mathbf{x}}_{t+1/t+1} = \hat{\mathbf{x}}_{t+1/t} + \mathbf{K}_{t+1}\bar{y}_{t+1/t}$	a-posteriori estimate of state vector $\mathbf{x}$
$\mathbf{P}_{t+1/t+1} = \left[\mathbf{I} - \mathbf{K}_{t+1}\mathbf{C}\right]\mathbf{P}_{t+1/t}$	a-posteriori error covariance estimate

4. Repeat 2 and 3 for all $t \in [1,...T]$

The procedure is schematically shown in Fig. B.1.

8.2 Maximum Likelihood Estimation of model parameters

The idea behind Maximum Likelihood Estimation is to compute the optimal parameters of the model by iteratively modifying them to minimize a likelihood function [5], [6], [7].

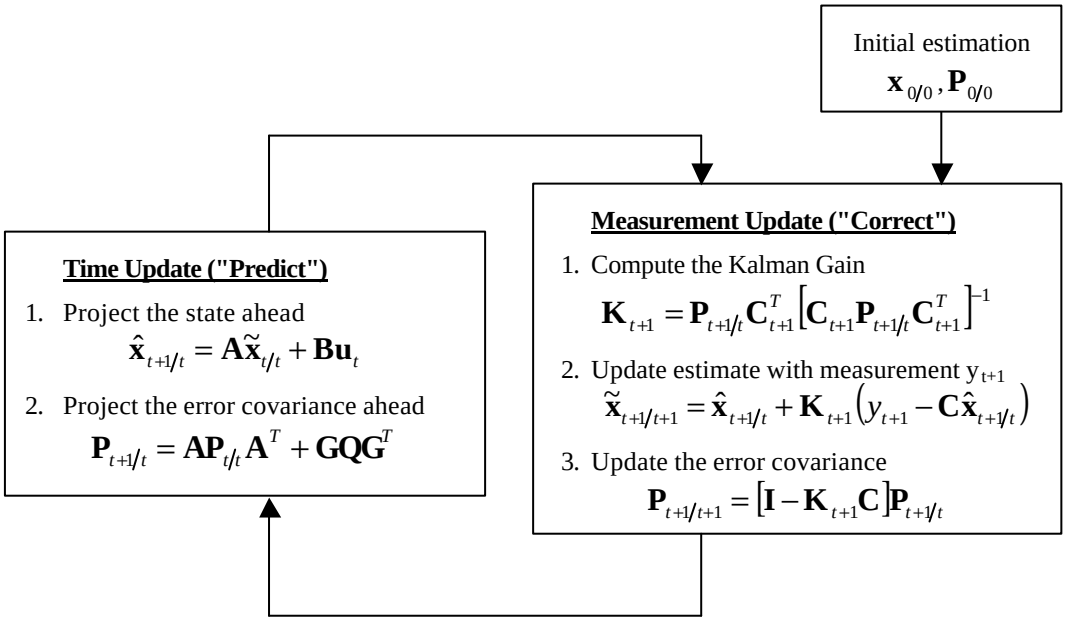

Fig. B.1 Kalman filter operation flowchart.

After constructing the model representation in the state space and setting up the KF procedure, we construct a vector of unknown parameters q that contains the unknown parameters of the model.

$$\mathbf{?} = \begin{bmatrix} a & k & s & s^d \end{bmatrix}^T$$

Using the covariance of the innovation process $\mathbf{N}_{t+1/t}$, obtained by the Kalman filter,

$$\mathbf{N}_{t+1/t} = \mathbf{C}\mathbf{P}_{t+1/t}\mathbf{C}^T$$

we can construct a log likelihood function J.

$$J = \log L = -\frac{1}{2}\sum_{t=1}^{T-1}\left[\bar{y}_{t+1/t}^T\mathbf{N}_{t+1/t}^{-1}\bar{y}_{t+1/t} + \log\left(\det\left(\mathbf{N}_{t+1/t}\right)\right)\right]$$

The procedure iteratively updates the parameter vector q according to the equation

$$\mathbf{?}^{i+1} = \mathbf{?}^i - r^i\mathbf{M}^{-1}\left(\mathbf{?}^i\right)\frac{\partial J\left(\mathbf{?}^i\right)}{\partial \mathbf{?}}$$

where **M(**a**)** is a Hessian matrix of the log likelihood function

$$\mathbf{M}(\mathbf{a}) = \frac{\partial^2 J(\mathbf{a})}{\partial a_i \partial a_j}$$

The procedure is schematically presented in Fig. B.2.

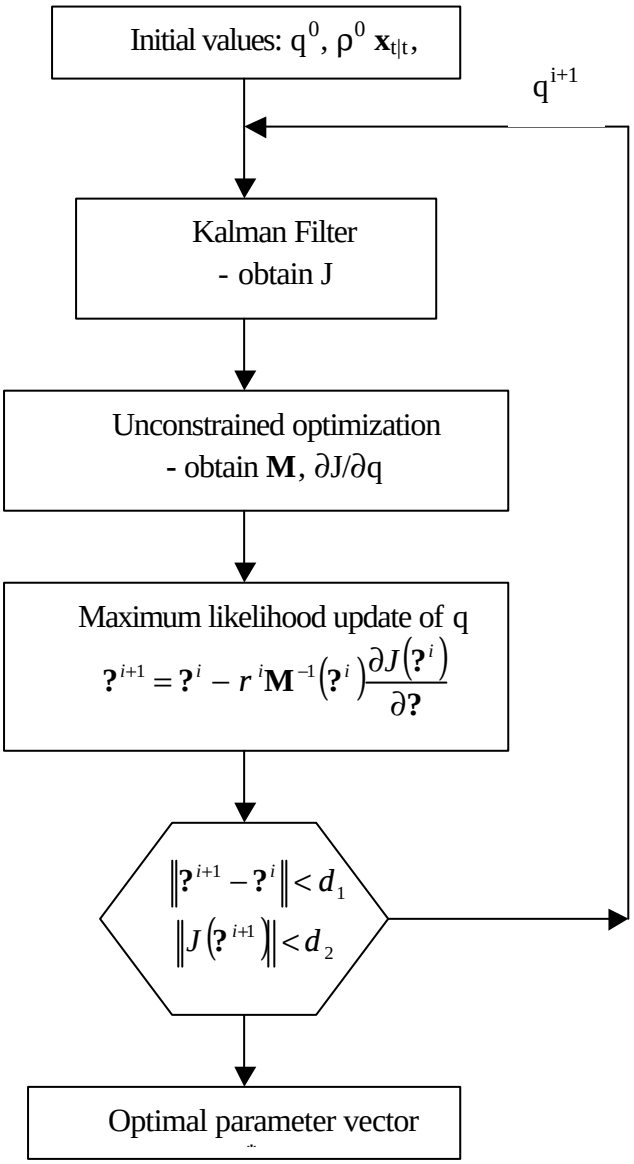

Fig. B.2 Maximum Likelihood estimation flowchart

9 Acknowledgements

We would like to gratefully acknowledge the financial support, provided by the MIT Consortium for "New Concepts and Software for Competitive Power Systems", that has made this research possible.

10 References

[1] Chui, C.K., Chen, G.: Kalman Filtering, with Real-Time Applications; Springer, New York, 1999, ISBN 3-540-64611-6.

[2] Welch, G., Bishop, G.: "An Introduction to the Kalman Filter", UNC-Chapel Hill, Working Paper TR 95-041, June 6, 2000.

[3] Mehra, R.K., "On the Identification of Variances and Adaptive Kalman Filtering"; *IEEE Trans. on Automatic Control*, Vol. AC-15, No. 2, April 1970, pp. 175-184.

[4] Mehra, R.K.: "Approaches to Adaptive Filtering", *IEEE Trans. on Automatic Control*, Vol. AC-17, No. 5, Oct. 1972, pp. 693-8.

[5] Bolland, P.J., Connor, J.T.: "A Constrained Neural Network Kalman Filter for Price Estimation in High Frequency Financial Data", *Int. Journal of Neural Systems*, Vol. 8, No. 4, August 1997, pp. 399-415.

[6] Burmeister, E., Wall, K.D.: "Kalman Filtering Estimation of Unobserved Rational Expectations with an Application to the German Hyperinflation"; *Journal of Econometrics*, Vol. 20, North-Holland Publ. 1982, pp. 255-284.

[7] Manoliu, M., Tompaidis, S.: "Energy Futures Prices: Term Structure Models with Kalman Filter Estimation"; Publication accessible over the www.caminus.com.

[8] Joy C., "Pricing, Modeling and Managing Physical Power Derivatives".

[9] Hull, J.: Options, Futures and other Derivatives, 4th ed., Prentice-Hall, 1999, ISBN 0-13-022444-8.

[10] G. H. Dunteman, Introduction to Multivariate Statistics. Sage Publications; Beverly Hills, 1984.

[11] G.H. Dunteman, Principal Component Analysis. Sage Publications; Newbury Park, 1989.

[12] B. Flury, Common Principal Component and related Multivariate Models. Wiley; New York, 1988, ISBN 0-471-63427-1

[13] S. K. Kachigan, Statistical Analysis – An Interdisciplinary Introduction to Univariate & Multivariate Methods. Radius Press, New York, 1986.

[14] Matlab Statistics Toolbox User's Guide, The MathWorks, Inc., 1999.

[15] Visudhiphan, P., Ilic, M.D., "Dynamic Game-based Modeling of Electricity Markets", *1999 IEEE PES Winter Power Meeting*, New York, New York City, February 1999.

[16] Visudhiphan, P., Ilic, M.D., "Dependence of Generation Market Power on the Demand/Supply Ratio: Analysis and Modeling", *2000 IEEE PES Winter Power Meeting*, Singapore, January 2000.

[17] Skantze, P., Chapman, J., Ilic, M.D., "Stochastic Modeling of Electric Power Prices in a Multi-Market Environment", *Transactions of IEEE PES Winter Meeting*, Singapore, January 2000.

[18] Ilic, M. D., Skantze, P., "Electric Power Systems Operation by Decision and Control: The Case Revisited", *IEEE Control Systems Magazine*, Vol. 20, No. 4, August 2000.

[19] Skantze, P., Ilic, M.D., "The Joint Dynamics of Electricity Spot and Forward markets: Implications on Formulating Dynamic Hedging Strategies", MIT Energy Laboratory Technical Report EL 00-005, MIT, November 2000.

[20] Skantze, P., Visudhiphan, P., Ilic, M.D., "Valuation of Generation Assets with Unit Commitment Constraints under Uncertain Fuel Prices", MIT Energy Laboratory Technical Report EL 00-006, MIT, November 2000.

[21] Baker, M. P., Mayfield, E. S., Parsons, J. E.: "Alternative Models of Uncertain Commodity Prices for Use with Modern Asset Pricing Methods", *The Energy Journal*, Vol. 19, No. 1.

[22] Mount, T.: "Market Power and Price Volatility in Restructured Markets for Electricity", *Proceedings of the Hawaii International Conference on System Sciences*, January 1999, Maui, Hawaii.

[23] Deng, S.: "Stochastic Models of Energy Commodity Prices and Their Applications: Mean-reversion with Jumps and Spikes", POWER report PWP-073, February 2000, Berkley, California.

[24] Schwartz, E.: "The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging", *The Journal of Finance*, Vol. 7, No. 3, July 1997.

[25] Backerman, S. R., Denton, M. J., Rassenti, S. J., Smith, V. L., "Market Power in a Deregulated Electrical Industry: An Experimental Study", Economic Science Laboratory, University of Arizona, Tuscon, AR.

[26] Rudkevich, A., Duckworth, M., Rosen, R.; "Modeling Electricity Pricing in a Deregulated Generation Industry: The Potential for Oligopoly Pricing in a Poolco", *The Energy Journal*, Vol. 19, No. 3.

[27] Baldick R., Gant R., Kahn E., "Linear Supply Function Equilibrium: Generalizations, and Limitations", POWER report PWP-078, August 2000, Berkley, California.

[28] Hobbs, B.F., Metzler, C.B., Pang, J.S., "Strategic Gaming Analysis for Electric Power Systems: An MPEC Approach", IEEE Trans. on Power Systems, Vol. 15, No. 2, May 2000, pp. 638-645.

[29] Rudkevich, A., "Supply Equilibrium in Poolco Type Power Markets: Learning All the Way", Draft, March 5, 1999.

[30] Borenstein S., Bushnell J., Wolak, F., "Diagnosing Market Power in California's Restructured Wholesale Electricity Market", August 2000.

[31] Moerch von der Fehr N.-H. , Harbord D., "Spot Market Competition in the UK Electricity Industry", Economic Journal, No. 103, May 1993, pp. 531-546. Blackwell Publishers, Cambridge, MA.