Modeling Real-Time Balancing Power Market Prices Using Combined SARIMA and Markov Processes

Magnus Olsson and Lennart Söder, *Member, IEEE*

***Abstract*—This paper describes modeling of real-time balancing power market prices by using combined seasonal auto regressive integrated moving average (SARIMA) and discrete Markov processes. The combination of such processes allows generation of price series with periods where no demand for balancing power exists. The purpose of the model is simulation of prices to construct scenario trees representing possible realization of the stochastic prices. Such scenario trees can be used in planning models based on stochastic optimization to generate bid sequences to the balancing market. The spread of the prices in the tree and the shape of the scenarios are of central importance. Model parameter estimation methods reflecting the demands on scenario trees have therefore been used. The proposed model is also applied to data from the Nordic power market. The conclusion of this paper is that the developed model is appropriate for modeling real-time balancing power prices.**

***Index Terms*—Markov processes, power prices, SARIMA processes, scenario generation, time series analysis.**

I. Nomenclature

λ_k	Real-time balancing power price, hour k.
ζ_k	Real-time balancing power quantity, hour k.
c_k	Real-time balancing price constraint, hour k.
z_k	Price constraint type indicator, hour k.
ξ_k	Spot market price, hour k.
α_k	Continuous part of price model, hour k.
β_k	Discrete part of price model, hour k.
δ_k	Price difference, hour k.
ν_k	Logarithm of δ_k.
$\Phi(z)$	Seasonal AR polynomial of order P.
$\phi(z)$	AR polynomial of order p.
$\Theta(z)$	Seasonal MA polynomial of order Q.
$\theta(z)$	MA polynomial of order q.
ω_k	White noise, hour k.
σ_ω	Standard deviation of ω_k.
s	Season length.
D	Seasonal difference model order.
d	Difference model order.
B	Backshift operator, i.e., $Bx_k = x_{k-1}$.
∇^d	Difference operator of order d.
∇_s^D	Difference operator of order D with season s.
$m_\diamond$	Average value of $\diamond$.
$\gamma(h)$	Autocovariance function (ACVF).
$\rho(h)$	Autocorrelation function (ACF).
ϕ_{hh}	Partial autocorrelation function (PACF).
a_h, b_h	Estimation weight parameters, lag h.
$\mathrm{Var}[\diamond]$	Variance of $\diamond$.
p_{ij}^t	Transition probability from state i to j after t steps in stage i.
$\#$	Cardinality operator.

Realizations of stochastic variables are indicated with tilde, e.g., $\tilde{\lambda}$. This paper considers upward and downward balancing, which are indicated by using the superscripts $\uparrow$and $\downarrow$ for upward and downward respectively.

Manuscript received July 26, 2006; revised September 27, 2007. Paper no. TPWRS-00482-2006.

The authors are with the School of Electrical Engineering, Royal Institute of Technology, 100 44 Stockholm, Sweden (e-mail: magnus.olsson@ee.kth.se).

Digital Object Identifier 10.1109/TPWRS.2008.920046

II. Introduction

In power systems that have been subjected to a liberalization process, the need for decision support tools for trading and operation has increased [1]. Focus when discussing power trading is usually on day-ahead and financial markets. However, the need for support tools also exists for real-time markets.

Important factors in planning tools for trading of electricity are the power prices, which at the time of planning are uncertain and thereby can be regarded as stochastic variables. Trading tools considering uncertainties can be based on stochastic optimization [2], [3], where the uncertain parameters are modeled using scenario trees. In this paper, a model of the real-time balancing power market prices, intended for generation of price scenarios for scenario trees, is presented. Also, existing model parameter estimation methods have been adjusted to apply to this application.

Stochastic processes, such as auto regressive integrated moving average (ARIMA) and seasonal ARIMA (SARIMA) processes [4], are established methods to model loads and market prices. Previously, models based on such processes have been used for load forecasts [5]–[7] and to forecast day-ahead spot market prices [8]–[12]. An overview of methods for forecasting loads and prices can be found in [13]. However, to capture the characteristics of balancing power market prices, e.g., periods with no demand, ARIMA or SARIMA processes alone are not sufficient. Historical time series can instead be used, as, e.g., in [14], but the number of representative scenarios are thereby limited.

In [15], an econometric analysis of the Nordic regulating power market prices is presented, but the suggested models are used for describing the underlying dependence between costs, spot prices and regulating quantities with the regulating power prices. Hence, the aims of the models presented in [15] and in this paper differ and therefore, the courses of action to model the prices are different.

Markov processes [16] are a well-known class of stochastic processes. Models based on Markov processes have previously been used for price modeling in, e.g., [17], [18] and [19].

To model the characteristics of real-time balancing power price series, this paper suggests a model based on SARIMA and discrete non-time-homogeneous Markov processes [16]. By applying this model, a sufficient number of representative scenarios can be generated in order to construct the scenario tree.

A. *Market Places*

The structure of liberalized power markets often includes a day-ahead spot market and a real-time market where unforeseen events can be balanced. Examples of power systems where such market places exists are the Nordic [20], Australian [21], Spanish [22] and PJM [23] systems. Trading on the electricity market is usually performed in trading periods. The length of the trading periods varies between systems: Whole hours are used, e.g., in the Nordic countries, Spain and USA, while half hours are used in the U.K. and New Zealand. In this paper, for simplicity, whole hours are assumed used as trading period length. The developed model is however applicable to any period length.

On the day-ahead spot market actors trade power for a number of hours in advance. The hourly demand and supply bids are submitted the day before the day of delivery. Aggregated supply and demand curves are then constructed for the different hours, and their intersection sets the spot prices and traded quantities.

In order to keep the balance between production and consumption in the system within the delivery hour, the system operator (SO) can use technical systems and trading. The technical systems automatically change the production at certain power stations equipped with frequency sensitive equipment according to grid frequency changes [24]. The technical system can be combined with a market place, where power can be purchased or sold in order to balance the system, or to free capacity in the power station equipped with frequency sensitive equipment. These market places are referred to by different names in different systems; e.g., real-time balancing market in PJM [23], real-time energy market in New England [25], frequency control ancillary services (FCAS) market in Australia [21], and regulating market in the Nordic system [20]. A compilation of balance management in Europe can be found in [22].

In the considered market structure, the SOs can balance the system by accepting bids consisting of a specified amount of power in MW and price per MWh. Balancing bids can either be upward bids, corresponding to the SO purchasing power to

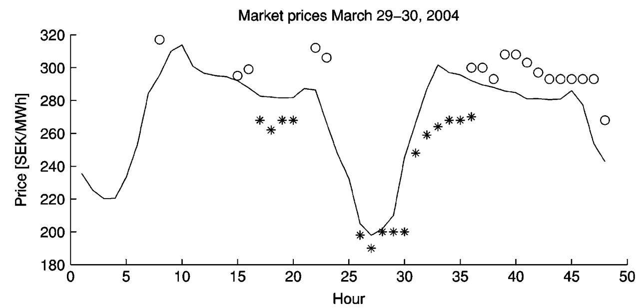

Fig. 1. Market prices from the Nordic electricity market March 29–30, 2004. Spot market prices are represented by the whole line, upward regulating market are indicated with "o," and downward with $*$."

increase the production in the system, or downward bids, corresponding to selling power to decrease the production in the system. The above considers changes in production, but also consumption changes can be traded.

The balancing power market price settlement differs between systems. Here, a marginal price principle is assumed, implying that all actors pay/are paid according to a price set by the most extreme bid accepted during the hour. Further, upward and downward balancing are assumed treated separately. Thereby, two prices are defined for each hour:

- *Upward balancing price:* The marginal price for upward balancing during the hour.
- *Downward balancing price:* The marginal price for downward balancing during the hour.

Countries having real-time balancing markets applying marginal price settlement includes the Netherlands [26], Australia [21], Spain [22], and the Nordic countries [20].

There might be situations where balancing bids of a certain type are not called during the hour. The corresponding price is then undefined. An example from the Nordic regulating power market can be studied in Fig. 1, where for the majority of hours, only one regulating price is defined. There are also a number of hours where no balancing price is defined at all, and an hour where both upward and downward balancing prices are defined (hour 36). This implies that there can be correlations between occasions when upward and downward prices are defined.

Hours where both upward and downward prices exists correspond to hours where both types of balancing were required. Upward and downward balancing are not performed simultaneously, but there can, e.g., be situations where upward balancing is required in the beginning of the hour and downward balancing at the end of the hour, and thereby both prices are defined.

Fig. 1 also shows another property of the Nordic regulating power prices: The price for regulating power is constrained by the spot price so that the upward price always is higher than or equal to the spot price, and the downward price always is lower than or equal to the spot price.

B. *Planning Tools and Scenario Trees*

Submission of bids to the market is preceded by a planning stage. When performing the planning it is important to consider the risks taken when trading. Tools considering the risks can be based on stochastic optimization [2], [3], which is a suitable tool for generation of optimal bids when uncertainties in market conditions are considered. Such tools are presented in [14], [27], [28] and [29].

In order to consider risks, uncertainties in parameters of the optimization problem must be regarded. Stochastic parameters, such as prices, are often modeled by using scenario trees representing possible realizations of the stochastic parameters. If the uncertain parameters are continuous stochastic variables, the scenario tree constitutes a discrete and finite approximation of the stochastic variables.

One method to create scenario trees is to generate a great number of representative scenarios and then reduce the size of the tree by using suitable methods. The purpose of the model presented in this paper is to generate such scenarios. A reduction method is described in [30] and [31].

An important property of price scenario trees is the span of prices in the tree at the different stages. This reflects the uncertainties of future prices during the different hours. Another important property is the correlation between prices for different hours since this represents the shape of the scenarios in the tree.

C. *Aim and Model Considerations*

The aim of this paper is to describe a real-time balancing power price model, which purpose is to generate scenarios for scenario trees for planning problems based on stochastic optimization. The following have been considered in the development of the model:

- different prices for upward and downward balancing;
- price variance;
- correlation between prices during different hours;
- correlation between occasions of existence of upward and downward prices.

The paper also presents a numerical example for the Nordic power market where estimation of model parameters is performed by using historical price series from this market.

III. Real-Time Balancing Power Price Model

The model for upward and downward balancing power prices are partly separate from each other. They are though based on the same mathematics and to avoid repetition, upward and downward balancing is not indicated unless necessary.

A. *Assumptions*

The developed model is based upon the following assumptions.

1) A market structure with a day-ahead spot market and a real-time balancing market as previously described.
2) The real-time balancing prices have upper or lower bounds, i.e., both downward and upward prices are bounded on one side.

The assumptions above are valid for several power systems, including the Nordic market [20], PJM [23] and New England [25]. Assumption 2) can be represented by bounds, e.g., set by the spot price as in the Nordic market or by a price cap as in PJM and New England.

B. *Price Model*

The developed balancing power price model considers the undefined prices by modeling the balancing power prices for hour k as

$$\lambda_k = \begin{cases} \alpha_k, & \beta_k = 1 \\ \text{not defined}, & \beta_k = 0 \end{cases} \tag{1}$$

where

- α_k is a model of the continuous behavior of the prices, i.e., not concerning the undefined prices;
- β_k is a model of the sequence of undefined and defined prices. Depending on the traded quantities, β_k takes the following values:

$$\beta_k = \begin{cases} 1, & \zeta_k > 0 \\ 0, & \zeta_k = 0 \end{cases} \tag{2}$$

where ζ_k denotes the traded quantities of balancing power.

An alternative to introduce α_k and β_k could be to let $\lambda_k = 0$ when no balancing is needed. However, this can cause problems when simulating downward prices since $\lambda_k^{\downarrow} \leq 0$ can be allowed depending on the bounds set on the prices. Hence, there will still be a need for distinguishing between the situations $\zeta_k > 0$ and $\zeta_k = 0$. Another possibility is to set $\lambda_k = \xi_k$, where ξ_k is the spot price, when no balancing is needed. This approach does not distinguish between the case that no balancing was needed and the case that the balancing price actually is the same as the spot price. Thus, for the same price both $\zeta_k = 0$ and $\zeta_k > 0$ can be valid. These two different cases should also be separated since they represent two different market situations.

1) Continuous Model: α_k are continuous variables that shall resemble the behavior of the prices when they are defined. The bounds on the prices are considered by modeling the difference between the bounds and the prices. Let z_k be a parameter indicating if the bounds are upper or lower by taking the value $z_k = -1$ for upper bounds and $z_k = 1$ for lower bounds. Further let the bounds be denoted by c_k. The continuous variables can then be expressed as

$$\alpha_k = c_k + z_k\delta_k \tag{3}$$

where $\delta_k \geq 0$ denote the difference between the bound and the price. Note that δ_k is a non-negative continuous stochastic variable.

A common method to ensure non-negativity of stochastic variables in simulations is to consider the logarithm of the stochastic variables [32]. Here $\delta_k \geq 0$, and hence the logarithm of the price difference can be simulated. However, the situation $\delta_k = 0$ must then be handled. Therefore, let $\nu_k = \ln(\delta_k + \varepsilon)$ where $\varepsilon > 0$ is a sufficiently small number. This implies that

$$\delta_k = \mathrm{e}^{\nu_k} - \varepsilon \tag{4}$$

which allows negative values of δ_k. However, as $\varepsilon \rightarrow 0$, the probability for negative δ_k moves towards zero.

The logarithm of the price differences $\nu_k^{\uparrow}$ and $\nu_k^{\downarrow}$ are considered independent of each other and thereby, no cross correlations between upward and downward balancing are con-

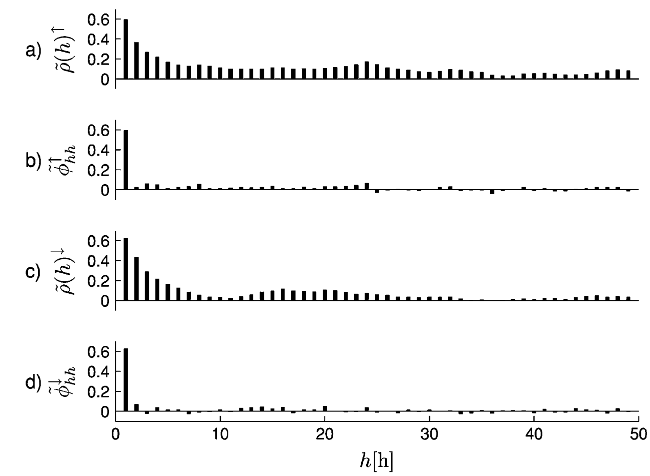

Fig. 2. Sample ACF, $\tilde{\rho}(h)$, and PACF, $\tilde{\phi}_{hh}$, for $\tilde{\nu}_k$ from the Nordic upward (a and b) and downward (c and d) regulating market prices for 2005.

sidered in this part of the price model. However, autocorrelations of the price differences for upward or downward balancing for different hours are regarded. An example of autocorrelations between subsequent hours can be studied in Fig. 2 where the sample autocorrelation (ACF) and partial autocorrelation (PACF) functions [4], [33] of $\tilde{\nu}_k^{\uparrow}$ and $\tilde{\nu}_k^{\downarrow}$ from the Nordic power market are plotted. If the price differences for different hours were not correlated with each other, $\tilde{\rho}(h)$ and $\tilde{\phi}_{hh}$ would have low values for $h > 0$.

To model $\nu_k^{\uparrow}$ and $\nu_k^{\downarrow}$, independent multiplicative SARIMA [34] models are used. The SARIMA $(p,d,q) \times (P,D,Q)_s$ model is defined by

$$\Phi(B^s)\phi(B)\nabla_s^D\nabla^d(\nu_k - m_\nu) = \Theta(B^s)\theta(B)\omega_k \qquad (5)$$

where $\nabla_s^D = (1 - B^s)^D, \nabla^d = (1 - B)^d$ and $\Phi(B^s), \phi(B), \Theta(B^s)$ and $\theta(B)$ are polynomials according to the following:

$$\begin{aligned}\Phi(B^s) &= 1 - \Phi_1 B^s - \Phi_s B^{2s} - \cdots - \Phi_P B^{Ps}\\ \Theta(B^s) &= 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \cdots + \Theta_Q B^{Qs}\end{aligned} \qquad (6)$$

and

$$\begin{aligned}\phi(B) &= 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\\ \theta(B) &= 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_p B^q.\end{aligned} \qquad (7)$$

Further, ω_k is a white noise sequence, $\omega_k \sim WN(0, \sigma_\omega)$, and m_ν denotes the mean value of ν_k. By letting $\Phi'(B) = \Phi(B)\phi(B), \Theta'(B) = \Theta(B)\theta(B)$ and $\eta_k = \nabla_s^D\nabla^d(\nu_k - m_\nu)$, the SARIMA model can be expressed as an ARMA model as follows:

$$\Phi'(B)\eta_k = \Theta'(B)\omega_k. \qquad (8)$$

2) Discrete Model: β_k are binary stochastic variables taking the values 1 or 0 indicating whether the price for hour k is defined or undefined. Thereby, the following four states can be identified:

1) Prices are neither defined for upward, nor downward balancing: $\beta_k^{\uparrow} = 0$ and $\beta_k^{\downarrow} = 0$.
2) Price defined for downward, but not for upward balancing: $\beta_k^{\uparrow} = 0$ and $\beta_k^{\downarrow} = 1$.
3) Price defined for upward, but not for downward balancing: $\beta_k^{\uparrow} = 1$ and $\beta_k^{\downarrow} = 0$.
4) Prices are defined for both upward and downward balancing: $\beta_k^{\uparrow} = 1$ and $\beta_k^{\downarrow} = 1$.

When modeling β_k it is important to consider the correlation between occasions when upward and downward balancing prices are defined. By using a Markov model it is possible to capture the correlation between upward and downward balancing, and also correlations between different hours. In this paper a non-time-homogeneous Markov [16] model is used to simulate the process $\{\beta_k^{\uparrow}, \beta_k^{\downarrow}\}$. The different states are associated with transition probabilities that depend on the number of steps in the process that have remained in the same state. These probabilities are denoted p_{ij}^t, where t is the number of steps in the same state. By randomizing according to p_{ij}^t, the series $\{\tilde{\beta}_k^{\uparrow}, \tilde{\beta}_k^{\downarrow}\}_{k=1}^K$ can be generated.

C. Scenario Generation

Generation of a price scenario of length K is performed by the following steps:

1) Generate $\{\tilde{\nu}_k^{\uparrow} - m_\nu^{\uparrow}\}_{k=1}^K$ and $\{\tilde{\nu}_k^{\downarrow} - m_\nu^{\downarrow}\}_{k=1}^K$ using (5).
2) Calculate $\{\tilde{\delta}_k^{\uparrow}\}_{k=1}^K$ and $\{\tilde{\delta}_k^{\downarrow}\}_{k=1}^K$ using (4).
3) Calculate $\{\tilde{\alpha}_k^{\uparrow}\}_{k=1}^K$ and $\{\tilde{\alpha}_k^{\downarrow}\}_{k=1}^K$ using (3).
4) Generate $\{\tilde{\beta}_k^{\uparrow}\}_{k=1}^K$ and $\{\tilde{\beta}_k^{\downarrow}\}_{k=1}^K$ by simulating the Markov process.
5) Generate $\{\tilde{\lambda}_k^{\uparrow}\}_{k=1}^K$ and $\{\tilde{\lambda}_k^{\downarrow}\}_{k=1}^K$ using (1).

IV. Model Parameter Estimation

The parameter estimation methods described in this paper require price series. The basis is a limited number of times series of prices that are assumed representative for the period to which the generated scenarios will be applied.

A. SARIMA Parameters

The parameters of the SARIMA model of $\{\nu_k - m_\nu\}$ must be estimated using methods that generates estimates relevant for the application, which in this case is generation of scenarios to a scenario tree. When applied in optimization problems, the spread of the scenario tree is of great importance, implying that the variance of the SARIMA model is essential. Also the shape of the price scenarios are important for the optimization. This means that the ACF, PACF and variance of the SARIMA model should be in focus when estimating the parameters. The parameters that need to be estimated are the following:

- Model orders $(p,d,q), (P,D,Q)$ and s.
- Coefficients of the polynomials, i.e., $\Phi_i, i = 1,\ldots,P, \phi_i, i = 1,\ldots,p, \Theta_i, i = 1,\ldots,Q$ and $\theta_i, i = 1,\ldots,q$.
- White noise variance, σ_ω^2.

The estimations of the model orders, the coefficients and the white noise variance are somewhat iterative. First, preliminary model orders are set, and then the coefficients and variance are estimated. The resulting model is then validated using validation data, whereafter modification of the model orders might be necessary, and a new estimation of coefficients and variance takes place, and so on.

1) Operator ∇ and calculation of $\tilde{\gamma}(h), \tilde{\rho}(h)$ and $\tilde{\phi}_{hh}$: When calculating $\tilde{\rho}(h)$ and $\tilde{\phi}_{hh}$ for a given series when considering the SARIMA model, the occasions where no prices are defined, i.e., $\tilde{\beta}_k = 0$, should not be considered. This because the SARIMA model should reflect the continuous behavior of the prices. Therefore, given a time series $\{\tilde{\nu}_k, \tilde{\beta}_k\}_{k=1}^{K}$, the auto covariance function (ACVF) of the balancing prices for lag h is calculated as

$$\tilde{\gamma}(h) = \frac{1}{\sum_{k=1}^{K} \tilde{\beta}_k} \sum_{k=1}^{K-h} (\tilde{\nu}_k - m_{\tilde{\nu}})(\tilde{\nu}_{k+h} - m_{\tilde{\nu}})\tilde{\beta}_k \tilde{\beta}_{k+h} \quad (9)$$

where $m_{\tilde{\nu}}$ denotes the mean value of $\tilde{\nu}_k$. By using values for $\tilde{\gamma}(h)$ obtained by using (9), values for $\tilde{\rho}(h)$ and $\tilde{\phi}_{hh}$ can be calculated by using conventional methods [33], [34].

In the estimation process, $\tilde{\rho}(h)$ and $\tilde{\phi}_{hh}$ will be calculated for differentiated time series for various values of d and D. Thus, $\nabla^d(\tilde{\nu}_k, \tilde{\beta}_k)$ needs to be calculated where the operator ∇ is defined as

$$\nabla(\tilde{\nu}_k, \tilde{\beta}_k) = (\tilde{\nu}_k - \tilde{\nu}_{k-1}, \tilde{\beta}_k \tilde{\beta}_{k-1}). \quad (10)$$

Thereby, ∇ works for $\tilde{\nu}_k$ according to the standard way, while for $\tilde{\beta}_k$ the differentiated values becomes one only if $\tilde{\beta}_k = 1$ and $\tilde{\beta}_{k-1} = 1$. Similarly, the operator ∇_s is defined as

$$\nabla_s(\tilde{\nu}_k, \tilde{\beta}_k) = (\tilde{\nu}_k - \tilde{\nu}_{k-s}, \tilde{\beta}_k \tilde{\beta}_{k-s}). \quad (11)$$

2) Model Orders: The model orders $(p, d, q), (P, D, Q)$ and s are estimated by using conventional methods described in, e.g., [4], [33] and [34]. These methods are based on observation of $\tilde{\rho}(h)$ and $\tilde{\phi}_{hh}$. The difference model orders (d, D) and the seasonal parameter s are set so differentiated data, $\{\nabla_s^D \nabla^d(\tilde{\nu}_k, \tilde{\beta}_k)\}_{k=1}^{K}$, becomes stationary. Appropriate initial values of the model orders (P, Q) and (p, q) are then set by studying the decrease and seasonal behavior of $\tilde{\rho}(h)$ and $\tilde{\phi}_{hh}$ for $\{\nabla_s^D \nabla^d(\tilde{\nu}_k, \tilde{\beta}_k)\}_{k=1}^{K}$.

3) Model Coefficients and Variance: Estimation of the coefficients of the polynomials $\Phi(B), \phi(B), \Theta(B)$ and $\theta(B)$, and the white noise sequence variance is performed in two steps: 1) estimation of $\Phi_i, i = 1 \cdots P, \phi_i, i = 1 \cdots p, \Theta_i, i = 1 \cdots Q, \theta_i, i = 1 \cdots q$ and 2) estimation of σ_ω.

a) Step 1: The first step is accomplished by minimizing the weighted square errors between sample and model ACF and PACF, i.e.,

$$\min_{\Phi_i, \phi_i, \Theta_i, \theta_i} \sum_{h \in \mathcal{H}} a_h [\rho(h) - \tilde{\rho}(h)]^2 + \sum_{h \in \mathcal{H}} b_h [\phi_{hh} - \tilde{\phi}_{hh}]^2 \quad (12)$$

where $a_h \geq 0$ and $b_h \geq 0$ are chosen weights. $\tilde{\rho}(h)$ and $\tilde{\phi}_{hh}$ are found through $\tilde{\gamma}(h)$, which is calculated according to (9) for $\{\nabla_s^D \nabla^d(\tilde{\nu}_k, \tilde{\beta}_k)\}_{k=1}^{K}$. $\rho(h)$ and ϕ_{hh} are the model ACF and PACF and are thereby functions of Φ_i, ϕ_i, Θ_i and θ_i. This expression can be regarded as an modification of the Yule–Walker estimation of parameters [34].

To ensure causality and invertibility of $\Phi'(z) = \Phi(z)\phi(z)$ and $\Theta'(z) = \Theta(z)\theta(z)$ respectively, constraints must be applied to the optimization. The process is causal if the roots of $\Phi'(z) = 0$ lies outside the unit circle. That is $\Phi'(z) = 0$ only when $|z| > 1$. The process is invertible if the corresponding applies for $\Theta'(z)$, i.e., $\Theta'(z) = 0$ only when $|z| > 1$.

b) Step 2: The second step of the parameter estimation is to estimate the variance of the white noise sequence, σ_ω^2. The variance of the ARMA$(P+p, Q+q)$ interpretation of the SARIMA model in (8) for time step k can be expressed as

$$\mathrm{Var}[\eta_k] = \sigma_\omega^2 \sum_{i=0}^{k} \psi_i, \quad \forall\, k \in \mathcal{K} \quad (13)$$

where the ψ_i's can be calculated by solving

$$\begin{aligned}(\psi_0 + \psi_1 z + \psi_2 z^2 + \cdots) \\ \times (1 - \phi_1' z - \cdots - \phi_{P+p}' z^{P+p}) \\ = (1 + \theta_1' z + \cdots + \theta_{Q+q}' z^{Q+q}). \end{aligned} \quad (14)$$

Thereby, the ψ_i coefficients are functions of Φ_i, ϕ_i, Θ_i and θ_i, and thus known at this stage.

The variance of the ARMA model should resemble the sample variance. By replacing the left hand side in (13) with the sample variance of $\{\nabla_s^D \nabla^d(\tilde{\nu}_k, \tilde{\beta}_k)\}_{k=1}^{K}$ this is accomplished. The sample variance should be calculated only with data where $\tilde{\beta}_k = 1$ in order to reflect the continuous behavior of the prices. Equation (13) thereby defines an over constrained system of linear equations where the only unknown variable is σ_ω^2, which can be solved by applying, e.g., the least square method.

B. Markov Parameters

The parameters of the discrete Markov model representing the discontinuous behavior of the balancing market prices, β_k, must also be estimated. These parameters consist of the transition probabilities, $\{p_{ij}^t\}_{i,j=1}^{4}$, which can be estimated using series of the market prices. Let $(\tilde{\beta}_k^{\uparrow}, \tilde{\beta}_k^{\downarrow})$ be pairs of binary variables indicating the states of the upward and downward prices obtained from a given price series at hour k. Let the variable $\tilde{e}_k$ be the state of the prices at stage k, i.e.,

$$\tilde{e}_k = \begin{cases} 1, & \text{if } (\tilde{\beta}_k^{\uparrow}, \tilde{\beta}_k^{\downarrow}) = (0\ 0) \\ 2, & \text{if } (\tilde{\beta}_k^{\uparrow}, \tilde{\beta}_k^{\downarrow}) = (0\ 1) \\ 3, & \text{if } (\tilde{\beta}_k^{\uparrow}, \tilde{\beta}_k^{\downarrow}) = (1\ 0) \\ 4, & \text{if } (\tilde{\beta}_k^{\uparrow}, \tilde{\beta}_k^{\downarrow}) = (1\ 1) \end{cases}$$
$$k = 1, \ldots, K. \quad (15)$$

By defining the sets

$$\begin{aligned}\mathcal{E}_{ij}^t = \{\tilde{e}_k : \tilde{e}_k = j, \tilde{e}_{k-1} \\ = i, \ldots, \tilde{e}_{k-t} = i, k = t+1, \ldots, K\}\end{aligned} \quad (16)$$

for $i, j = 1, \ldots, 4$, the transition probabilities can be calculated as

$$p_{ij}^t = \frac{\#\mathcal{E}_{ij}^t}{\sum_{l=1}^{4} \#\mathcal{E}_{il}^t}, \quad i, j = 1, \ldots, 4. \quad (17)$$

V. Case Study

In this case study, historical hourly price series from the Nordic spot and regulating power markets for 2005 were

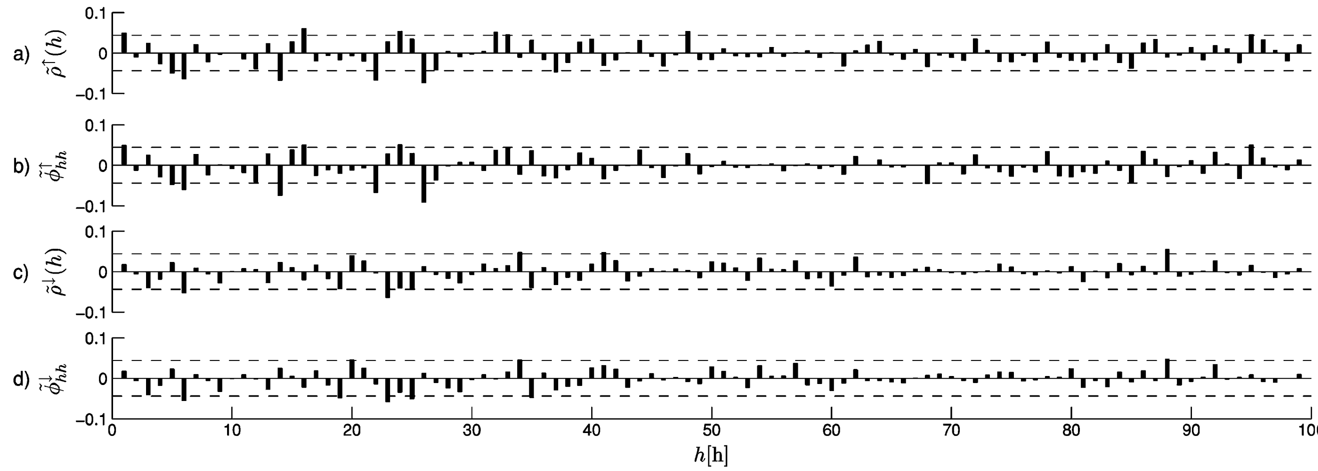

Fig. 3. The $\tilde{\rho}(h)$ and $\tilde{\phi}_{hh}$ calculated for the residuals for upward and downward regulation in the case study. Graphs a) and b) corresponds to upward and c) and d) to downward regulation. The horizontal lines represent the $\pm 1.96/\sqrt{N}$ interval.

used for parameter estimation and validation. This means that the estimated model will reflect the situation on the Nordic regulating market in 2005. The data consists of whole weeks, resulting in data series of 8568 spot, upward and downward regulating prices. Hence, the available historical data set is $\{\tilde{\xi}_k, \tilde{\alpha}_k^{\uparrow}, \tilde{\beta}_k^{\uparrow}, \tilde{\alpha}_k^{\downarrow}, \tilde{\beta}_k^{\downarrow}\}_{k=1}^{8568}$.

On the Nordic market, the spot market quantities and prices are released before submission of bids to the real-time market. The constraints on the real-time market prices are set by the spot prices for the different hours so that $\lambda_k^{\uparrow} \geq \xi_k \forall\, k : \zeta_k^{\uparrow} > 0$ and $\lambda_k^{\downarrow} \leq \xi_k \forall\, k : \zeta_k^{\downarrow} > 0$ [20]. This implies that (3) can be written as

$$\begin{aligned} \alpha_k^{\uparrow} &= \xi_k + \delta_k^{\uparrow}, \\ \alpha_k^{\downarrow} &= \xi_k - \delta_k^{\downarrow}. \end{aligned} \tag{18}$$

A. SARIMA Model

Before estimating the SARIMA model parameters, the historical price differences, $\tilde{\delta}_k$, are calculated and then logarithmized to get $\tilde{\nu}_k$. Thus, $\tilde{\nu}_k = \ln(\tilde{\delta}_k + \varepsilon)$, where $\varepsilon = 0.1$ in this case study. The estimation of model orders and parameters was conducted according to the methods described in Section IV-A. The set $\mathcal{H}$ in (12) was set to $\{1, \ldots, 6, 24, \ldots, 27, 48, \ldots, 51\}$ for upward and $\{1, \ldots, 6, 22, \ldots, 27\}$ for downward. Further, the weights $a_h = b_h = 1$ in were used. After iteration with estimation of coefficients and diagnostics checking, the final SARIMA models became the following:

- $\nu_k^{\uparrow}$ is modeled with a SARIMA $(1,2,3) \times (1,0,2)_{24}$ process with the following polynomials:

$$\begin{aligned} \Phi^{\uparrow}(B) &= 1 + 0.0347B^{24} \\ \phi^{\uparrow}(B) &= 1 - 0.6238B \\ \Theta^{\uparrow}(B) &= 1 + 0.0852B^{24} + 0.0818B^{48} \\ \theta^{\uparrow}(B) &= 1 - 1.4184B + 0.4443B^2 + 0.1107B^3 \end{aligned} \tag{19}$$

  and with $\sigma_{\omega}^{\uparrow} = 0.0528$.

- $\nu_k^{\downarrow}$ is modeled with a SARIMA $(3,1,3) \times (1,0,1)_{22}$ process with the following polynomials:

$$\begin{aligned} \Phi^{\downarrow}(B) &= 1 - 0.0717B^{22} \\ \phi^{\downarrow}(B) &= 1 - 0.2553B - 0.2061B^2 + 0.0628B^3 \\ \Theta^{\downarrow}(B) &= 1 - 0.0040B^{22} \\ \theta^{\downarrow}(B) &= 1 - 0.3173B - 0.2114B^2 + 0.0219B^3 \end{aligned} \tag{20}$$

  and with $\sigma_{\omega}^{\downarrow} = 0.2292$.

The residuals [34] were generated for the models, whereafter $\tilde{\rho}(h)$ and $\tilde{\phi}_{hh}$ were calculated. The results can be studied in Fig. 3. The horizontal lines corresponds to the interval $\pm 1.96/\sqrt{N}$, i.e., that if the series is a white noise sequence, 5% of the values can be expected to fall outside the interval. Less than 4.5% of the values for $\tilde{\rho}(h)$ and $\tilde{\phi}_{hh}$ falls outside the interval, which indicates a good model fit. This is also supported by the Ljung–Box test [34], which support the models for both upward and downward regulation at the 95% level.

B. Markov Model

The transition probabilities of the non-time-homogeneous Markov process for a specific state i, $\{p_{ij}^t\}_{j=1}^4$, have different values for $t = 1, \ldots, T_i$. For $t \geq T_i$ the probabilities are truncated by using the mean probabilities for state i. The motivation for this is that the main deviations from the mean value of the probabilities exists for $t = 1, \ldots, T_i$, but also that the amounts of historical data decreases rapidly for $t \geq T_i$, which makes the estimation of the probabilities uncertain. In this case study, T_i for the different states are $T_1 = 3, T_2 = 4, T_3 = 4$ and $T_4 = 2$. The transition matrices $P^t = \{p_{ij}^t\}_{i,j=1}^4$ for $t = 1, 2$ are given in the following:

$$P^1 = \begin{pmatrix} 0.6244 & 0.1949 & 0.1759 & 0.0048 \\ 0.0558 & 0.9116 & 0.0209 & 0.0116 \\ 0.0795 & 0.0230 & 0.8870 & 0.0105 \\ 0.0380 & 0.2658 & 0.6329 & 0.0633 \end{pmatrix} \tag{21}$$

$$P^2 = \begin{pmatrix} 0.6751 & 0.1523 & 0.1701 & 0.0025 \\ 0.1097 & 0.8240 & 0.0383 & 0.0281 \\ 0.1297 & 0.0425 & 0.8160 & 0.0118 \\ 0 & 0.6000 & 0.4000 & 0 \end{pmatrix}. \tag{22}$$

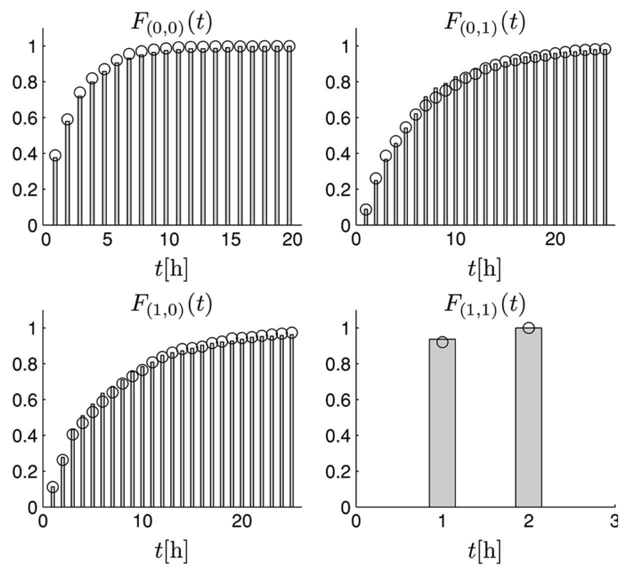

Fig. 4. Distribution functions, $F_{(\tilde{\beta}^{\uparrow},\tilde{\beta}^{\downarrow})}(t)$, for the time t for the different states, $(\tilde{\beta}^{\uparrow},\tilde{\beta}^{\downarrow})$. The bars correspond to historical data and the markers "o" to data generated with the developed Markov model.

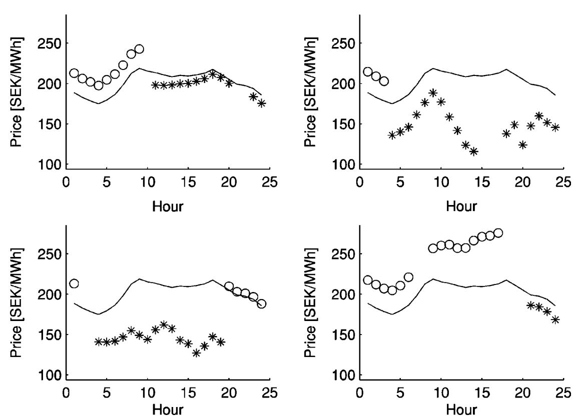

Fig. 5. Examples of scenarios generated by the developed model. The spot market prices are from January 20, 2005. Spot market prices are represented by the whole line, upward regulating prices are indicated with "o," and downward with "$*$."

Diagnostics of the Markov model was performed by comparison between distributions computed for historical and generated data for the time the process remained in the different states, $F_{(\tilde{\beta}^{\uparrow},\tilde{\beta}^{\downarrow})}(t)$. These distributions are presented in Fig. 4. As the figure shows do the distributions for the model and historical data coincide well.

The reason that $F_{(1,1)}(t)$ in Fig. 4 have values only for $t \leq 2$ is that the historical data only contains situations where the process remained in this state for one or two steps. Hence, the probability that $t > 2$ is very small.

C. Price Model

Price scenarios can now be generated according to Section III-C, where the SARIMA model polynomials are given by (19) and (20), and $\{\beta_k^{\uparrow}, \beta_k^{\downarrow}\}$ are modeled with a Markov process with transition probabilities $\{p_{ij}^t\}_{i,j=1}^4$. A few examples of generated scenarios are presented in Fig. 5.

VI. Discussion

This paper presents a model for generation of real-time balancing power price scenarios based on SARIMA and discrete non-time-homogeneous Markov processes. The development of such a model can be motivated by the following.

- If considering the present market situation, historical time series can be used to represent the real-time balancing market prices. However, if a scenario tree will be constructed, a large number of scenarios should be used in order to get a tree that well approximates the desired stochastic variables. It can thereby be difficult to find enough historical time series. Using a model to generate scenarios can solve this problem.
- If significant changes in the power system will be examined, a model representing the new market situation is needed. Such a model can be developed by modeling the present situation and validate the model using historical data, and then adjust the model to reflect the new situation.

A. Parameter Estimation

The presented estimation method was chosen because of the focus on the desired model properties $\rho(h)$, ϕ_{hh} and σ_ω. A more standard approach would be using a maximum likelihood estimator, which however do not explicitly focus on the model properties above. By using the suggested estimation method, the correct behavior of model ACF, PACF and variance can be ensured.

The parameters Φ_i, ϕ_i, Θ_i and θ_i are estimated by using a method based on minimization of the squares of $\tilde{\rho}(h) - \rho(h)$ and $\tilde{\phi}_{hh} - \phi_{hh}$, which can be regarded as a modification of the Yule–Walker estimator [34]. The estimation problem is thereby a nonlinear and non-convex optimization problem, implying that reaching a global optimum can not be assured [35]. However, this is not a problem if the reached optimum is sufficient for the application. The model diagnostics in the case study shows a good model fit and supports the choice of parameter estimation method.

B. Impact on Real-Time Balancing Market

The demand for real-time balancing power depends on unforeseen events in the system, such as outages in production units, and on forecast errors for, e.g., wind power production and load. In the Nordic system, the load during winter heavily depends on the temperature because of electric heating, and thereby will temperature forecast errors have a significant effect on the demand for balancing power.

The prices for balancing power partly depends on the demand. Strong correlation between prices subsequent hours can be explained by, e.g., forecast errors for wind or temperature. Other factors affecting the price are the water situation in any hydro power system and the market structure. The possibility of trading on the intra-day market makes it possible for actors to adjust their traded quantities according to updated weather forecasts, etc. If this possibility does not exist, all forecast errors from forecasts performed before submission of bids to the spot market must be handled on the real-time balancing market.

The developed model is a statistical price model, not explicitly taking the market structure or other underlying factors into consideration. In the presented case study, data for a whole year was used for estimation of parameters, resulting in a model behaving according to the average of the prices during that year. In reality, more information about the current situation is available, which makes it possible to adapt the model to the actual market situation. This can be performed by choosing historical estimation data from situations resembling the current one, or by adjusting the model parameters to get the desired model behavior.

VII. Conclusion

The described model is built upon a combination of SARIMA and Markov process, both well known and widely used in different applications and areas. However, the combination of the two is a new approach of modeling real-time balancing power prices. This paper concludes that the generated price scenarios shows the desired behavior and resemble real-time series of balancing power prices. It can therefore be concluded that the developed model is suitable to use for generation of real-time balancing power price scenarios.

References

[1] W. H. Dunn, M. A. Rossini, and B. Avramovic, "Impact of market restructuring on power systems operation," *IEEE Comput. Appl. Power*, vol. 8, no. 1, pp. 42–47, Jan. 1995.

[2] S. W. Wallace and S.-E. Fleten, , A. Ruszczynski and A. Shapiro, Eds., *Stochastic Programming in Energy*, ser. Handbooks in Operations Research and Management. Amsterdam, The Netherlands: North-Holland, 2003, vol. 10.

[3] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*. New York: Springer-Verlag, 1997.

[4] G. E. P. Box and G. M. Jenkins, *Turn- Series Analysis: Forecasting and Control*. San Francisco, CA: Holden-Day, 1976.

[5] I. Moghram and S. Rahman, "Analysis and evaluation of five short-term load forecasting techniques," *IEEE Trans. Power Syst.*, vol. 4, no. 4, pp. 1484–1491, Nov. 1989.

[6] S. J. Huang and K. R. Shih, "Short-term load forecasting via ARMA model identification including non-Gaussian process considerations," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 673–679, May 2003.

[7] N. Amjady, "Short-term hourly load forecasting using time-series modeling with peak load estimation capability," *IEEE Trans. Power Syst.*, vol. 16, no. 3, pp. 498–505, Aug. 2001.

[8] J. Contreras, R. Espinola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1014–1020, Aug. 2003.

[9] F. J. Nogales, J. Conlreras, A. J. Conejo, and R. Espinola, "Forecasting next-day electricity prices by time series models," *IEEE Trans. Power Syst.*, vol. 17, no. 2, pp. 342–348, May 2002.

[10] A. M. Calmarza and J. I. de la Fuente, "New forecasting method for the residual demand curves using lime series (ARIMA) models," in *Proc. Int. Conf. Probabilistic Methods Applied to Power Systems (PMA/'S) 2002*, Naples, Italy, Sep. 2002.

[11] M. Zhou, Z. Van, Y. X. Ni, G. Li, and Y. Nie, "Electricity price forecasting with confidence-interval estimation through an extended ARIMA approach," *Proc. Inst. Elect. Eng., Gen., Transm., Distrib.*, vol. 153, no. 2, pp. 187–195, Mar. 2006.

[12] A. J. Conejo, M. A. Plazas, R. Espinola, and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 1035–1042, May 2005.

[13] D. W. Bunn, "Forecasting loads and prices in competitive power markets," *Procc. IEEE*, vol. 88, no. 2, pp. 163–169, Feb. 2000.

[14] S.-E. Fleten and E. Pettersen, "Constructing bidding curves for a price-taking retailer in the Norwegian electricity market," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 701–708, May 2005.

[15] K. Skytte, "The regulating power market on the Nordic power exchange Nord pool: An econometric analysis," *Energy Econ.*, vol. 21, no. 4, pp. 295–308, Aug. 1999.

[16] J. R. Norris, *Markov Chains*. Cambridge, U.K.: Cambridge Univ. Press, 1997.

[17] A. M. Gonzalez, A. M. S. Roque, and J. Garcia-Gonzalez, "Modeling and forecasting electricity prices with input/output hidden Markov models," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 13–24, Feb. 2005.

[18] E. Ni and P. B. Luh, "Optimal integrated bidding and hydrothermal scheduling with risk management and self-scheduling requirements," in *Proc. 3rd World Congr. Intelligent Control and Automation*, Hefsi, China, 2000.

[19] G. Gutiénez-Alcarax and G. B. Sheblé, "Electricity market price dynamics: Markov process analysis," in *Proc. Int. Conf. Probabilistic Methods Applied to Power Systems (PMAPS) 2004*, Ames, IA, Sep. 2004.

[20] Nordic Grid Code 2004, Nordel, Jun. 2004. [Online]. Available: http://www.nordel.org.

[21] "Operating procedure: Frequency control ancillary services," National Electricity Market Management Company (NEMMCO), Jul. 2005. [Online]. Available: http://www.nemmco.com.au.

[22] Current State of Balance Management in Europe, ETSO, Dec. 2003. [Online]. Available: http://www.etso-net.org.

[23] Scheduling Operations, PJM Manual ll: PJM, 2007. [Online]. Available: http://www.pjm.com.

[24] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation and Control*. New York: Wiley, 1984.

[25] Market Operations, ISO New England, Inc., Apr. 2006. [Online]. Available: http://www.iso-ne.com.

[26] The Imbalance Pricing System as at 01-01-2001, revised per 26-10-2005, TenneT, Oct. 2005. [Online]. Available: http://www.tennet.nl.

[27] S. A. Gabriel, A. J. Conejo, M. A. Plazas, and S. Balakrishnan, "Optimal price and quantity determination for retail electric power contracts," *IEEE Trans. Power Syst.*, vol. 21, no. 1, pp. 180–187, Feb. 2006.

[28] H. Song, J. Lawarree, and R. W. Dahlgren, "Optimal electricity supply bidding by Markov decision process," *IEEE Trans. Power Svst.*, vol. 15, no. 2, pp. 618–624, May 2000.

[29] M. Denton, A. Palmer, R. Masiello, and P. Skantze, "Managing market risk in energy," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 494–502, May 2003.

[30] J. Dupacova, N. Growe-Kuska, and W. Romisch, "Scenario reduction in stochastic programming—an approach using probability metrics," *Math.Program.*, vol. Ser. A 95, pp. 493–511, 2003.

[31] H. Heitsch and W. Romisch, "Scenario reduction algorithm in stochastic programming," *Comput. Optim. Appl.*, vol. 24, pp. 187–206, 2003.

[32] D. G. Luenberger, *Investment Science*. Oxford, U.K.: Oxford Univ. Press, 1998.

[33] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and its Applications*. New York: Springer-Verlag, 2000.

[34] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*. New York: Springer-Verlag, 1991.

[35] S. G. Nash and A. Sorer, *Linear and Nonlinear Programming*. New York: McGraw-Hill, 1996.

Magnus Olsson was born in Mjölby, Sweden, on October 2, 1975. He received the M.Sc. degree in vehicle engineering with specialization in systems engineering and operations research in May 1999 from the Royal Institute of Technology (KTH), Stockholm, Sweden, where he is currently pursuing the Ph.D. degree in management of hydropower and wind power on the daily electricity market at the School of Electrical and Systems Engineering.

Lennart Söder (M'92) was born in Solna, Sweden, in 1956. He received the M.Sc. and Ph.D. degrees in electrical engineering from the Royal Institute of Technology (KTH), Stockholm, Sweden, in 1982 and 1988, respectively.

He is currently a Professor in electric power systems at KTH. He also works with projects concerning deregulated electricity markets and integration of wind power.