LUT
Lappeenranta
University of Technology

Sergey Voronin

PRICE SPIKE FORECASTING IN A COMPETITIVE DAY-AHEAD ENERGY MARKET

Thesis for the degree of Doctor of Science (Technology) to be presented with due permission for public examination and criticism in the Auditorium of the Student Union House at Lappeenranta University of Technology, Lappeenranta, Finland on the 1st of November, 2013, at noon.

Acta Universitatis
Lappeenrantaensis 530

Supervisor Professor Jarmo Partanen
Department of Electrical Engineering
Institute of Energy Technology (LUT Energy)
LUT School of Technology
Lappeenranta University of Technology
Finland

Reviewers Professor Risto Lahdelma
Department of Energy Technology
Aalto University
Finland

Professor Ivar Wangensteen
Department of Electric Power Engineering
Norwegian University of Science and Technology
Norway

Opponent Professor Risto Lahdelma
Department of Energy Technology
Aalto University
Finland

ISBN 978-952-265-461-8
ISBN 978-952-265-462-5 (PDF)
ISSN-L 1456-4491
ISSN 1456-4491

Lappeenrannan teknillinen yliopisto
Yliopistopaino 2013

Abstract

Sergey Voronin
Price spike forecasting in a competitive day-ahead energy market
Lappeenranta 2013
177 pages
Acta Universitatis Lappeenrantaensis 530
Diss. Lappeenranta University of Technology
ISBN 978-952-265-461-8, ISBN 978-952-265-462-5 (PDF), ISSN-L 1456-4491,
ISSN 1456-4491

Electricity price forecasting has become an important area of research in the aftermath of the worldwide deregulation of the power industry that launched competitive electricity markets now embracing all market participants including generation and retail companies, transmission network providers, and market managers.

Based on the needs of the market, a variety of approaches forecasting day-ahead electricity prices have been proposed over the last decades. However, most of the existing approaches are reasonably effective for normal range prices but disregard price spike events, which are caused by a number of complex factors and occur during periods of market stress.

In the early research, price spikes were truncated before application of the forecasting model to reduce the influence of such observations on the estimation of the model parameters; otherwise, a very large forecast error would be generated on price spike occasions. Electricity price spikes, however, are significant for energy market participants to stay competitive in a market. Accurate price spike forecasting is important for generation companies to strategically bid into the market and to optimally manage their assets; for retailer companies, since they cannot pass the spikes onto final customers, and finally, for market managers to provide better management and planning for the energy market.

This doctoral thesis aims at deriving a methodology able to accurately predict not only the day-ahead electricity prices within the normal range but also the price spikes. The Finnish day-ahead energy market of Nord Pool Spot is selected as the case market, and its structure is studied in detail.

It is almost universally agreed in the forecasting literature that no single method is best in every situation. Since the real-world problems are often complex in nature, no single model is able to capture different patterns equally well. Therefore, a hybrid methodology that enhances the modeling capabilities appears to be a possibly productive strategy for practical use when electricity prices are predicted.

The price forecasting methodology is proposed through a hybrid model applied to the price forecasting in the Finnish day-ahead energy market. The iterative search procedure

employed within the methodology is developed to tune the model parameters and select the optimal input set of the explanatory variables.

The numerical studies show that the proposed methodology has more accurate behavior than all other examined methods most recently applied to case studies of energy markets in different countries. The obtained results can be considered as providing extensive and useful information for participants of the day-ahead energy market, who have limited and uncertain information for price prediction to set up an optimal short-term operation portfolio.

Although the focus of this work is primarily on the Finnish price area of Nord Pool Spot, given the result of this work, it is very likely that the same methodology will give good results when forecasting the prices on energy markets of other countries.

Keywords: day-ahead electricity prices, price spikes, feature selection, hybrid methodology
UDC 621.3:658.8.011.1:338.534:51.001.57:519.2

Acknowledgements

This study was carried out at the Department of Electrical Engineering, Institute of Energy Technology (LUT Energy) at Lappeenranta University of Technology) between 2009 and 2013.

First of all, I would like to express my deepest gratitude to my supervisor Professor Jarmo Partanen for his valuable guidance and giving me an opportunity to be his student.

I thank the preliminary examiners of this doctoral thesis, Professor Risto Lahdelma from Aalto University and Professor Ivar Wangensteen from Norwegian University of Science and Technology for examining the manuscript and giving fruitful comments, which have significantly enhanced the work.

I would express my thanks Dr. Hanna Niemelä and Peter Jones for improving the language of the thesis and the journal papers.

I would like to thank all the LUT colleagues who have helped me in making this research a success. Special thanks are due to Dr. Matylda Jabłońska and Dmitry Kuleshov for discussions and valuable advices.

My sincere gratitude to all people who have created a perfect atmosphere during my studying and living in Lappeenranta.

My special thanks go to my father Vyacheslav and mother Liudmila for their love and support. This work would not be possible without their trust in me.

Finally, I express my gratitude to Polina for her love, great support and understanding during the years.

Sergey Voronin
7th September 2013
Lappeenranta, Finland

Dedicated

to my beloved parents

Contents

List of publications supporting the present monograph

The results described in the work were presented in part in accepted and submitted articles. The articles and the author's contribution in them are summarized below:

Publication I

Voronin S. and Partanen J. (2012), "A hybrid electricity price forecasting model for the Finnish electricity spot market," in *Proceedings of the 32nd Annual International Symposium on Forecasting*, (ed. R. Hyndman), Boston.

The author is the primary author of this conference paper. A hybrid methodology for the separate prediction of normal range electricity market prices and price spikes is presented. The author also gave an oral presentation in the conference.

The content of the conference article is presented in Chapter 4 starting from page 75.

Publication II

Voronin S., Partanen J. and Kauranne T. (2013), "A hybrid electricity price forecasting model for the Nordic electricity spot market," *International Transactions on Electrical Energy Systems*. Published online, DOI: 10.1002/etep.1734.

The author is the primary author of this journal article. The article is an extended version of the conference paper described above.

The content of the journal article is presented in Chapter 4 starting from page 75.

Publication III

Voronin S. and Partanen J. (2013), "Forecasting electricity price and demand using a hybrid approach based on wavelet transform, ARIMA and neural networks," *International Journal of Energy Research*, Published online, DOI: 10.1002/er.3067.

The author is the primary author of this journal article. A hybrid forecast method for the simultaneous prediction of price and demand in the day-ahead energy market is proposed in the paper.

The content of the journal article is presented in Chapter 5 starting from page 102.

Publication IV

Voronin S. and Partanen J. (2013), "Price forecasting in the day-ahead energy market by an iterative method with separate normal price and price spikes frameworks," *Energies*, in review.

The author is the primary author of this journal article. An iterative forecasting methodology composed of two modules separately applied to the prediction of normal prices and price spikes is proposed in the paper.

The content of the journal article is presented in Chapter 6 starting from page 116.

Abbreviations

ACF	Autocorrelation Function
AIC	Akaike Information Criterion
AMAPE	Adapted Mean Average Percentage Error
ARCH	Autoregressive Conditional Heteroscedasticity
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
ARMAX	Autoregressive Moving Average with exogenous variable
BIC	Bayesian Information Criterion
BM	Brownian Motion
CHP	Combined Heat and Power
CI	Computational Intelligence
EM	Expectation Maximization (algorithm)
FT	Fourier Transform
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
GMM	Gaussian Mixture Model
KNN	K-Nearest Neighbor
kWh	Kilowatt Hour
LM	Levenberg-Marquardt (learning algorithm)
LSQ	Least Squares (method)
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MLE	Maximum Likelihood Estimation
MLP	Multilayer Perceptron

MSE	Mean Square Error
MW	Megawatt
MWh	Megawatt Hour
NN	Neural Network
OU	Ornstein-Uhlenbeck
PACF	Partial Autocorrelation Function
pdf	probabilty density function
PNN	Probability Neural Network
PSO	Particle Swarm Optimization
RBF	Radial Basis Function
RF	Random Forest
RSS	Residual Sum of Squares
RVM	Relevance Vector Machine
SARIMA	Seasonal Autoregressive Integrated Moving Average
SDE	Stochastic Differential Equation
SDI	Supply-Demand Index
Std	standard deviation
TSO	Transmission System Operator
TWh	Terawatt Hour
WN	White Noise
WT	Wavelet Transform

1 Introduction

This chapter provides a basic background for the study addressed in this doctoral thesis. The motivation for the work is presented and previous works in the field are reviewed. The time framework for a day-ahead energy market of Nord Pool Spot is introduced. Finally, the outline of the work is given and the main scientific contributions are identified.

1.1. Motivation and background

The power markets worldwide have been strictly regulated during the most part of the 20th century, but over the last decades, they have undergone a significant restructuring and deregulation.

Before deregulation, that is, within vertically integrated power systems, electricity prices were usually regulated and the consumers were offered predetermined tariffs. The attempts to design well-functioning competitive markets that give players the correct incentives were supposed to improve production efficiency and limit market power, since in competitive electricity markets, participants have the option of trading electricity. Hence, in deregulated electricity markets, more freedom is left to the players.

One of the most pertinent questions for deregulation programs, in the light of the key objectives such as reducing electricity prices while keeping the lights on, is how to arrange the electricity trading between the generators and the buyers in the wholesale market. There is no ready-made answer to this question as there are different electricity market structures and regulatory policies in different countries. It is possible, however,

to identify two main market arrangements from the several models implemented around the world, namely the power pool and bilateral contract in parallel to a voluntary power exchange (Barroso et al., 2009).

Companies acting on the power exchange require accurate electricity price forecasts to have an opportunity to optimize the use of their portfolio by bidding or hedging against price volatility in order to get the highest possible profit. For example, generating companies acting on the power exchange compete with each other in serving the consumers' demand and have the opportunity to optimize the use of their production portfolio by pricing and bidding their available production capacity into the market. On the other hand, demand-side participants look for feasible options to avoid the high electricity market prices during peak hours. Moreover, price forecasts are of great importance for system operators, who are responsible for keeping the grid in balance. Besides, market participants are interested not only in price prediction but also in knowing the uncertainty of the forecast, which plays a significant role in decision making.

Certain unique characteristics of electricity markets make the electricity price forecasting more complex than the price forecasting of other commodities. Electric power cannot be stored economically, and further, transmission congestion influences the exchange of power. Unlike electricity demand series, electricity price series can exhibit variable means, major volatility, and significant outliers. Because of the extreme volatility reflected in price spikes, electricity price modeling and forecasting face a number of challenges. Thus, applications used to forecast the prices of other commodities are only of limited validity in electricity price forecasting and may produce large errors.

The Finnish day-ahead energy market of Nord Pool Spot is selected as the case market. The prices in the Nordic energy market are highly volatile but are not purely stochastic and, therefore, can be explained, at least partly, by background variables. Drivers affecting the prices on the market are, for example, temperature and wind power forecasts, as well as power plant availability and transmission congestions. Electricity prices on the Nord Pool Spot market are, in the long run, significantly influenced by the water level in the reservoirs of the Norwegian and Swedish hydropower plants.

With a growing proportion of energy trading being carried out on Nord Pool Spot and with the expanding geographical areas that this power exchange covers, the need for advanced market price forecasting methods has increased. Thus, prior information on market price fluctuations is a crucial concern for market participants. Short-term operation scheduling in a competitive electricity market is a challenging task because of the uncertainty associated with the future electricity prices. This approach is particularly efficient if the price forecast is of a high accuracy.

This doctoral thesis addresses the issue of forecasting day-ahead electricity market prices through development a forecasting model where an optimal input feature set and

model parameter setting are analytically selected to predict not only prices within the normal range but also price spikes.

1.2. **Objectives of the thesis**

The objectives of the thesis are:

- to review a structure of a selected case market;
- to detect a set of candidate explanatory variables that are probably influencing the day-ahead electricity price volatility and spikes;
- to investigate models built on classical (e.g. time series, stochastic, regime-switching), modern (e.g. neural networks), and hybrid (e.g. classical time series plus neural networks) approaches recently applied to case studies of price forecasting on day-ahead energy markets in different countries;
- to examine existing feature selection techniques and construct their combinations to find the best feature selection approach resulting in the highest price forecasting accuracy;
- to derive the methodology for the analysis and prediction of day-ahead electricity price signals within not only the normal range but also price spikes;
- to verify the methodology on actual data extracted for a case market, and
- to apply the obtained price forecasts to a short-term scheduling of a single market consumer.

1.3. **Previous work**

Electricity market price forecasting is a relatively new area of research, unlike the electricity demand forecasting problem (Hippert et al., 2001). Based on the needs of the market, a variety of approaches to forecast electricity prices have been proposed over the last decades.

The first group of models applied to electricity price forecasting within the context of competitive electricity markets is based on simulation of power system equipment (transmission congestions, losses, etc.) and the related cost information (marginal generation costs, heat rates, or fuel efficiencies) (Bastian et al., 1999; Fu and Li, 2006). A major drawback of this approach is the requirement of a large amount of real-time data on the existing equipment. Nevertheless, the simulation methods presented could very well be effective if used by market operators and regulators, who have the authority to collect precise equipment and operational information.

The second group is game-theory-based models, which focus on the impact of bidder strategic behavior on electricity prices. It has been stated that electricity market prices are closely related to the bidding and pricing strategies of the market participants (Guan et al., 2001; Bajpai and Singh, 2004; Chandarasupsang et al., 2007; Sadeh et al.,2009).

The third approach is based on stochastic modeling. A modified version of the geometric Brownian motion was proposed as a jump diffusion model for the stochastic modeling of electricity prices (Barlow, 2002). The robustness of various diffusion models in the case of electricity prices has been evaluated in (Barz and Johnson, 1998). The main conclusion was that the geometrical mean reverting jump-diffusion models provide the best performance and that all models without jumps appear inappropriate for modeling electricity prices. It should be noted that the main disadvantage of the stochastic modeling approaches arises from difficulties involved in incorporating physical characteristics of power systems, such as losses and transmission congestions, into mathematical (financial) models, which may produce a significant mismatch between the model output and the actual power market.

The fourth approach is based on time series models and includes two major branches: regression-based models and artificial intelligence (AI) models such as neural networks (NN) and fuzzy logic. Regression models are considered to be functions of past price observations and exogenous explanatory variables such as electricity demand and meteorological conditions. Much work has been done on electricity price forecasting with an autoregressive moving average (ARMA) approach, transfer function, and dynamic regression (Nogales et al., 2002; Contreras et al., 2003). To overcome the restrictions of linear models and to account for nonlinear patterns observed in real problems, several classes of nonlinear models have been proposed. These include threshold autoregressive (TAR-type) models (Robinson, 2000; Rambharat et al., 2005) and an autoregressive conditional heteroscedasticity (ARCH) model by Engle (Engle, 1987) and its extended version GARCH (Bollerslev, 1986; Garcia et al., 2005; Karandikar, 2009). More recently, AI models have been suggested as an alternative to the above mentioned regression-based forecasting models. Among AI models, NNs with different structures and training algorithms have been applied to electricity price forecasting (Szkuta, 1999; Nasr et al., 2001; Zhang, 2003; Zhang and Qi, 2005; Amjady, 2006; Taylor, 2006, Catalão et al., 2007; Mandal et al., 2007; He and Bo, 2009). The main strength of AI models is their flexible nonlinear modeling capability.

Linear-based models and nonlinear models have both achieved successes in their own linear or nonlinear domains. However, none of them is a universal model that is suitable for all circumstances. For example, the approximation of ARMA models to complex nonlinear problems may not be adequate, and the use of NNs to model linear problems has yielded mixed results. Since it is difficult to thoroughly know the characteristics of the data in a real problem, a hybrid methodology that has both linear and nonlinear modeling capabilities would appear to be a possibly productive strategy for practical use. It is almost universally agreed in the forecasting literature that no single method is best in every situation; largely due to the fact that real-world problems are often complex in nature, and no single model is able to capture different patterns equally well. By combining different models, different aspects of the underlying patterns may be captured. Researchers have compared various adaptive and nonadaptive linear and potentially nonlinear models and concluded that hybrid models consisting of multivariate adaptive linear and nonlinear models outperform other models for many

variables (Swanson and White, 1997). A model combining NN and ARMA has been developed (Tseng et al., 2002). The model outperformed single ARMA and NN in terms of performance accuracy measures. A hybrid model for day-ahead price forecasting, composed of linear and nonlinear relationships of prices and explanatory variables such as electricity demand was developed (Wu and Shahidehpour, 2010). A day-ahead price forecasting model was implemented by a hybrid intelligent system composing of a NN model and a genetic algorithm with an enhanced stochastic search procedure (Amjady and Hemmati, 2007).

Most of the existing approaches forecasting electricity prices are reasonably effective for normal range electricity prices but disregard price spike events, which are caused by a number of complex factors and occur during periods of market stress. These stressed market situations are associated with extreme meteorological events, unusually high demand or, more often, unexpected shortfalls in supply, caused for example by generator failures (Becker et al., 2007). In the early research, price spikes were truncated before application of the forecasting model to reduce the influence of such observations on the estimation of the model parameters; otherwise, a very large forecast error would be generated on price spike occasions (Yamin et al., 2004; Rodriguez et al., 2004; Weron, 2006).

In addition to a normal price behavior analysis, an improved analysis of spikes is important for market participants to stay competitive in a competitive market. GARCH was tested to simulate price spikes in an original price series (Keles, 2012). Spikes were incorporated into a Markov-switching model by proposing different regimes; regular and spiky (Becker et al., 2007). Spikes were introduced into diffusion models by the addition of a Poisson jump component with time varying parameters (Jabłońska et al., 2011). Data mining techniques have been applied to the spike forecasting problem (Lu et al., 2005; Zhao et al., 2007a). Most of the approaches proposed for the problem of price spike forecasting were not able to produce spikes with heights and occasions usually observed in an original price series.

Most of the work on electricity market price forecasting is concentrated on improving forecast accuracy rather than the effects of price forecast inaccuracy on market participants. Only a few approaches have been reported in the literature to deal with the problem of future price uncertainty in operation planning in competitive environments (Plazas et al., 2005; Carrion et al., 2007; Li et al., 2007). The obtained price forecasts were used in scenario-based techniques employed to derive optimal operational strategies (Zareipor et al., 2010).

1.4. **Forecasting time framework**

In most cases, the analysis presented in this work relies on hourly data. When hourly observations are not available, or for simplicity, average daily or weekly values are entered.

The time framework to forecast the electricity prices in the Nord Pool Spot day-ahead energy market is illustrated in Figure 1.1 and explained below.

The day-ahead price forecast for day D is required on day D-1 (bidding: 12:00 CET). As soon as the noon deadline to submit bids has passed, all purchase and sell orders are aggregated into two curves for each delivery hour of day D; an aggregate demand curve and an aggregate supply curve. The system price for each hour of day D is determined by the intersection of the aggregate supply and demand curves, which represent all bids and offers for the entire Nordic region and are published by the system operator on day D-1 (clearing: between 12:30 and 13:00 CET). Hence, actual price data up to 24 hours of day D-1 are available on day D-2. Therefore, when bidding for day D, price data up to hour 24 of day D-1 are considered known. As a result, the actual forecast of day-ahead prices for day D can take place between the clearing hour for day D-1 of day D-2 and the bidding hour for day D of day D-1. A detailed description of how a day-ahead market in the Nordic region works can be found in (Nord Pool Spot, 2013a).

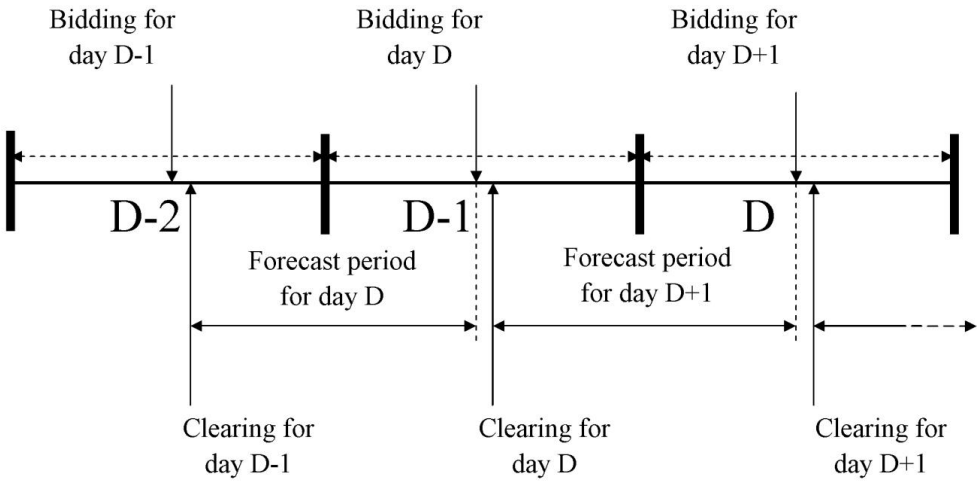

Figure 1.1. Time framework to forecast market prices in the Nord Pool Spot day-ahead energy market.

In multistep ahead prediction, the predicted price value of the current step is used to determine its value in the next step, and this cycle is repeated until the price values of the whole forecast horizon are predicted.

1.5. **Scientific contribution**

A day-ahead electricity price forecasting model is developed. The main contributions are shortly as follows:

- Classical and most recently used forecasting methodologies and their combinations are surveyed and applied to price prediction in a case energy market.
- Different feature selection techniques and their combinations are studied. The technique (combination of techniques) resulting in the most accurate price forecasting is selected.
- The forecasting methodology that is able to predict both normal range prices and price spikes with a high accuracy is proposed.
- The obtained price forecast is applied to produce an optimal short-term operation scheduling of a single market costumer.

1.6. Outline of the thesis

Chapter 2 describes the deregulated electricity markets in the Nordic region. The structure and characteristics of the electricity supply and demand in the Nordic market, the functioning of the power exchange Nord Pool Spot, and the formation of the day-ahead electricity prices are introduced.

Chapter 3 discusses the application of the classical time series approaches, stochastic and regime-switching processes to deal with the problem of day-ahead price forecasting.

Chapter 4 presents the application of a NN model as an example of modern nonlinear approaches. A hybrid methodology implying a merging of classical and modern approaches for separate normal range price and price spike forecasts is introduced.

Chapter 5 describes the process of tuning the model parameters and selection of an optimal input set through an iterative search procedure. A hybrid methodology for simultaneous prediction of price and demand in the day-ahead energy market is presented.

Chapter 6 presents a novel iterative forecasting methodology with separate normal price and price spike forecasting frameworks. This methodology is built on the findings made within the research and implemented as a combination of different forecasting engines.

Chapter 7 provides discussion and future prospects.

2 Nordic electricity market

This chapter gives an insight into the electricity market structure of the Nordic countries. Section 2.1 reviews the reasons behind the process of electricity market deregulation. Section 2.2 presents the main features of electricity to distinguish it from other commodities. Section 2.3 introduces the structure of the Nordic electricity market and the principles of a day-ahead electricity price formation.In Sections 2.4–2.5, statistics for electricity generation and consumption in the Nordic region are presented.

2.1 Deregulation

Generation, transmission, and distribution of electrical energy require huge capital investments for operation, maintenance, and expansion (Yan, 2009). In some countries, crown corporations were established and given a monopoly of generation, transmission, and distribution of electrical energy within prespecified geographical boundaries. Such a monopoly structure guaranteed a decent return on the huge investment that a single entity or a crown corporation would typically make.

Regulation became part of the electricity industry to protect the consumer from the inevitable consequences of a monopoly industry. However, the regulated electricity market is still a monopoly but watched by the government. In a regulated electric market, that is, in a vertically integrated system, local consumers have no other choice for electricity service but the local provider, and therefore, the electricity price is high and services are usually limited.

In the late 1970s, Chile was the first country to introduce competitive forces into the electricity market. This gradually led other countries to consider the option of a deregulated electricity market (Wolak, 1997). Deregulation refers to the reduction or elimination of government control allowing the generation and retail to be competitive while the transmission is kept under government control. The reason to keep the transmission sector under regulation is to establish a fair competitive environment where all competitors have equal access to the transmission network. In a deregulated market, instead of only one generation provider in a local area, there is a number of generation providers. Therefore, consumers have many options for their electricity providers and development of an optimal operation portfolio.

2.2 Electricity as a commodity

There are certain features of electricity that set it apart from an ordinary commodity and consequently, result in special power system economics. Electricity cannot be stored in economically considerable quantities. As a continuous flow, electric energy has to be consumed at the same time as it is produced. Therefore, there must be an instant balance between electricity supply and demand in the electricity market. Thus, while the store affects the aggregate demand for the majority of commodities, this effect does not exist for electric energy. The nonstorability of electricity also leads to the requirement of reserve capacity in an electric power system.

One of the key features of electricity as a commodity is the necessity for the electric energy transmission infrastructure, that is, an electric power network. From that point of view, electricity may be considered a network-based commodity.

Electric energy is uniform by nature; it is a commodity that cannot practically be differentiated in terms of product or brand as in the classic economic theory. Electric energy can be differentiated by different sources of origin (e.g. hydro, nuclear, thermo power), voltage level, and power quality characteristics (e.g. voltage and frequency deviations, reliability of supply); yet there are no physical means by which a producer that actually generated a unit of electricity (a kWh) delivered to a consumer can be recognized.

As an essential commodity, electricity is characterized by a relatively inelastic demand. This means that if the price for electricity suddenly doubles, the demand for electricity will not considerably decrease because of the absence of substitute goods.

2.3 Structure of the Nordic electricity market and price formation

The Nordic region has considerable experience in deregulated electricity markets. The Nordic electricity market was formed in 1993 in conjunction with the deregulation of the electricity markets in the region. The derivatives and energy markets were separated in 2002 to establish the power exchange Nord Pool Spot, which currently operates in

Norway, Denmark, Sweden, Finland, Estonia, Lithuania, and Latvia (Nord Pool Spot, 2013b).

The main objective of Nord Pool Spot is to balance the generation of electricity with the electricity demand, precisely and at an optimal price, that is, by equilibrium point trading. The optimal price represents the cost of producing one kWh of power from the most expensive source needing to be employed in order to balance the system. Two different physical operation markets are organized in Nord Pool Spot: Elspot and Elbas.

2.3.1 Elspot market

The Elspot market is the day-ahead energy market, where market participants submit bids for sale or purchase of electricity in the next day's 24-hour period. It is possible to submit hourly bids, block bids, and flexible hourly bids in the Elspot market. All bids consist of a price and a volume. The hourly bid specifies the amount of electricity a participant wishes to buy or sell at different prices in a certain hour. The hourly bid sets at least the highest buying or selling volume and a price limit for it, and the lowest buying or selling volume and a price limit for it. The bid may include up to 62 price steps in addition to the minimum and maximum price limits set by Nord Pool Spot. Electricity volumes between each adjacent pair of submitted price steps are linearly interpolated by Nord Pool Spot.

The participants send their bids for the following operation day before deadline at 12:00 CET. Once the market prices have been announced, the market participants receive a notification of the accepted bids and the hourly commitments of the following operation day.

2.3.2 System price

After the daily trading cycle in the Elspot market, the day-ahead system price is calculated for the following day. This price is transparent, liquid, and equal for all market participants. The system price can be used as a reference price for any financial electricity market contracts. The system price is formed at every hour of the following day. To get these hourly system prices, hourly demand and supply curves are built by combining all the selling and buying bids for each hour of the following day. The system price is obtained as the point where the demand and supply curves intersect. Figure 2.1 qualitatively shows the aggregated supply and demand curves.

The aggregated supply curve is presented in the chart with different power generation methods. The width of the bars corresponds to the generation capacity of each production form. The shaded areas illustrate the increase in the production costs of electricity caused by the price of emission allowances. The curve has various steps as a result of different costs of individual generation forms. If the demand intersects the supply curve, for example, in the coal condensing part of the curve, then hydro, nuclear power, combined heat and power (CHP), and coal condensing are used to meet the

electricity demand. In the system price calculation, the possible restrictions for the transmission capacity between different geographical areas of the Nordic countries are left out. In other words, the system price is formed with the assumption that the transmission capacities between Norway, Sweden, Finland, Denmark, Estonia, Lithuania, and Latvia are infinite. This is the reason why the system price is also denoted "the unconstrained market clearing price" that balances the sale and purchase in the exchange area.

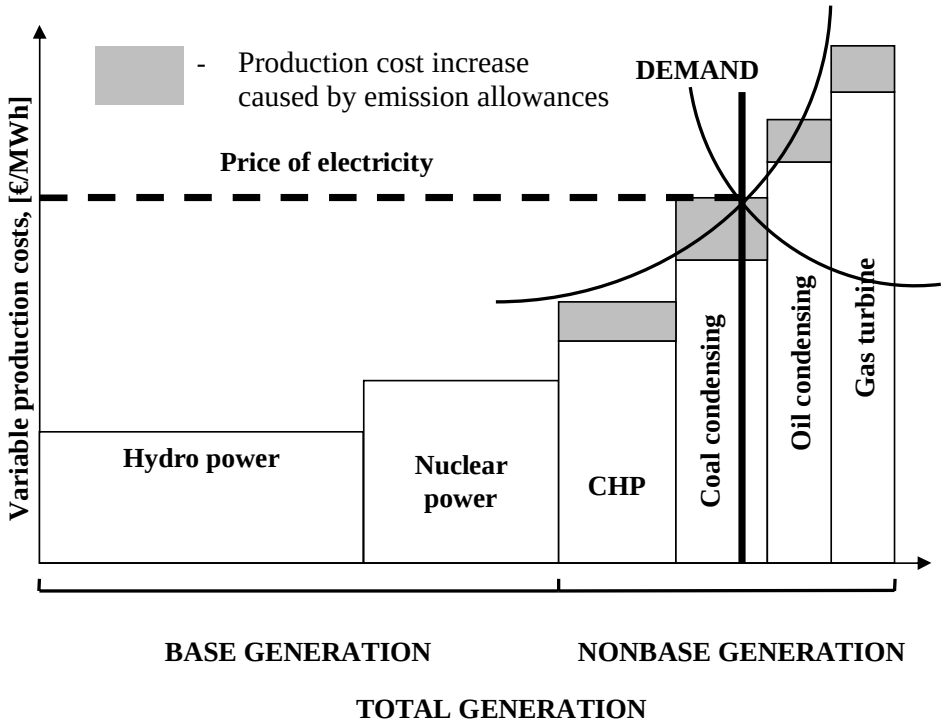

Figure 2.1. Formation of the day-ahead system price.

2.3.3 **Area price**

The total Nordic market is divided into 15 bidding areas: five in Norway, four in Sweden, two in Denmark, one in Finland, Estonia, Lithuania, and Latvia. Figure 2.2 presents the current geographic structure of the Nord Pool Spot market with a division into possible price areas when grid congestions occur.

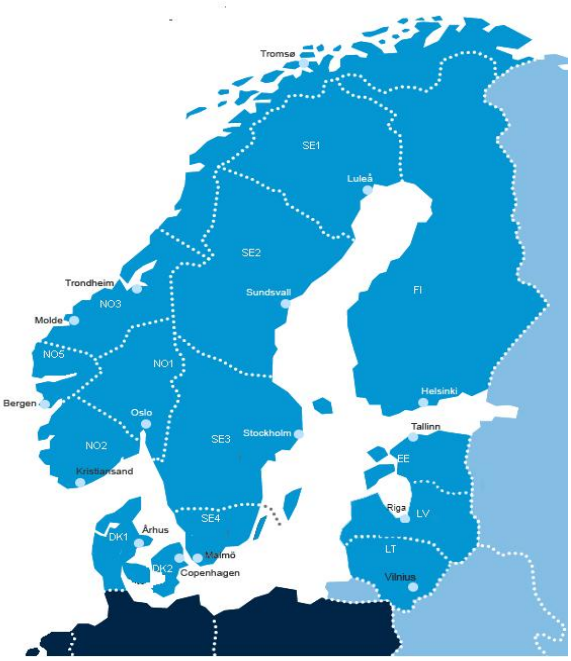

Figure 2.2. Nord Pool Spot price areas (source: Nord Pool Spot, 2013c).

An insufficient grid capacity is an obstacle for a uniform price for the whole Nordic region. An area price is formed on the basis of the demand and supply curves aggregated for the specific bidding area taking into account the transmission capacity between different bidding areas.

For the sake of simplicity, the formation of the area price in a market composed of two market areas is considered. The principle is the same for the actual fifteen bidding areas in the Nordic electricity market. In Figure 2.3, area level supply/demand curves for two areas are shown.

There is large overproduction in area A and short supply in area B when the electricity price is equal to the system price. If the amount of required electricity import to area B from area A is more than the transmission capacity, it is not possible to completely meet the overdemand in area B. In this case, the supply curve (area B) is transferred the amount of the transmission capacity to the right. Area price is read on the vertical axis at the intersection of the demand curve and the new supply curve. As a result, the price in area B is higher than the system price. In the overproduction area A, the situation is similar. If the amount of desirable export is over the transmission capacity, the area price for area A is set below the system price. The import to area with a production deficit equals the export from the area with excess supply.

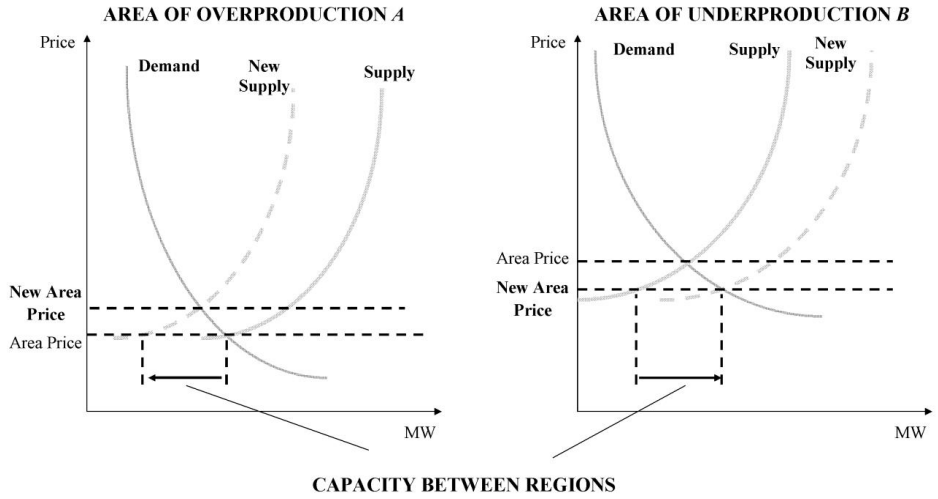

Figure 2.3. Formation of the area price in a two-area market.

If the transmission flow in the system price equilibrium does not exceed the available physical transmission capacity, the area prices are equal to the system price. The Finnish day-ahead area prices are equal to the day-ahead system prices in most cases over the period 1999–2013 (see Figure 2.4).

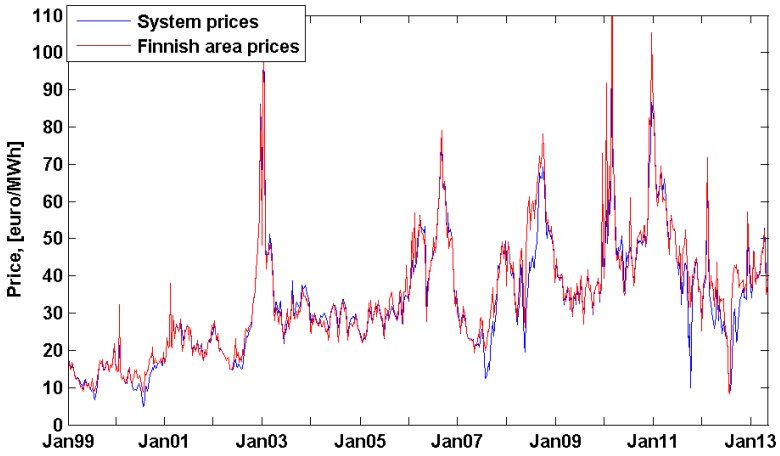

Figure 2.4. System prices versus area prices (in Finland) over the period 1999-2013 (weekly averages) (source: Nord Pool Spot, 2013d).

2.3.4 Elbas market

Some of the market and physical processes taking place up to the physical delivery after the Elspot market results have been settled should be considered in more detail. The time period between the physical delivery hour and the Elspot price settlement is long (36 hours at the most). There are many factors causing a change in the consumption and the generation situation, and thus, a market player needs an opportunity of trading during these hours. The Elbas market is an intraday continuous real-time physical market for electric power trading 24 hours a day, 365 days a year. The Elbas market is used to match the supply and demand up to one hour prior to the delivery in the case of unexpected situations or changes after the Elspot market trading.

2.3.5 Regulation power market

The regulating or balance power market is a tool for the Nordic transmission system operators (TSOs) to maintain the system balance between electricity production and consumption in real time. The balance between electricity production and consumption is described by the power system frequency. With the help of the regulating power market, a system operator can adjust the production or load based on the operational situation whenever necessary. There are two types of participants in the balance market. The first one is the active participants, the second one is the passive participants.

The active participants are producers or consumers who have an opportunity to regulate their generation or consumption in case of a request from the TSOs. There are some requirements for the active participants who operate in the balance market for the regulation of generation or consumption.

The holders of production or loads have an opportunity to submit bids for the regulating power market. The volumes of the bids are based on the holder's capacity that can be regulated. The balance providers get a right to participate in the regulating power market according to the balance service agreement. Other holders of capacity can also participate in the regulating market through their balance provider or by signing a separate regulating power market agreement with the TSO. There is a limit for the volume that is given in the bids and the responding time for regulation. The regulating bids shall be submitted to the TSO no later than 30 minutes before the operational hour. The minimum volume of the regulating bid is 10 MW, which should be implemented in 10 minutes after the request. In other words, prior to maintaining the physical balance, that is, balance regulation, the TSO regularly accepts bids, in other words, volume (power in MW) and price, from balance providers who are willing to quickly (within 10 minutes) increase or decrease their level of production or consumption (Fingrid, 2013a).

The regulation price is determined in accordance with the most expensive measure taken during upward regulation (the balance service purchases electricity), or the cheapest measure taken during downward regulation (the balance service sells electricity) applied during the hour. In other words, the up-regulation price is formed

based on the price of the most expensive up-regulation used, but at the least the price for the price area. All balance providers who were requested by the TSO for up-regulation during the hour obtain a price for the agreed energy in accordance with the up-regulation price. The down-regulation price is formed based on the price of the cheapest down-regulation used, but at the most the price for the price area. All balance providers who were requested by the TSO for down-regulation during the hour pay the down-regulation price for the agreed energy. The final regulation price applies to all balance providers who were selected to regulate the balance upwards or downwards.

2.3.6 Financial market

The Nordic financial market allows trading of financial contracts such as forward and futures with delivery periods up to six years in advance. None of these contracts entails physical delivery, and they are all settled in cash against the system price in the day-ahead market.

The system price plays a key role in the Nordic financial market. The majority of the standard financial contracts are settled by comparing the average system price for the period in question with the hedge price stated in the contract. There is mutual insurance in alliance to obligations that both parties have taken out. The difference in prices is multiplied by the volume in the contract, and this amount of money is transferred between the parties of the financial contract. However, not all financial contracts are settled against the system price, but there are also financial contracts with reference to the specific area prices.

2.4 Electricity demand

The total energy consumption in the Nordic countries can be divided into several user groups. The main groups are industry, housing, transport, and agriculture. Figure 2.5a introduces the structure of electricity consumption in the Nordic market in 2010, when the total energy consumption was 1 177 TWh, which is equal to about 8% of the energy consumption in the EU-27 (International Energy Agency, 2012).

Each consumer group can be characterized by its own demand profile, the shape of which typically slowly varies over time. The most stable electricity demand is caused by the energy-intensive industry sector. The reason for this is that the industrial plants operate continuously throughout the year with the exception of short interruptions. The electricity demand of the household sector instead is not stable through the year. As winters are often cold in the Nordic area, a household's electricity consumption is notably higher in winter when electric heating is widely used. In summer, the household demand for electricity is rather low as summers in the Nordic region are mild, and consequently, there is little need for air-conditioning. For the sake of visibility, Figures 2.6 and 2.7 present the relation between prices, total electricity consumption and opposite average values of temperature in the Nordic region since there is an explicit

negative correlation between temperature and prices, and temperature and electricity consumption.

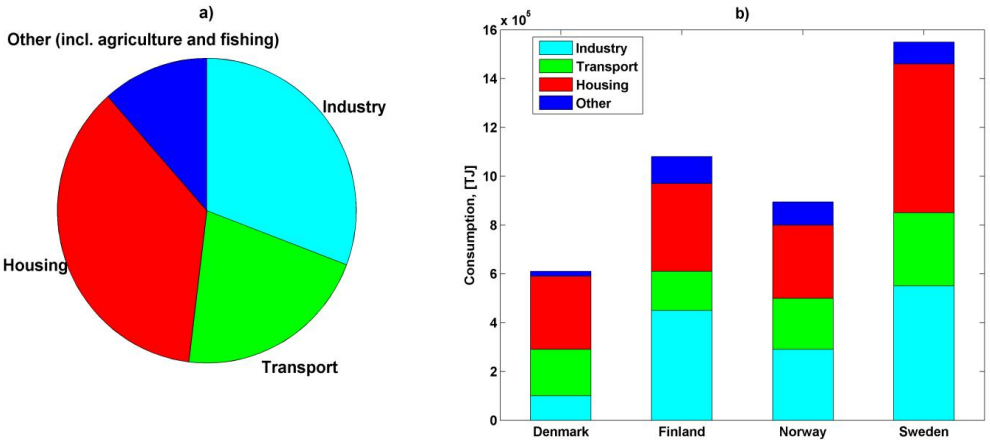

Figure 2.5. a) Nordic energy consumption by sector, 2010; b) Nordic energy consumption by sector and country, 2010 (source: International Energy Agency, 2012).

The public sector is mostly composed of transport and services, and its demand is significantly higher on weekdays compared with weekends. Electricity consumption of this sector decreases considerably during holidays.

The Nordic electricity market is presented by the electricity markets of Norway, Sweden, Finland, and Denmark. Each of these countries has quite similar demand characteristics (see Figure 2.5b). In Finland and Sweden, the forest-based industry is highly important. Metal manufacturing is of particular importance in Norway. The cold climate, combined with a history of low-cost and easy access to electricity, has resulted in high rates of electricity consumption for heating, particularly in Norway, Sweden, and Finland.

Despite having a relatively decarbonized electricity supply, the Nordic region has slightly higher per capita greenhouse gas emissions than other industrialized countries in Europe and Asia. This is due in part to the cold climate and prevalence of energy-intensive industry. The Nordic countries have set ambitious targets for emissions reductions by 2050.

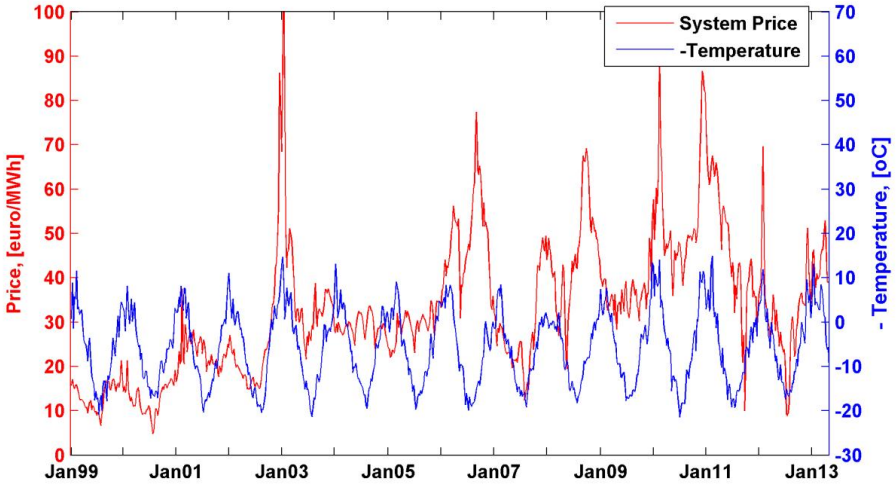

Figure 2.6. Nord Pool Spot day-ahead system prices versus temperature over the period 1999–2013 (source: Weather Underground, 2013; Nord Pool Spot, 2013d).

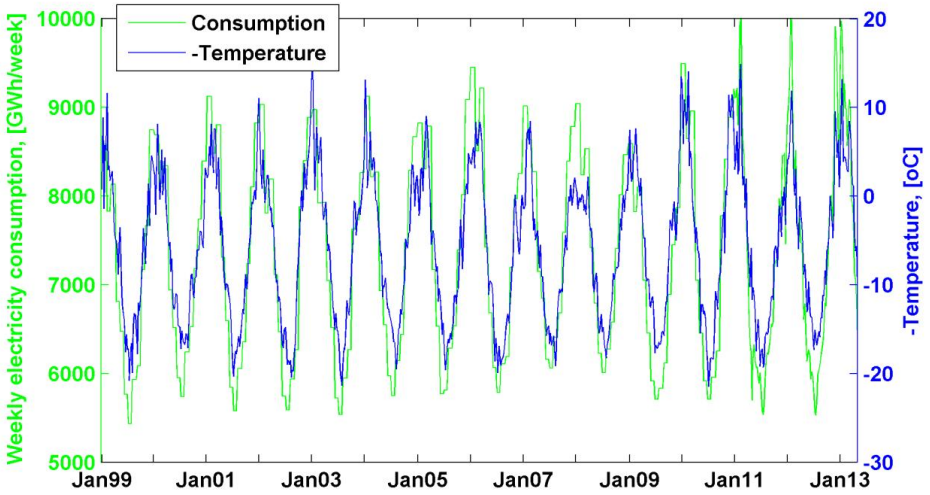

Figure 2.7. Total electricity consumption in the Nordic region versus temperature over the period 1999–2013 (source: Weather Underground, 2013; Nord Pool Spot, 2013d).

2.5 Electricity supply

Hydropower, nuclear power, conventional condensing power, CHP, and wind power may be considered the most important forms of electricity generation in the Nordic region.

A third of the energy supply in the Nordic region comes from renewable sources. The largest of these are biomass and waste, which are used to generate electricity, heat, and transport fuels in Sweden, Finland, and Denmark (see Figure 2.8a). Renewable electricity in the region is also generated from hydropower in Norway, as well as a growing share of wind power. With nuclear power in Sweden and Finland, almost half of the region's energy is CO2-free. Oil is still the largest single energy source, because of its central role as a transport fuel.

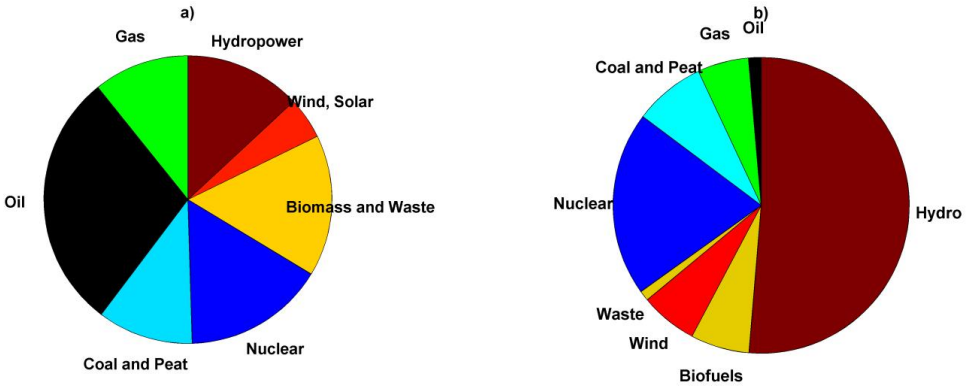

Figure 2.8. a) Nordic total primary energy supply 2011; b) Nordic electricity production 2011 (source: International Energy Agency, 2012).

As a whole, the Nordic electrical system is hydro dominant. More than a half of the overall electricity consumption is covered with hydropower generation (see Figure 2.8 b). The amount of hydropower fluctuates from year to year depending on the annual inflow that is determined by precipitation and the amount of melting snow. So, the annual energy available in the Nordic electrical system varies with the fluctuation of the annual water level.

Biomass is burned in CHPs across Finland and Sweden, while Denmark has the highest share of wind power in the world (see Figure 2.9).

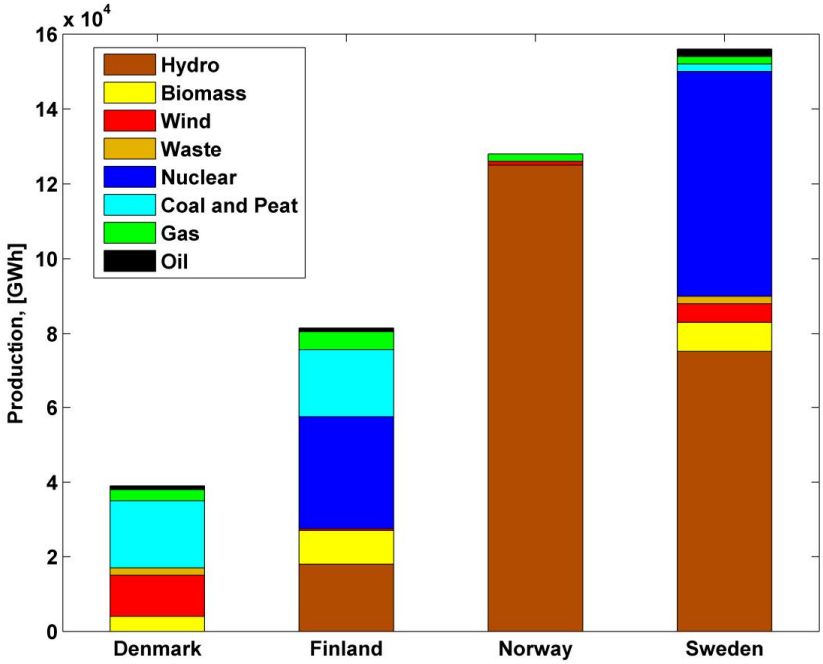

Figure 2.9. Electricity production 2011 (source: International Energy Agency, 2012).

Individually, the Nordic countries have very different, but complementary electricity mixes. This is made possible by the common Nordic grid connecting Norway, Sweden, Finland, and Denmark.

Since over a half of the generation capacity in the Nordic market is based on hydro units, a factor representing hydro reservoir in the area can be considered to determine the electricity price. In the long-run, however, electricity prices are more correlated with the variation in the hydro reservoir content than the absolute value of this variable (Jabłońska et al., 2012). The time series of both the day-ahead system price and the deviation of the Scandinavian hydrological situation from normal are plotted in Figure 2.10. The deviation is calculated as the difference between the mean value indicated as the average between the minimum and maximum possible hydro storage over the last 10-year history and the hydrological situation in a given week. The Nordic market has shown that the deviations of water levels from normal have been clearly reflected in the electricity day-ahead prices till 2005 when the emissions trading of the EU was introduced.

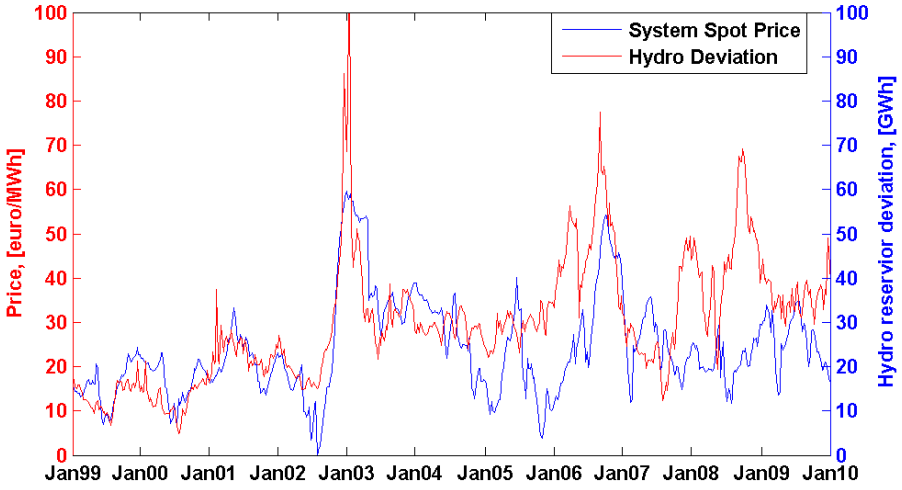

Figure 2.10. System prices versus deviation of the hydrological situation over the period 1999-2010 (source: Nord Pool Spot, 2013d).

3 Classical approaches to the modelling and forecasting of electricity prices

This chapter reviews a number of classical models and their application to the Finnish day-ahead electricity price behavior simulation and forecasting. In particular, Section 3.1 gives a basic statistic of prices over the last decade. Section 3.2 introduces techniques to define spike samples within a given series. Sections 3.3–3.4 discuss deterministic factors that have an impact on day-ahead electricity prices and propose a multivariate linear regression model with varying parameter estimates. Section 3.5 presents details and application of ARMA-based models. In Section 3.6, the mean-reverting Ornstein-Uhlenbeck model is presented, with both white and colored noise. ARMA-based and mean-reverting models both enhanced with a regime-switching technique are presented in Section 3.7.

3.1 Basic statistics of the Finnish day-ahead electricity prices

The Finnish day-ahead electricity prices over the period from 1 Jan 1999 to 31 Dec 2010 are illustrated in Figure 3.1 a. A first look to Figure 3.1a reveals a quite erratic behavior of the day-ahead prices. The series is clearly nonstationary, that is, its mean value does not remain constant over time. The price log-return series is used to get stationarity and based upon the following formula

$$r_h = \ln \frac{X_h}{X_{h-1}} \tag{3.1}$$

where r_h is return for any time h, X_h is the price value at moment h, X_{h-1} is the price value at moment $h-1$. The variance in the series is not constant, which is clearly seen in Figure 3.1b representing the price log-returns. This feature is called heteroscedasticity.

Both the original prices and the price log-returns have evident spikes and mean reversion characteristics. The presence of spikes and mean-reversion is generally explained by the use of expensive generators entering the market when the demand increases (see Figure 2.1). Similarly, a decrease in demand will cause the prices to decrease when expensive generators leave the market.

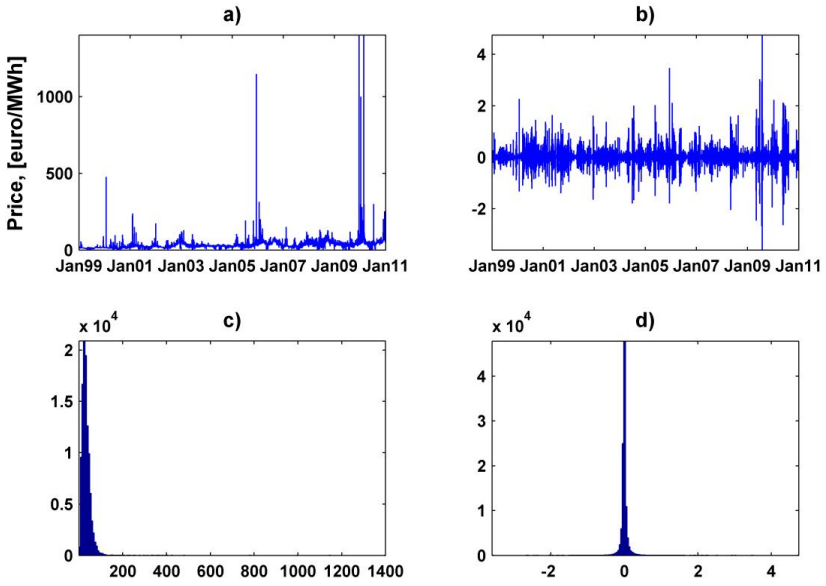

Figure 3.1. a) Original prices; b) Price log-returns; c) Histogram of the original prices; d) Histogram of the price log-returns.

The values of the most important distribution parameters of both the price and log-return series are collected in Table 3.1. With a mean value of 32.55 euro/MWh, the original price series reached maximum and minimum values of 1400.1 euro/MWh and 0 euro/MWh, respectively, during the sample period. This shows a huge spread of magnitudes over the given sample period. On the other hand, the returns seem to be of a relatively small range when compared with the prices, but this is a result of logarithmic operation. The prices for the winter and fall seasons show very similar mean values which, in turn, are higher than the price mean values for the spring and summer seasons. The standard deviations of sample prices show that the prices of the winter season are at least twice as volatile as those of the three other seasons.

In general, comparing the given probability distributions of both the prices and the price log-returns with the normal probability distribution, it is easily seen that neither the prices nor the log-returns follow the normal distribution. The original prices and price log-returns series show very high leptokurtosis (see Figure 3.1c, 3.1d). It indicates that extremely low and high values of the series have a much higher probability of

occurrence than those values that are due to a normal distribution with the same variance. The degree of asymmetry of the original prices and the price log-returns is not as high as the leptokurtosis. Both the series are positively skewed.

Table 3.1. Basic statistics of the prices and the price log-returns.

	Original prices, [euro/MWh]					Price log returns
	All seasons	Winter	Spring	Summer	Fall	All seasons
Mean	32.95	36.89	28.49	31.35	35.16	0.00
Std	22.61	35.77	13.55	17.18	16.01	0.11
Maximum	1400.11	1400.11	149.52	300.04	199.90	4.74
Minimum	0.00	3.87	0.28	0.00	2.19	-3.60
Skewness	18.87	18.70	0.79	1.64	0.95	1.79
Kurtosis	940.98	589.89	4.24	14.80	5.01	120.39

The interdependencies in the price series are verified. The autocorrelation functions (ACF) and the partial autocorrelation functions (PACF) of both the original prices and the price log-returns are plotted (see Figure 3.2).

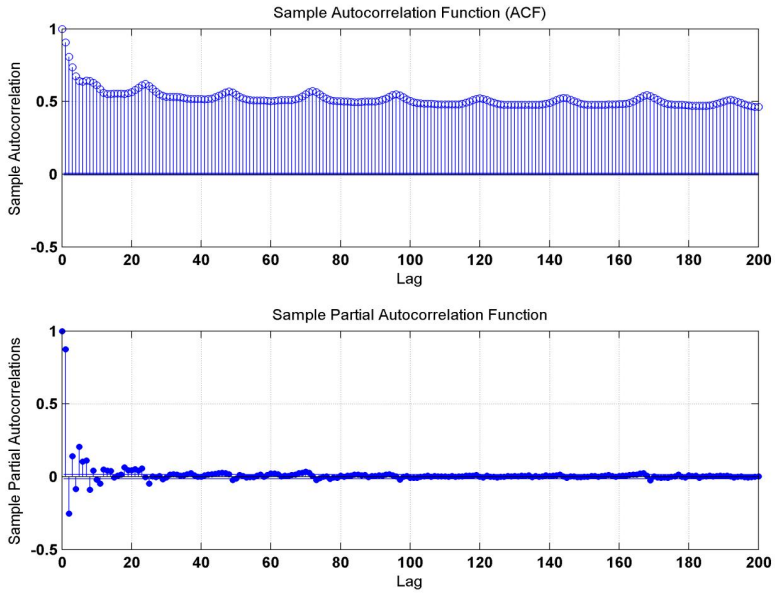

Figure 3.2. ACF (top) and PACF (bottom) of the prices.

The ACF of the prices seem to die out very slowly, whereas the PACF plot reveals a very significant spike at lag 1. The price log-returns are significantly positively

autocorrelated at several lags multiple of 24 indicating strong seasonal patterns (see Figure 3.3).

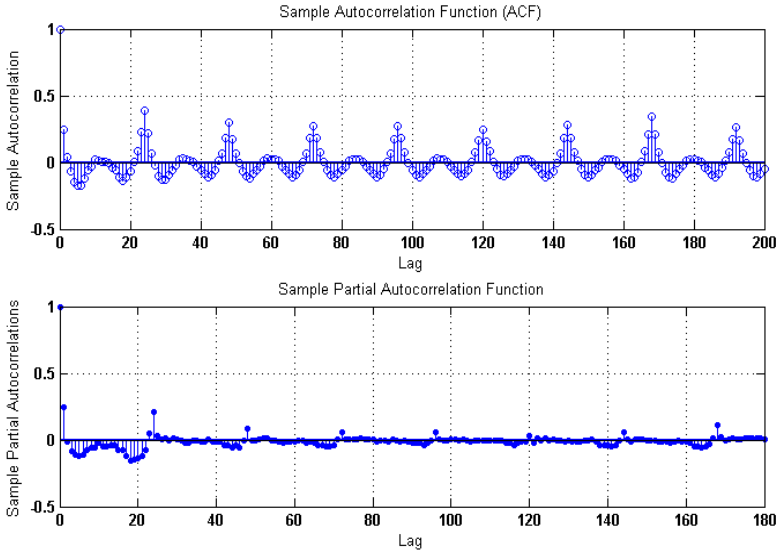

Figure 3.3. ACF (top) and PACF (bottom) for the price log-returns.

3.2 Electricity price spikes

For the purposes of the study of price spikes, a spike definition is formulated. A price spike can be defined as a price that surpasses a specified threshold. However, the main questions are how high the threshold should be and whether the threshold should have a fixed or time-dependent value. Specification of the threshold characteristics is a challenging task. Some authors suggest the use of fixed log-price change thresholds (Bierbraurer et al., 2004), a varying original or log-price range threshold (Cartea and Figueroa, 2005), or a fixed original price range threshold (Amjady and Keynia, 2010).

It is advisable to use a varying threshold value since the very volatile character of electricity prices usually requires the use of a varying threshold instead of one global value to cut off global outliers. Two different approaches to define spikes within a given series are applied within the study:

- A varying threshold is iteratively calculated. The whole given series is filtered with values that are out of the range defined by the mean value μ and the *n* time standard deviation σ of the whole given series at the specific iteration as [μ-*n*·σ μ+*n*·σ]. On the second iteration, the corresponding mean value and standard deviation of the remaining series is again calculated: those values that are now out of the range are

filtered again. The process is repeated until no further values can be filtered. Then, a spike value is calculated as a difference between the corresponding values of original and adjusted series and considered as upper or lower spikes.

- A time-varying threshold is calculated as was proposed in one of the previous studies (Jabłońska, 2008). Further, a spike is understood as an observation that is out of the range defined by the mean value μ and the n time standard deviation σ of the neighborhood data of the specific length w as $[\mu\text{-}n\cdot\sigma \;\; \mu\text{+}n\cdot\sigma]_w$. Here, the neighborhood data are understood as a set of observations before and after the given observation. Therefore, very high and very low values of the given series can be indicated and considered as upper or lower spikes, respectively. Then, a spike value is calculated as a difference between the given observation (defined as a spike) and the mean value μ of the corresponding neighborhood interval of length w.

Since the importance (i.e. economic impact) of upper price spikes for market participants is much higher than that of lower spikes, in the further study, only upper price spikes are considered with a few exceptions (see Section 3.7).

Figure 3.4 shows the results obtained when the two above-mentioned spike-defining approaches are used. As an example, upper price spikes are extracted given n = 3 and w = 6 months (4380 hours). The clustering character of the price spikes is visible.

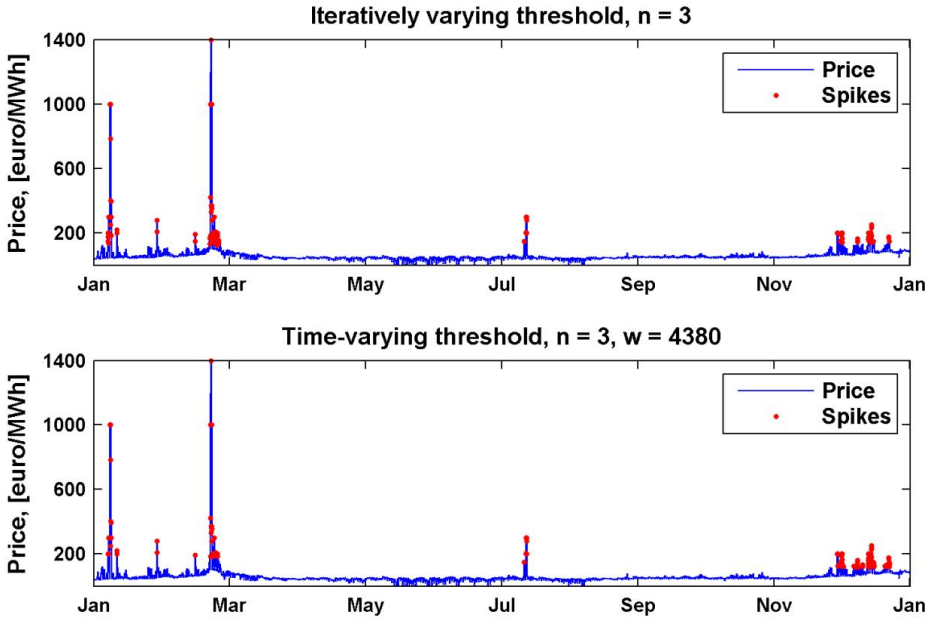

Figure 3.4. Spike samples extracted from the original hourly prices of the year 2010 when iterative (top) and time-varying (bottom) thresholds are used.

For both approaches, the spike size distributions are constructed and plotted in Figure 3.5. Moreover, the empirical normalized histograms are compared with an exponential distribution having a parameter λ (red curve in Figure 3.5) equal to the mean value of the extracted spikes. As can be seen, the magnitude of spikes can be roughly approximated by an exponential distribution.

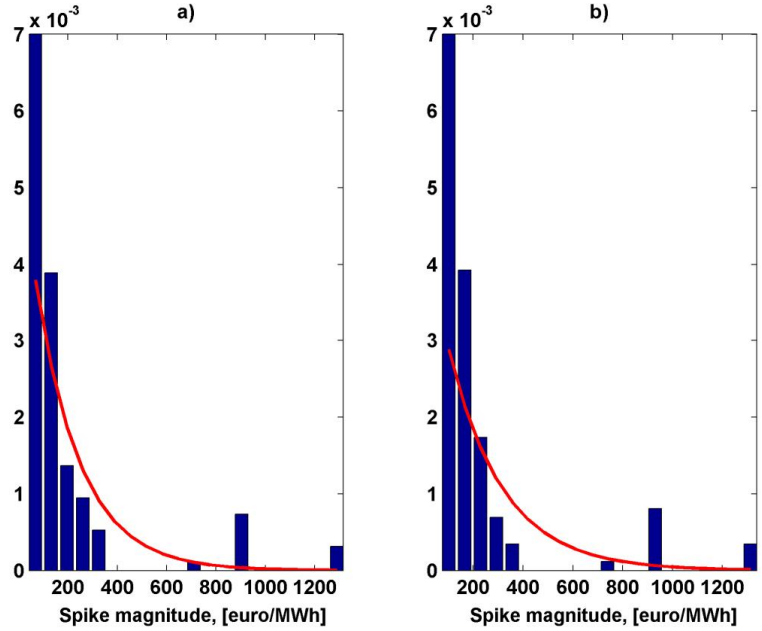

Figure 3.5. Distribution of spike magnitude in the original hourly prices of the year 2010 obtained by the approaches using a) iterative threshold given n = 3 and b) time-varying threshold given n = 3 and w = 4380 hours.

3.3 Deterministic factors

Prices in the electricity market are highly volatile but are not purely stochastic and, therefore, can be explained, at least partly, by background information. As mentioned, electricity prices are influenced by many factors, such as historical prices, electricity demand, weather conditions, imports, generation outages, and operational reserves (Catalão, 2007). Some of the factors are more important than others.

3.3.1 Trend and seasonality

It is clearly seen that the Finnish day-ahead electricity prices exhibit different types of periodicity (see Section 3.1). They mostly arise as a result of an electricity demand change under specific climate conditions, such as temperature and the number of

daylight hours. Distinguishing between on-peak and off-peak electricity prices, or among prices corresponding to different time periods, such as seasons, is indeed important in power markets (Lucia and Schwarz, 2000). In some countries, also the supply side shows seasonal variations in output. Hydro units, for example, are heavily dependent on precipitation and snow melting, which varies from season to season. These seasonal fluctuations in demand and supply translate into seasonal behavior of electricity prices, and day-ahead electricity prices in particular (Weron, 2006).

As a result, the prices of the Finnish day-ahead energy market are known to have three main types of periodicity: daily, weekly, and annual. The first two types are due to regular variations in demand between different hours of the day (morning and evening peaks) and different days of the week (business day–weekend structure). The latter type of periodicity reveals long-term annual fluctuations; high prices in wintertime and low prices during the summer.

The functional relationship between these components can assume different forms. The classical decomposition in which a series is seen as the sum or product of trend, seasonal, and irregular components may be considered. Hence, there are two main options for a decomposition model:

$$X_h = T_h \cdot S_h \cdot I_h \text{ (multiplicative)} \tag{3.2}$$

or

$$X_h = T_h + S_h + I_h \text{ (additive)} \tag{3.3}$$

where X_h is the original data, T_h stands for the trend, and S_h and I_h for the seasonal and irregular components at moment h, respectively.

These approaches allow separation of the underlying patterns in the data series from the irregular components.

The above-mentioned deterministic components are modeled with the help of functions. The parameters of the functions are estimated from historical data. One of the approaches to account for both an annual price fluctuation and a trend can be given as a sinusoid with a linear trend (Weron, 2006):

$$S_{annual,h} = A \cdot \sin(\frac{2\pi}{8760}(h + B)) + C \cdot h + D \tag{3.4}$$

The estimates of the parameters A, B, C, and D at moment h can be obtained through a least squares fit (LSQ).

After removing the trend and the annual seasonality, the remaining series is used for the hourly/weekly seasonal cycles. A very simple method, which, in many cases, produces

good results consists of finding the "average" day (or any other detected period). The average may be taken to be the arithmetic mean or the median, that is, the 0.5 quantile. In the latter case, single large spikes do not influence the average very much as the median is more robust to outliers than the arithmetic mean (Weron, 2006). The idea is to rearrange the corresponding time series into a matrix with rows of length H (e.g. 24 element rows for a daily period detected in the hourly data; 168 element rows for a weekly period, etc.) and take the arithmetic mean or median of the data in each column. Then, for a given seasonality of length L, its respective seasonal indices are calculated as the following mean or median values. For the mean:

$$S_h = \mathrm{E}(S_h, S_{h+H}, S_{h+2H}, ..., S_{h+vH}) \tag{3.5}$$

where $h = 1,..,H$ and v is the number of all corresponding seasonal cycles within the total data horizon.

As mentioned above, intra-day and intra-week regular patterns are mainly determined by business activity, and they might change along the year following changes in the electricity demand across seasons. Figure 3.6 displays the average weekly seasonal cycle throughout sample prices over the period from 1 Jan 1999 until 31 Dec 2012. There is a clear difference in the shapes and mean levels between weekdays and weekends. The days of the week, in turn, are divided into weekdays, weekends, and holidays (see Figure 3.7). The following holidays in Finland are taken into account: Midsummer Day, Epiphany (6 Jan), May Day, Ascension Day, Christmas, New Year, and Independence Day (6 Dec).

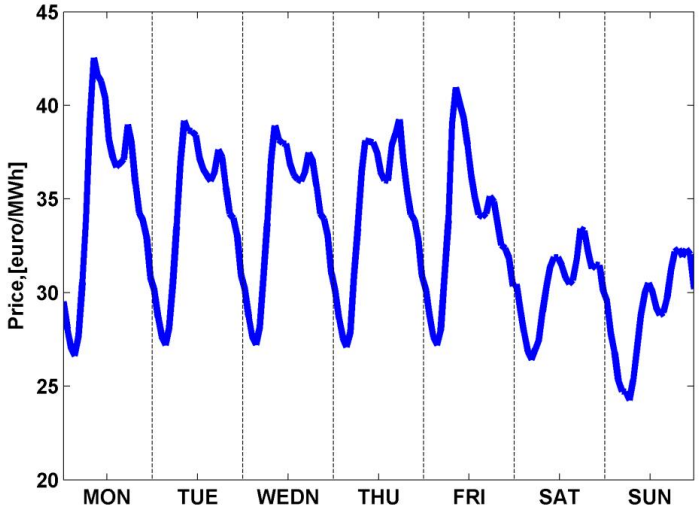

Figure 3.6. Hourly average pattern throughout the week for the period 1999–2012.

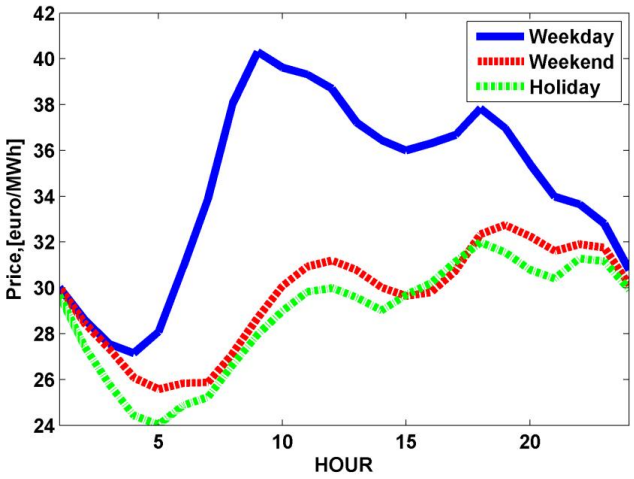

Figure 3.7. Hourly average patterns for working days, weekends, and holidays for the period 1999–2012.

The average price pattern for weekdays indicates higher prices during peak hours (08:00–12:00 and 17:00–20:00) especially over a winter season (see Figure 3.8). The shapes and mean values of weekend/holiday patterns are notably smoother and lower, respectively.

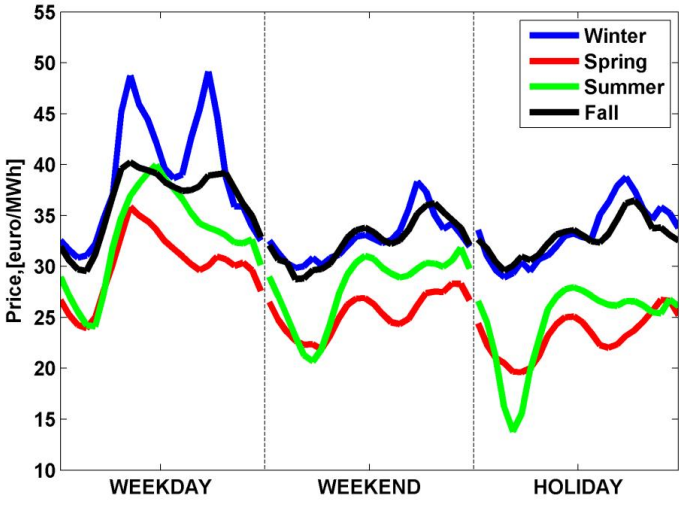

Figure 3.8. Hourly average patterns for weekdays, weekends, and holidays across seasons for the period 1999–2012.

The irregular component of a given price series can be calculated by extracting from the seasonally adjusted series. Afterwards, different methods can be applied for the irregular component of the electricity price process to simulate and forecast the given series.

3.3.2 External factors affecting the electricity prices in the Nordic region

It is widely known that the winter in the Nordic countries is often very cold. Electricity demand is higher when the atmospheric temperature rises or falls from a base "comfortable" level; temperature-dependent demand variations are more extreme if the humidity is higher, since moisture increases the heat retention capability of air (Willis, 2002). Atmospheric pressure variations generally cause air temperature variations, and as a consequence, demand variations. The effect of temperature and other weather-related variables can be incorporated in the electricity demand.

The level in the water reservoirs in the Nordic region is at minimum before the spring flood. Whole electricity demand cannot be covered by cheap hydroelectric power, and more expensive means of production must be used. Therefore, in order to understand the market state, instead of the overall generation capacity, the production capacity of different technologies such as hydro, thermal, and nuclear may be considered.

A part of the total electricity generation and consumption structures can be combined into nonbase electricity demand (Calmarza and de la Fuente, 2002). As can be seen in Figure 3.9, the hydro and nuclear power productions are rather constant and, therefore, show a low correlation with electricity prices in the short-run. The nonbase electricity demand is obtained by subtraction of the nuclear power and hydro power generation from the total electricity demand. The new explanatory variable is the part of the total electricity demand that is not covered by the base generation consisting of nuclear and hydro power generation.

Formally, in hourly resolution, the explanatory variable is defined as:

$$\textit{Nonbase Demand}(h) = \textit{System Demand}(h) - \textit{Hydro power}(h) - \textit{Nuclear power}(h) \quad (3.6)$$

where $h = 1,\ldots,24$ (hour of a day).

The value of the nonbase electricity demand is high if the total electricity demand is high or there is a lack of base generation (the values of hydro or/and nuclear power generation are low). The value of the nonbase electricity demand is low if the electricity demand is low or there is a high level of hydro and/or nuclear power generation. This explanatory variable covers all possible cases and presents an adequate hourly variation.

A transmission capacity excess or surplus is another important factor influencing the electricity price development. Two regimes of the system can be considered, where one

of the regimes is the regular regime and the other one, the nonregular regime, is the capacity-limited regime existing when the total available transmission capacity is not able to cover the required capacity.

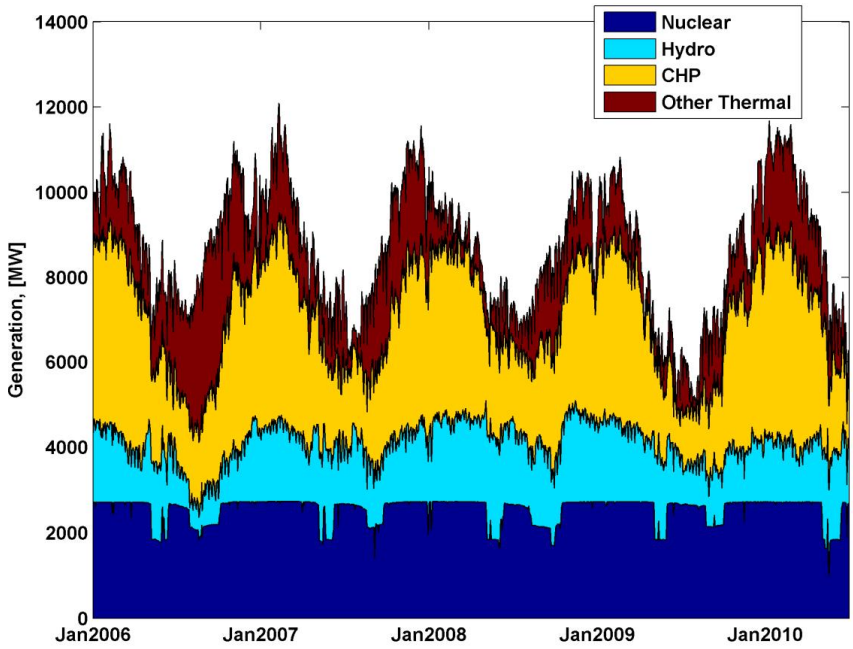

Figure 3.9. Structure of the Finnish energy production during the period 2006–2009.

Unit outage information, although clearly of importance, is not usually considered in the study because it is generally proprietary and not available to all market participants in real time.

3.4 Linear regression

To characterize the impact of selected independent variables (deterministic factors) on a dependent variable (price), a regression model can be considered. A simple linear regression model is selected for the problem of forecasting because of its capability of estimating time series. Its fast convergence (approach to results) using a limited number of available observations validates this choice. The three sets of variables involved in a regression model are:

- the dependent variable X to be predicted by the model,
- independent predictor(s) V, and

- unknown parameters β to be estimated.

Linear regression forecasting models are generally in the following format

$$X_h = V_h\beta + \varepsilon_h \tag{3.7}$$

where X_h is the value to be predicted at time h, $V_h = (1,v_{1h},\ldots, v_{kh})$ is a vector of k explanatory variables at time h, $\beta = (\beta_0, \beta_1,\ldots,\beta_k)$ is the vector of coefficients, and ε_h is a random error term at time h, $h = 1,\ldots,H$.

The most common way to find the model parameters (β) is to construct LSQ function $L(\beta)= \sum_{h=1}^{H}(X_h - (V_h \cdot \beta))^2$ considered as the difference between the forecasted and actual values to be minimized.

The standard assumptions of the time series regression are:

- $E(\varepsilon_h) = 0$,
- $Cov(\varepsilon_h,\varepsilon_t) = 0$, $0 \le t < h$, i.e., the residuals are not autocorrelated, and
- $Var(\varepsilon_h) = \sigma^2 < \infty$, i.e., the residuals are homoscedastic with a finite variance.

3.4.1 Forecast evaluation methods

Several evaluation criteria may be used to examine the accuracy of the results obtained from a forecasting model. Mean square error (MSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) were frequently considered to evaluate the performance of the forecast results in the literature.

The evaluation criteria are listed below:

$$MSE = \sum_{h=1}^{H}(\hat{X}_h - X_h)^2 / H \tag{3.8}$$

$$MAE = \sum_{h=1}^{H}|\hat{X}_h - X_h| / H \tag{3.9}$$

$$MAPE = \left(\sum_{h=1}^{H}\left[|\hat{X}_h - X_h| / X_h\right] / H\right)\cdot 100\% \tag{3.10}$$

Here $\hat{X}_h$ is the predicted value at time h, X_h is the actual value at time h, and H is the number of predictions.

The main disadvantage of the MAPE criteria is the adverse effect accruing from small actual values. If the actual value is small, Eq. (3.10) will contribute large terms to the

MAPE even if the difference between the actual and forecast values is small. Therefore, in some cases, the use of an adapted MAPE (AMAPE) is preferred:

$$AMAPE = \left(\sum_{h=1}^{H} \left[| \hat{X}_h - X_h | / (\sum_{h=1}^{H} X_h / H) \right] / H \right) \cdot 100\% \tag{3.11}$$

In the further study, MAPE and AMAPE criteria are generally used with one exception when MSE and MAE are also given (see Appendix D.5).

3.4.2 **Regression model building**

A significant relationship between two or more variables may simply mean that they are following the same trend without any further underlying relationship between them, a phenomenon more commonly known as spurious correlation. Therefore, before estimating the desired model, the price series and the explanatory variables are initially detrended and deseasonalized through the series decomposition. When having the dependent and independent variables properly aligned, LSQ optimal regression model is estimated (Jabłońska et al., 2012).

As an example, the regression model is estimated using the Finnish day-ahead electricity prices over the period from 1 Jan 2006 to 31 Dec 2009 to forecast prices 24 hours ahead. Selected explanatory variables are the nonbase electricity demand and the total electricity net import in Finland (Fingrid, 2013b). Here, the actual values of independent variables are used.

It should be borne in mind that the price series have local trends since market conditions evolve with time and, hence, the use of a long training period may result in significant inaccuracies. Therefore, the fit of the regression model is preferred not to be done globally on the whole data set at once, but in a moving regression fashion. Such a forecasting strategy provides an opportunity to account for more local trends dependent on other variables not available for the study and, therefore, being still consistent after the series decomposition is done. On the other hand, when a very short training interval is used, the model may not capture essential features of the considered series. Hence, selection of an appropriate training interval for a forecasting model is a challenging task and usually depends on the price characteristics of a case market.

In the experiment, every day historical price data of a specific horizon are used to estimate the model parameters and project the resulting prices 24 hours ahead. The prices of three years over the period from 1 Jan 2007 to 31 Dec 2009 are used as a test set. Table 3.2 presents numerical results obtained from the regression models using different horizons of a moving training period. As can be seen from Table 3.2, the best forecasting performance corresponds to the regression model with a moving training period of two months and results in MAPE value of 14.65%. The price forecast path produced by the regression model with a moving training period of two months is presented against the realized prices over the test interval in Figure 3.10a. The

respective residual series are given. The moving regression parameter estimates for the normalized values of the nonbase demand and the net import are illustrated in Figures 3.10b, c.

Table 3.2. Residual statistics for moving regression forecasting.

Training interval length	Mean	Std	Skewness	Kurtosis	MAPE, [%]
One month	1.15	7.15	-0.80	16.13	18.82
Two months	0.69	7.81	-0.18	8.12	14.65
Six months	0.03	8.00	-0.26	10.77	16.98
One year	0.40	8.75	-0.78	9.12	21.28

As can be seen, the moving regression parameter estimates differ over the modeling period. The electricity prices rise since there are a high nonbase electricity demand and a lack of import of cheap electricity mainly transferred from Russia.

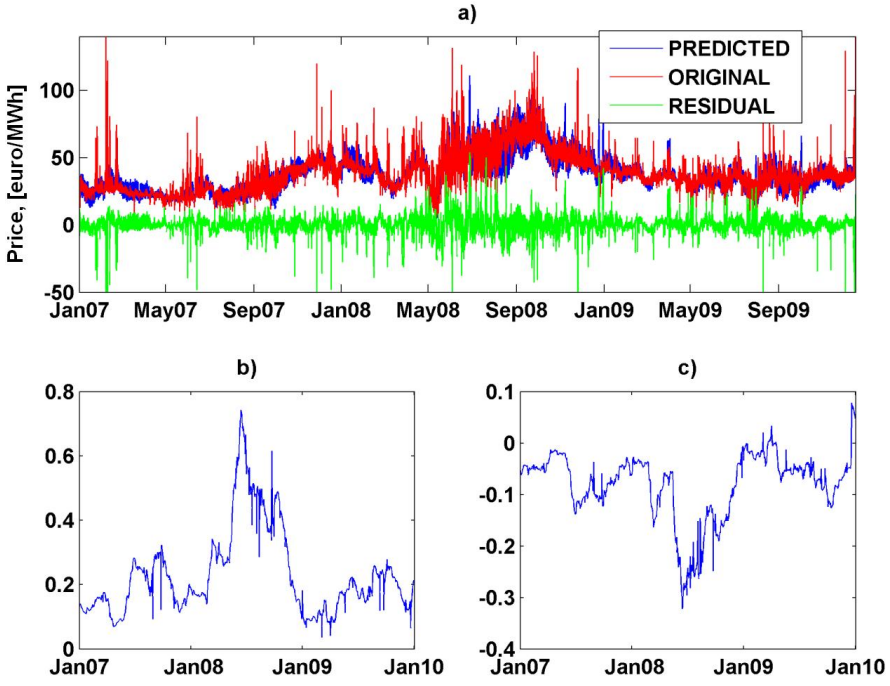

Figure 3.10. a) Moving regression forecasting; b) Nonbase demand; c) Net import.

3.4.3 **Summary**

Overall, a multivariate linear regression approach is not sufficient for modeling and forecasting actual electricity price behavior mainly because of its inability to capture the nonlinear characteristics of real prices. The residuals obtained from the regression model fit are prone to outliers and present a nonconstant mean level over the testing period (see Figure 3.10 a)).

3.5 The Box-Jenkins methodology

A time series is a sequence of data points at regular intervals (hourly, daily, monthly, annually). Provided that electricity prices are mean reverting, the Box-Jenkins methodology is used (Box and Jenkins, 1970). The Box-Jenkins model alters the time series to make it stationary. Besides a decomposition approach, the model is able to pick out trends from the time series itself, typically using autoregression (AR), moving average (MA), and seasonal differencing.

3.5.1 **ARMA model**

An autoregressive (AR) model attempts to model the current observation based on the previous realizations of a given process. An AR model of order p is denoted by AR(p). The AR(p) model for a stationary time series $\{X_h \,|h=1,2,\ldots,H\}$ is defined as:

$$X_h=\varphi_1 X_{h-1}+ \varphi_2 X_{h-2}+\ldots+ \varphi_p X_{h-p}+a_h \tag{3.12}$$

where φ are the AR coefficients, a_h is the error term $\{a_h\}\sim WN(0,\sigma^2)$, and a_h is uncorrelated with a_t for all $h<t$.

A moving average (MA) model is a linear regression of the current value of the series against the previous values of process errors. The MA model of order q is denoted by MA(q):

$$X_h=\theta_1 a_{h-1}+ \theta_2 a_{h-2}+\ldots+ \theta_q a_{h-q}+a_h \tag{3.13}$$

where θ are the MA coefficients and a_h is the error term $\{a_h\}\sim WN(0,\sigma^2)$.

ARMA is a combination of the AR and MA models. The model is then referred to as the ARMA(p,q), where p and q are the orders of the AR and MA models, respectively. The ARMA(p,q) is defined as:

$$X_h-\varphi_1 X_{h-1}-\varphi_2 X_{h-2}-\ldots- \varphi_p X_{h-p} = \theta_1 a_{h-1}+ \theta_2 a_{h-2}+\ldots+ \theta_q a_{h-q}+a_h \tag{3.14}$$

where all the terms have the previous meanings.

By using another time series that is known to covariate with the data under consideration, one can improve the prediction performance of the future values. The addition of an external input to a model is called using an exogenous variable in the time series modeling process. An ARMA(p,q) model with an exogenous factor is denoted by ARMAX(p,q,b):

$$X_h = \sum_{i=1}^{p} \varphi_i X_{h-i} + \sum_{i=1}^{q} \theta_i a_{h-i} + \sum_{i=1}^{b} \eta_i U_{h-i+1} + a_h \tag{3.15}$$

where b and η_i are referred to as the lag and the coefficient of the exogenous variable U_{h-i+1}.

If a given time series shows evidence of nonstationarity (trend, seasonality), initial differencing can be applied to remove the nonstationary characteristics. The model is, then, referred to as an integrated autoregressive moving average (ARIMA) or a seasonal ARIMA (SARIMA). The differenced time series is produced by subtracting the time series with lagged values from itself, the first-order lag operator is defined as

$$X_h - X_{h-1} = (1-B)X_h \tag{3.16}$$

where B is the backward shift operator $BX_h=X_{h-1}$

The differencing operator can be applied several times if necessary to obtain a stationary time series. When dealing with seasonal data it is preferred to use a seasonal differencing operator:

$$X_h - X_{h-S} = (1-B^S)X_h \tag{3.17}$$

where S is the period of the seasonal data.

For nonnegative integers d and D_S, the series X_h is a SARIMA(p,d,q)(P,D_S,Q) process with a period S if the differenced series $Y_h= (1-B)^d(1-B^S)^{Ds}X_h$ is an ARMA process defined by

$$\varphi_p(B)\Phi_P(B^S)Y_h = \theta_q(B)\theta_Q(B^S)a_h \tag{3.18}$$

where $\varphi_p(B)= (1- \varphi_1B- \varphi_2B^2-\ldots- \varphi_pB^p)$, $\Theta_q(B)= (1+ \Theta_1B+\Theta_2B^2+\ldots+\Theta_qB^q)$ are the regular, and $\Phi_P(B^S)= (1- \Phi_1B^S- \Phi_2B^{2S}-\ldots- \Phi_PB^{PS})$, $\Theta_Q(B^S)= (1+\Theta_1B^S+\Theta_2B^{2S}+\ldots+\Theta_QB^{QS})$, are the seasonal polynomials in B, p,q are regular orders of the AR and MA polynomials, P,Q are seasonal orders of the AR and MA polynomials, d is the number of regular differences, and D_S is the number of seasonal differences.

3.5.2 Preparing Box-Jenkins models

The Box-Jenkins approach uses an iterative model building strategy consisting of four steps. In the first step, the structure of the model is identified. Application of the ACF and PACF of the sample data is a basic tool to identify the order of the ARMA best model, which is then estimated by the maximum likelihood (ML) method in the second step. Description of the ML estimation method is given in Appendix A. The parameters of the model are estimated such that an overall measure of errors is minimized. The goodness-of-fit is tested on the estimated model residuals in the third step. If the model is not adequate, a new tentative model should be identified. Forecast future outcomes are obtained in the fourth step (Box and Jenkins, 1970).

When evaluating different models, it is important to be able to deduce which of the competing models best fits the data. The Akaike Information Criterion (AIC) is a measure that is used to compare models with each other; the AIC rewards models for a good fit and penalizes models for complexity. The AIC is defined as follows:

$$AIC = 2k + \left[\ln\left(\frac{RSS}{n_{obs}} \right) \right] \tag{3.19}$$

where k is the number of free parameters, n_{obs} is the total number of observations, and RSS is the residual sum of squares.

Bayesian information criterion (BIC) is closely related to the AIC but has a larger penalty term than in the AIC:

$$BIC = n_{obs} \ln(\sigma^2) + k \ln(n_{obs}) \tag{3.20}$$

where k and n_{obs} have the previous meaning and σ^2 is the error variance.

For both approaches, the aim is to choose the model order that provides the minimum values of AIC and BIC.

3.5.3 ARCH/GARCH modeling

ARMA-based models are used in many applied problems. The basic assumptions of the error terms of the models include zero mean and constant variance. In practice, the homoscedasticity assumption of constant variance sometimes does not hold. Such time series are called heteroscedastic. Thus, when the error terms are autocorrelated, the ML estimator of the ARMA model coefficients is no longer asymptotically unbiased and consistent. It is agreed that the electricity price time series present nonconstant deviations over time as demonstrated in Figure 3.1. Hence, the autoregressive conditional heteroscedasticity (ARCH) model was introduced (Engle, 1987). In this model, the conditional error variance σ^2 is considered as time dependent ARCH(r):

$$\sigma_h^2 = C + \sum_{i=1}^{r} \alpha_i a_{h-i}^2, \quad \text{at time } h, h = 1,2,3,...,H \tag{3.21}$$

where $a_h = \varepsilon_h \sigma^2$ is an error term produced by ARMA at time h, $\varepsilon_h = N(0,1)$ and C is a variance constant.

As ε_h is white noise, which is assumed to be normally distributed, a_h will also be normally distributed with a zero mean and the variance σ^2. In practical applications, the current variance sometimes appears to be dependent not only on past squared disturbances, but also on the past variance of the errors. Such an extended model was introduced and comes as a GARCH(r,s) model (Bollerslev, 1986):

$$\sigma_h^2 = C + \sum_{i=1}^{r} \alpha_i a_{h-i}^2 + \sum_{i=1}^{s} \beta_i \sigma_{h-i}^2, \quad \text{at time } h, h = 1,2,3,...,H \tag{3.22}$$

The application of a GARCH model is an iterative procedure similar to the ARMA procedure and includes iteratively: order determination, parameter estimation, and model diagnostic checking.

3.5.4 Price modeling and forecasting with SARIMA+GARCH

The process of ARMA-based model building is presented. The model adequacy and forecasting accuracy are evaluated with actual data from the Finnish day-ahead energy market of Nord Pool Spot. The main attention is focused on a particular period of hourly real-time electricity prices during the period from 16 Sep 2009 to 21 Nov 2009. The whole data set is divided into training (60 days) and testing (7 days) sets. Hence, the moving 24 hours ahead out-of-sample forecasts are generated from the estimated models over the testing period from 15 Nov 2009 to 21 Nov 2009.

A preliminary inspection of the ACF and PACF of the corresponding price log-returns indicates the presence of seasonality with respect to the hourly electricity prices (see Figure 3.11).

Besides the ACF/PACF analysis, the AIC and BIC values are estimated for a tentative model. Examples of model structures and their respective AIC/BIC values are presented in Table 3.3.

Given that the ARMA modeling process requires a stationary time series, nonseasonal first differencing and seasonal differencing of the twenty-fourth order are needed to render the electricity prices. A further examination of the ACF/PACF of the resulting stationary time series on electricity prices, and the AIC/BIC values of the corresponding residuals (see Table 3.3) indicated the following SARIMA specification:

$$(1-B)(1-B)^{24}(1-\varphi_1 B)(1-\Phi_1 B^{24} - \Phi_2 B^{168})price_h = (1+\theta_1 B)(1+\Theta_1 B^{24})a_h \tag{3.23}$$

The model diagnostics obtained for the SARIMA model given in Eq. (3.23) is reported in the second column of Table 3.4. The corresponding coefficients of the model parameter estimates and their standard errors are presented in Appendix B. The residuals are free of serial correlation based on the chi-square Ljung-Box Q-statistics. However, the chi-square test statistic for autoregressive conditional heteroscedasticity is statistically significant at the 5% level. The invertibility conditions for the respective nonseasonal and seasonal terms are satisfied.

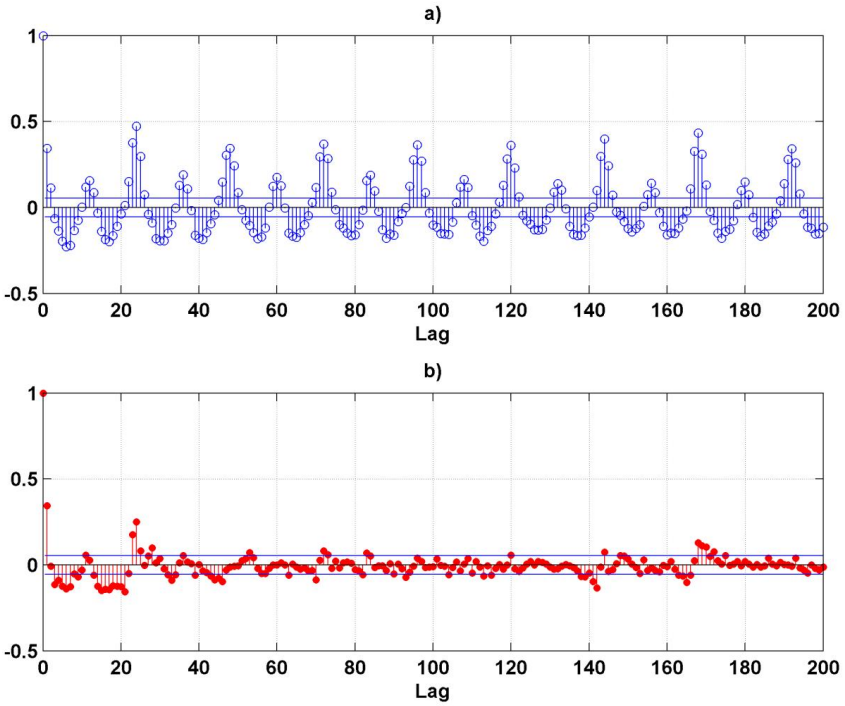

Figure 3.11. a) ACF of the price log-returns; b) PACF of the price log-returns.

Table 3.3. AIC/BIC results for ARMA based models estimated on the training set.

Model	AIC	BIC
ARMA(1,1)	4603	4616
ARIMA(1,1,1)	4602	4616
SARIMA(1,1,1)(1,1,1)$_{24}$	4240	4265
SARIMA(1,1,1)((1,7),1,1)$_{24}$	4195	4224
SARIMA(1,1,1)(1,1,(1,7))$_{24}$	4235	4264
SARIMA(1,1,1)((1,7),1,(1,7))$_{24}$	4196	4230

To recognize the presence of the autoregressive conditional heteroscedasticity in the residuals, a SARIMA+GARCH model is estimated. The AIC/BIC values are compared for an extensive range of different SARIMA+GARCH models. Examples of model structures and their corresponding AIC/BIC values are given in Table 3.5.

Table 3.4. Model diagnostics and MAPE values for SARIMA and SARIMA+GARCH models estimated for original and adjusted price series.

Model diagnostics:	Original price series		Adjusted price series	
	SARIMA (1,1,1) $((1,7),1,1)_{24}$	SARIMA(1,1,0) $(7,1,1)_{24}$ + GARCH(1,1)	SARIMA(1,1,1) $((1,7),1,1)_{24}$	SARIMA $(1,1,0)(7,1,1)_{24}$ +GARCH(1,1)
LB_Q	59.21	64.32	66.08	66.79
	[0.11]	[0.06]	[0.06]	[0.06]
ARCH	156.17	97.12	104.34	66.79
	[0.00]	[0.00]	[0.00]	[0.06]
MAPE, [%]	5.83	4.62	4.05	3.65

Notes: Probability values are reported in brackets. LB_Q is the Ljung-Box Q-statistic to test for serial correlation in the residuals. ARCH tests for autoregressive conditional heteroscedasticity in the residuals.

The methodology results in the following SARIMA+GARCH model:

$$(1-B)(1-B)^{24}(1-\varphi_1 B)(1-\Phi_1 B^{24}-\Phi_2 B^{168})price_h = (1+\Theta_1 B^{24})a_h \tag{3.24}$$

$$\sigma_h^2 = C + \sum_{i=1}^{1} \alpha_i a_{h-i}^2 + \sum_{j=1}^{1} \beta_i \sigma_{h-j}^2, \text{ at time } h, h = 1,2,3,...,H \tag{3.25}$$

The third column of Table 3.4 reports the model diagnostics obtained for the SARIMA+GARCH given in Eqs.(3.24)–(3.25). The residuals are free of both serial correlation but still indicate presence of the heteroscedasticity at the 5% level. All the model parameter estimates are statistically significant at the 5% level (see Appendix B).

To limit the volatility of the given price series, electricity price spikes are extracted from the original price series with parameters w = 720 hours; n = 3 (see Section 3.2). The proposed SARIMA/SARIMA+GARCH models are estimated on the adjusted price series. The results of the models are reported in the fourth and fifth columns of Table 3.5. As can be seen, the residuals are free of both serial correlation and autoregressive conditional heteroscedasticity for the case of SARIMA+GARCH model estimated on the adjusted price series. The model parameter estimates are all statistically significant at the 5% level and presented in Appendix B.

In order to assess the ability of the models to capture the actual behavior of prices the forecasted price curves are presented against original ones (see Figure 3.12). The MAPE

values used to examine the forecasting performance of the models over the testing period are reported in Table 3.4.

Table 3.5. Obtained results of AIC/BIC for SARIMA+GARCH models estimated on the training set.

Model	AIC	BIC
SARIMA(0,1,0)(7,1,1)$_{24}$+GARCH(1,1)	3879	3898
SARIMA(1,1,1)((1,7),1,1)$_{24}$+GARCH(1,1)	3872	3901
SARIMA(1,1,0)((1,7),1,1)$_{24}$+GARCH(1,1)	3865	3889

Truncation of the spikes before application of the forecasting model helps to reduce the influence of such observations on the estimation of the method parameters. Such a strategy results in an improvement in the model forecasting performance over the testing period, which supports the previous studies (see Figure 3.12). However, this finding is mainly reasonable for the case when no spikes exist on the forecast period.

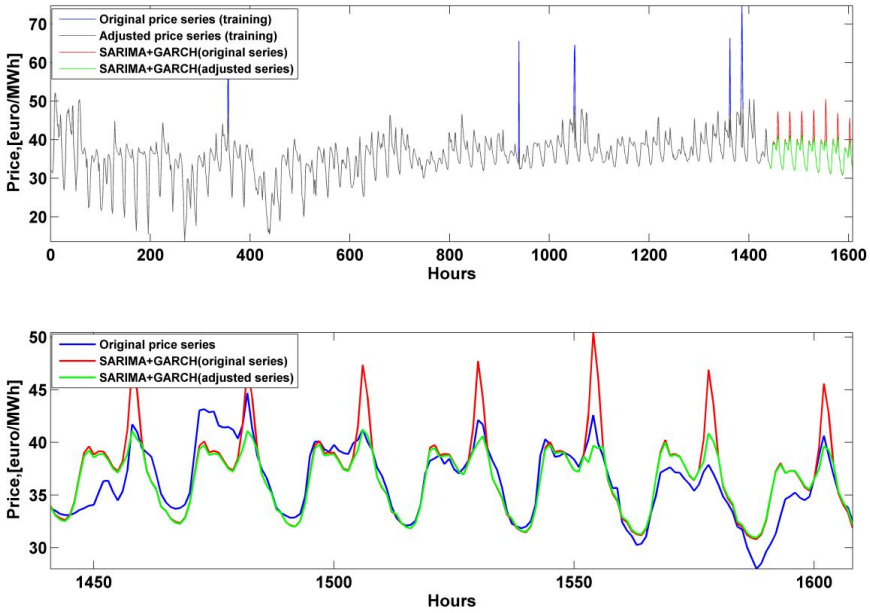

Figure 3.12. SARIMA+GARCH one-day ahead forecast over the period 15 Nov 2009 — 21 Nov 2009.

3.5.5 Summary

It is shown that accurate prediction of day-ahead electricity prices with (S)AR(I)MA/ (S)AR(I)MA+GARCH models is not generally possible because of the inability of the models to estimate high volatility and spike clustering presented in the original price time series. Therefore, to avoid an undesirable effect of the presence of spike samples in the training data set, those samples should be extracted from the corresponding set. Further, a possible approach to capture the actual price behavior would be a separate prediction of adjusted price series and spikes with the use of different forecasting engines.

3.6 Stochastic differential equations – Ornstein-Uhlenbeck process

One of the approaches to model electricity prices is based on stochastic modeling. A stochastic process is a family of random variables $X(h,\omega)$ of two variables $h \in H$, $\omega \in \Omega$ on a common probability space (Ω,F,P), which assumes real values and is P-measurable as a function of ω for a fixed h. The parameter h is interpreted as time, with H being a time interval and $X(h,\cdot)$ represents a random variable on the above probability space Ω, while $X(\cdot, \omega)$ is called a sample path or trajectory of the stochastic process.

3.6.1 Stochastic process

A stochastic process $(W_h)_{h\geq 0}$ is defined as Brownian motion (BM) if has the following characteristics:

- $W_0=0$, that is, BM starts at zero.
- $(W_h)_{h\geq 0}$ is a process with homogeneous and independent increments, i.e., distribution of future changes does not depend on past realizations.
- Any increment W_h-W_t is normally distributed with a mean zero and the variance h-t, $0 \leq t < h$, i.e., the variance increases linearly with the length of time interval.
- The paths of $(W_h)_{h\geq 0}$ are continuous but nowhere differentiable.

3.6.2 Ornstein-Uhlenbeck process

The mean-reversion (MR) process is one of the most applied stochastic processes to simulate electricity prices (Gibson and Schwartz, 1990; Hirsch, 2009; Möst and Keles , 2010). Therefore, it can be considered an alternative to the Box-Jenkins time series models. The MR process called Ornstein-Uhlenbeck (OU) (Uhlenbeck and Ornstein, 1930) can be formulated for the price changes with the following stochastic differential equation (SDE):

$$dX_h = k(\mu - X_h)dh + \sigma dW_h \tag{3.26}$$

The first term $k(\mu\text{-}X_h)$ of Eq. (3.26) describes the drift component. The parameter k determines the reversion rate of the stochastic process to its long-term mean μ. The essence of the mean-reversion concept for the case of a price time series is the assumption that any stochastic price fluctuations are temporary and the price will tend to move to the mean price over time. As mentioned above, in the electricity markets, the price fluctuations and the mean reversion are generally explained by entering expensive generators as a result of an extreme meteorological situation, power plant outages and transmission congestions.

The second term σdW_h corresponds to the standard Brownian motion. The stochastic driver is the Wiener process movement $dW_{h=}\ \varepsilon_h dh^{1/2}$, where ε_h is a standard normally distributed random variable.

3.6.3 Calibration of SDE

The SDE is solved by Euler discretization (Lari-Lavassani et al., 2001), applying Ito's Lemma with the following exact solution (Karatzas and Shreve, 2000):

$$X_{h+1} = X_h \cdot e^{-k\delta} + \mu\left(1 - e^{-k\delta}\right) + \sigma\sqrt{\frac{1 - e^{-2k\delta}}{2k}} \cdot \varepsilon_h,\ \ \varepsilon_h \sim N(0,1) \tag{3.27}$$

The substitutions $a=e^{-k\delta}$, $b=\mu(1-e^{-k\delta})$, $\sigma_\varepsilon = \sigma\sqrt{(1-e^{-2k\delta})/2k}$ and lead to the equation $X_{h+1} = aX_h+b+\varepsilon_h$, $\varepsilon_h \sim N(\mu_\varepsilon, \sigma_\varepsilon)$ whereas δ is the time difference between h and $h+1$, here one hour.

The parameters a,b,σ_ε are determined by ML or LSQ. The resubstitution of the parameters a,b,σ_ε results in the original parameters of the exact solution k,μ,σ. With the help of the estimated parameters, the exact solution of the SDE is applied to generate the price path.

3.6.4 OU process to simulate electricity prices

In the first step, prices are logarithmized and the price logs are passed to the simulation tool instead of the prices themselves. The logarithm is used as it limits the volatility and leads to a variance stabilization. Since the electricity prices display typical patterns, the models developed to describe the behavior of electricity prices should capture the deterministic components (trend, daily, weekly, and annual cycles) of electricity prices.

The deterministic patterns (daily, weekly, annual seasonality) are removed from the log-price series. The remaining stochastic component is then used to estimate the parameters of the corresponding stochastic process. Finally, the deterministic components are added to the simulated stochastic component, and then, the simulated price logs are retransformed receiving a simulated price path.

Model parameter estimates are calculated for the stochastic component extracted from the logarithmized Finnish day-ahead electricity prices of the years 2007–2009. At a closer inspection of Figure 3.13 it becomes evident that the simulated price path partly follows the actual series. Rather, this is a consequence of the excessive "jumpiness" of an optimal mean-reverting model. The residuals emerging from this optimal mean reverting model are normally distributed. Since an Ornstein-Uhlenbeck model is always normally distributed by definition, this property is transferred to the model residuals when there are frequent spikes in the simulated series that do not coincide with the spikes in the actual series (Naeem, 2009).

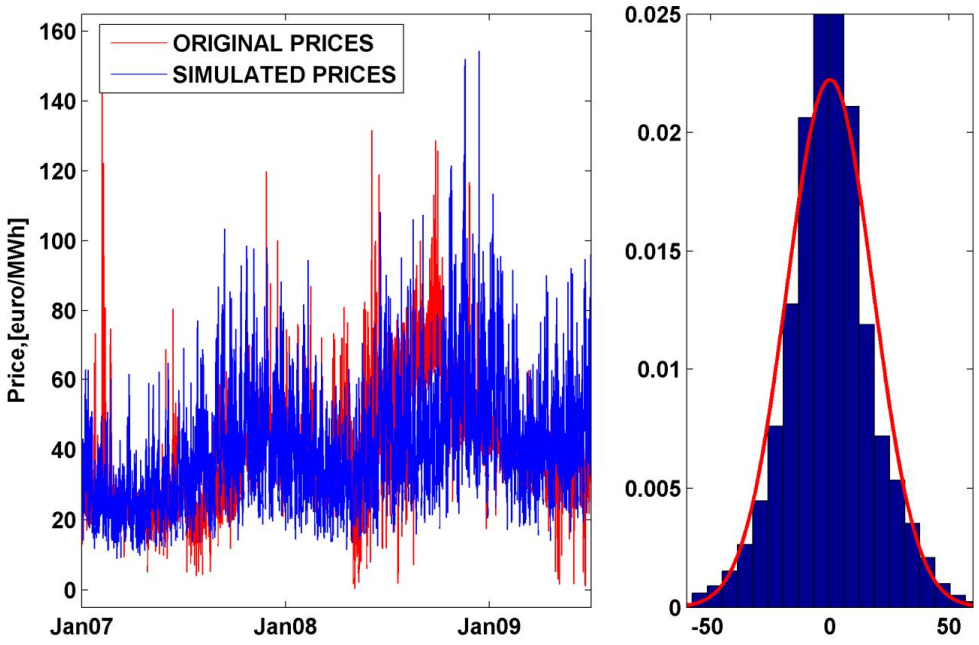

Figure 3.13. Ornstein-Uhlenbeck simulation (left) and normalized histogram of the model residuals with normal distribution (right).

Relevant statistics of the original and simulated prices are collected in Table 3.6. To achieve a more robust result, an expected value for the measurements is determined based on 50 simulations for the OU process.

It should be concluded that the conventional mean-reverting Ornstein-Uhlenbeck model, even when calibrated optimally with the actual electricity market prices, is not able to capture the statistical characteristics of the actual series.

Table 3.6. Basic statistics for original Finnish day-ahead electricity prices and price paths simulated by the Ornstein-Uhlenbeck process.

	Original prices, [euro/MWh]	Simulated prices, [euro/MWh]
Mean	39.67	40.01
Std	15.64	17.33
Skewness	0.86	1.22
Kurtosis	3.93	5.49
Maximum	150.30	162.62
Minimum	0.34	8.64

3.6.5 **OU process with colored noise**

The mean-reversion process driven by an exponential colored noise can be formulated for the price changes with the following SDE (Mtunya, 2010):

$$dX_h = k(\mu - X_h)dh + \sigma\zeta_h dh \tag{3.28}$$

The terms of Eq. (3.28) have the same meanings as in Eq. (3.26), ζ_h is an exponentially colored noise process generated to mimic the behavior of both the spikes and the usual volatility of the prices. The colored noise process ζ_h produces a sequence of correlated random variables $\zeta(h_1)$, $\zeta(h_2)$,… with the same standard deviation in each. Colored noise is a Gaussian process, and it is well known that this process can be completely described by their mean and covariance functions (Arnold, 1974).

The Ornstein-Uhlenbeck process is extended and repeatedly integrated to obtain the colored noise of the first and second orders forcing along the series:

$$d\zeta_1(h) = -\frac{1}{\tau}\zeta_1(h)dh + \alpha_1 dW_h \tag{3.29}$$

$$d\zeta_2(h) = -\frac{1}{\tau}\zeta_2(h)dh + \frac{1}{\tau}\alpha_2\zeta_1(h)dh \tag{3.30}$$

where τ is the correlation time for colored noise (in the case of hourly data τ = 24, indicating price hourly seasonality); α_1 and α_2 are the diffusion constants; W_h is a Wiener process with $dW_h \sim N(0,dh)$.

The system of Eqs. (3.29)–(3.30), with $\zeta(0)=0$ (i.e., starting with no noise) and $t<h$, has the following solutions:

$$\zeta_1(h) = \alpha_1 \int_0^h e^{-\frac{(h-t)}{\tau}} dW_t \tag{3.31}$$

$$\zeta_2(h) = \frac{1}{\tau}\alpha_1\alpha_2\int_0^h e^{-\frac{(h-t)}{\tau}}(h-t)dW_t \tag{3.32}$$

All the relevant process parameters are estimated by the ML methodology. The system of Eqs. (3.31)–(3.32) generates a stationary, zero-mean, correlated Gaussian process $\zeta_2(h)$. The generated colored noise process $\zeta_2(h)$ is applied to Eq. (3.28) to model the price. Therefore, the mean-reverting log-price equation is as

$$dX_h = k(\mu - X_h)dh + \sigma\zeta_2 dh \tag{3.33}$$

With the use of colored noise forces, the correlation of the noise terms that influence the price time series is modeled more accurately, and it becomes possible to take into account the spiking characteristics and volatility clustering of the prices.

3.6.6 OU process with colored noise to simulate electricity prices

Prices are logarithmized and deterministic patters are removed. The corresponding stochastic component of the price logs are passed to the simulation tool, deterministic patterns are added, and the simulated price logs are retransformed receiving a simulated price path. Simulation of the Finnish day-ahead electricity prices of the years 2007–2009 with the use of the MR process driven by an exponential colored noise is presented in Figure 3.14.

The relevant statistics of the original and simulated prices based on 50 simulations are collected in Table 3.7.

Table 3.7. Basic statistics for the original Finnish day-ahead electricity prices and price paths simulated by the Ornstein-Uhlenbeck process with colored noise.

	Original prices, [euro/MWh]	Simulated prices, [euro/MWh]
Mean	39.67	41.97
Std	15.64	21.80
Skewness	0.86	1.26
Kurtosis	3.93	5.26
Maximum	150.30	182.47
Minimum	0.34	5.89

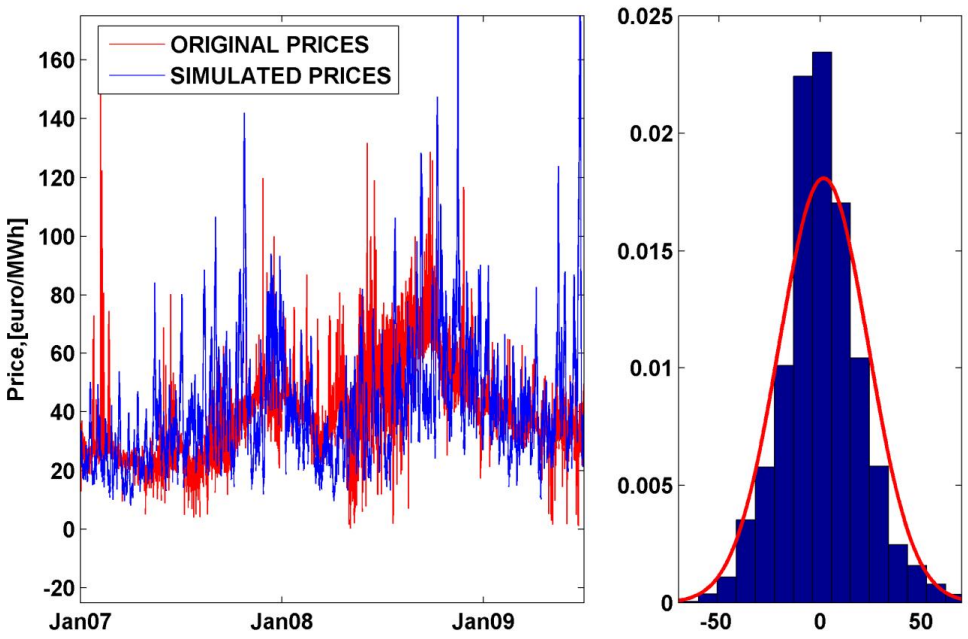

Figure 3.14. Ornstein-Uhlenbeck with colored noise simulation (left) and normalized histogram of the model residuals with normal distribution (right).

The process driven by colored noise produces prominent spike groups. However, the trajectory of the price path simulated by the process with colored noise partly captures the original price behavior. As can be seen in Figure 3.14, the spike groups are clustered and usually exist more often and for a longer time period than in actual case.

3.7 Regime-switching model

Different models based on MR, ARMA, and GARCH processes applied to the electricity price modeling and simulation are evaluated and compared.

As in the previous section, the simulation of electricity prices is formed on an extended modeling approach considering both stochastic and deterministic components of the price process derived from the Finnish day-ahead energy market of Nord Pool Spot. First, the deterministic components are modeled and removed from logarithmized historical price series. The resulting stochastic residuals are then used to estimate the parameters of each stochastic process.

As the presence of spikes is one of the main characteristics of electricity prices, a regime-switching approach is applied to distinguish the nonspiky and spiky behavior of prices. Both upper) and lower spikes of a given series are considered. Finally, deterministic patterns are added to the simulated stochastic component. The forecasting methodology is illustrated in Figure 3.15.

A regime-switching approach is implemented into the forecasting model to simulate the transition of prices between the normal and spike regimes. To combine the different regimes with a common approach, transition probabilities between the regimes and probabilities remaining in the same regime are calculated based on historical data. Therefore, if regime 1 is the normal regime, regime 2 is the upper jump regime, and regime 3 is the lower spike regime, respectively. Then, the matrix of transition would come as:

$$T = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \tag{3.34}$$

The rows of the matrix sum up to one. All cases of the transition matrix are as follows:

- If the process is in regime 1,
  - it can remain in the normal regime (p_{11}),
  - it can move into the upper jump regime (p_{12}), or
  - it can move into the lower jump regime (p_{13}).
- If the process is in regime 2,
  - it can move into the normal regime (p_{21}),
  - it can remain in the upper jump regime (p_{22}), or
  - it can (not) move into the lower regime (p_{23}). $p_{23}=0$ is plausible for electricity prices, and it can be observed from historical data.
- If the process is in regime 3,
  - it can move into the normal regime (p_{31}),
  - it can (not) move into the upper regime (p_{32}), or
  - it can remain in the lower jump regime (p_{33}).

The upper jumps are iteratively defined as values above the level $\mu+3\cdot\sigma$, while the lower jumps are defined as values that are below the level $\mu-3\cdot\sigma$. Here, μ and σ are the mean and variance values of a corresponding stochastic component of historical prices (see Section 3.2). The regime-switching model for the upper and lower spikes is separately applied for working and nonworking days of different yearly seasons and the transition probabilities are determined for each case, as the number of jumps and the length of jump groups can differ for different day types.

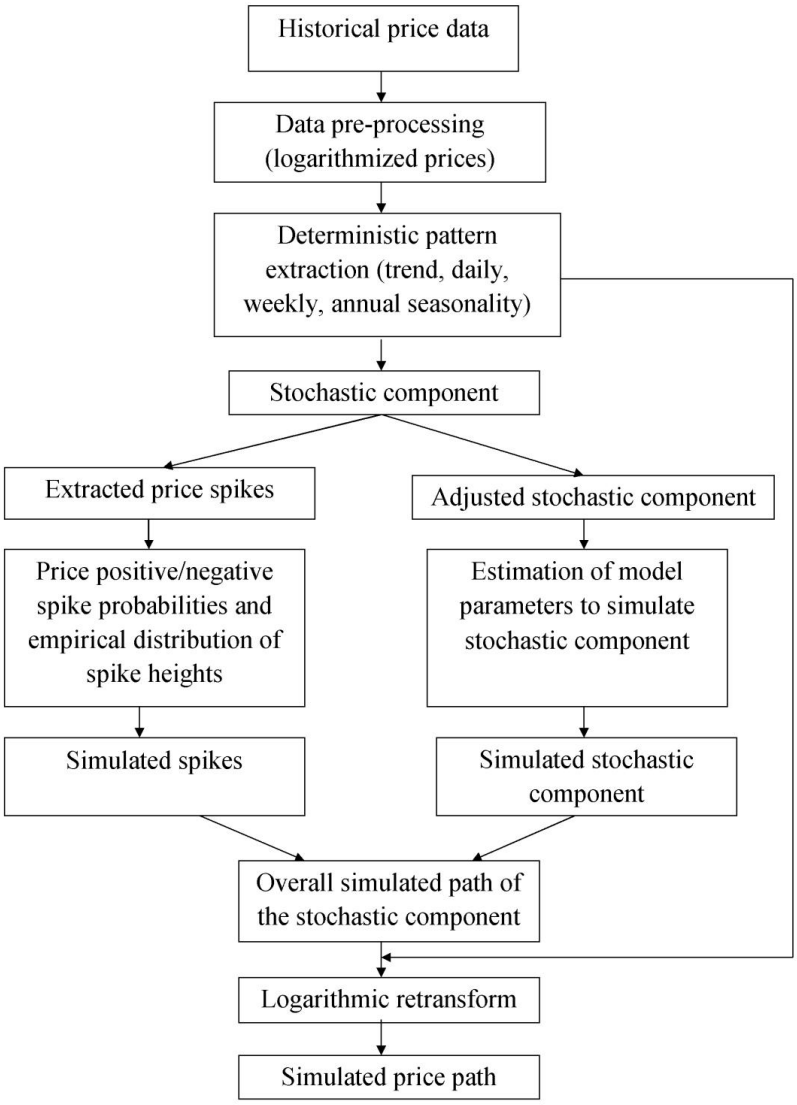

Figure 3.15. Flowchart of the proposed forecasting methodology.

Historical price market data from the period of 1 Jan 2002 to 31 Dec 2009 show that negative spikes are mostly observed in the night and morning hours. The distribution of lower spikes over the week and year has a maximum on Sundays from May to July and on December, respectively (see Figure 3.16). To sum up, the lower spikes have appeared so far during Sundays, the off-peak period, which comprises the time between 00:00 and 08:00 hours, summer, and partly in the winter seasons of the year.

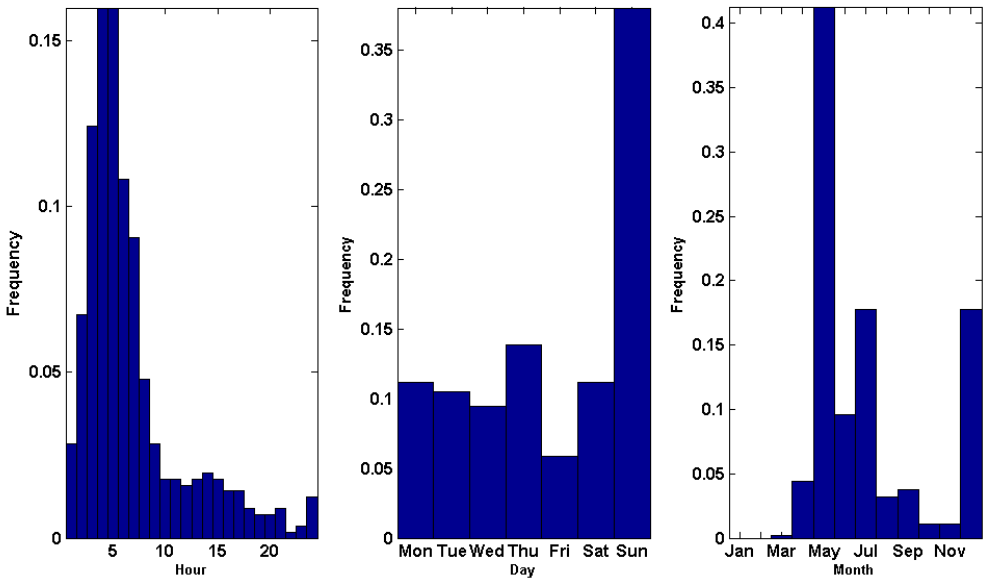

Figure 3.16. Normalized histograms of the occurrence of lower price spikes in the Finnish day-ahead energy market of the years 2002–2009 on different hours, days of the week, and months.

On the other hand, the upper spikes are mostly observed in the day hours (see Figure 3.17). The distribution of upper spikes over the week is observed almost uniformly over the working days. The distribution of upper spikes over the year is mostly observed during winter months. To sum up, the upper spikes have appeared during all working days, the on-peak hour period, which comprises the time between 09:00–12:00 and 15:00–19:00 hours on weekdays, and in the winter season of the year.

Spike magnitudes ($spike_{upper}$ / $spike_{lower}$) are sampled from the empirical distribution functions obtained from historical data. The sampled spike heights are added to a simulated normal regime in the case of the upper spike regime, and subtracted in the case of the lower regime. The spike regime can be described for a time point $h+1$ as:

$$X_{upper\_spike,h+1} = X_{normal,h+1} + spike_{upper,h+1} \tag{3.35}$$

$$X_{lower\_spike,h+1} = X_{normal,h+1} - spike_{lower,h+1}. \tag{3.36}$$

For example, if a MR process is used for the normal regime, the upper regime is modeled as:

$$X_{\text{upper\_spike},h+1} = X_h \cdot \mathrm{e}^{-\mathrm{k}\delta} + \mu\left(1 - \mathrm{e}^{-\mathrm{k}\delta}\right) + \sigma\sqrt{\frac{1 - \mathrm{e}^{-2\mathrm{k}\delta}}{2\mathrm{k}}} \cdot \varepsilon_h \sim N(0,1) + \mathrm{spike}_{upper,h+1} \tag{3.37}$$

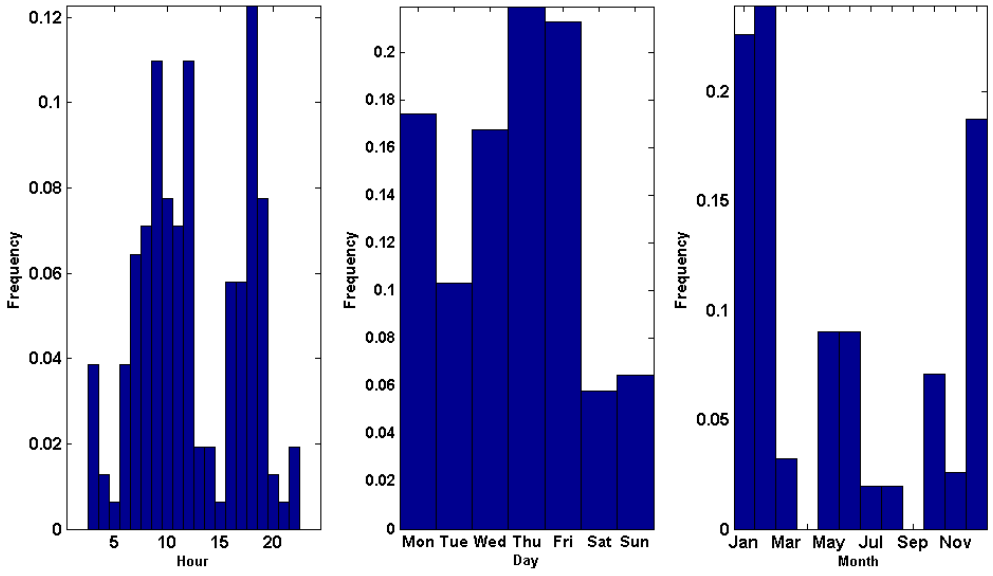

Figure 3.17. Normalized histograms of the occurrence of upper price spikes in the Finnish day-ahead energy market of the years 2002–2009 on different hours, days of the week, and months.

Afterwards, the deterministic components are added again to the stochastic component; the logarithmic simulated path is retransformed to the original range to receive the simulated electricity prices.

The weekly and daily price cycles are very important, as the ACF for the price series shows considerable autocorrelation between the values of the same hours of different days and between the same days of different weeks (see Figure 3.18a). The detrended and deseasonalized price series obtained from the original price is not periodic, even though it still displays some patterns (see Figure 3.18 a,b). As one of the approaches to capture the characteristics of the detrended/deseasonalized series, an ARMA(2,1) model is implemented. Figure 3.18c shows the PACF of the residual series after the ARMA(2,1) model is fitted to the detrended/deseasonalized series. It can be concluded that the model adequately captures the patterns of the data. The PACF values of the squared residuals at several lags are larger than the bounds, which suggests that the residual series have a condition heteroscedasticity (see Figure 3.18d). Finally, ARMA(2,1)+GARCH(1,1) is obtained to model the given process.

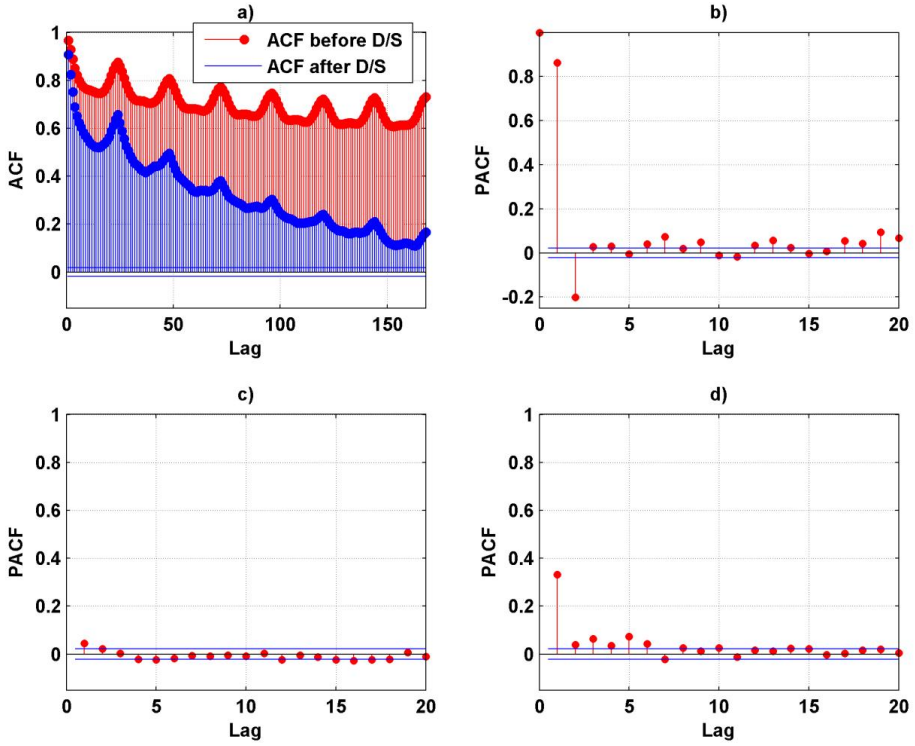

Figure 3.18. a) ACF of the price logs before and after detrending/deseasonalizing; b) PACF of the price logs after detrending/deseasonalizing; c) PACF of the residuals obtained from the ARMA(2,1); d) PACF of the squared residuals obtained from the ARMA(2,1).

After calibrating the models, a number of experiments are carried out to evaluate the goodness-of-fit of each model for an in-sample price path simulation. Figure 3.19 presents single simulated price paths obtained from the MR, GARCH, SARIMA, and ARMA+GARCH models. Based on a general graphical comparison, the results prove to resemble well the true data behavior. The simulated electricity price curves capture daily, weekly, and annual cycles. This is generally caused by the initial removal and addition of the above-mentioned deterministic components before the MR, GARCH, and ARMA+GARCH models are implemented. The SARIMA model adequately captures seasonal patterns to simulate real prices. Price jumps are also generated within the simulated price paths. The MR property is well captured by the models.

Besides the visual investigation, a more detailed statistical comparison of the prices simulated by one of the models with respect to the true series is performed. Table 3.8 presents statistical measurements for simulated price paths obtained from the MR model and original price series. To achieve a more robust result, an expected value for the

measurements is determined based on 50 simulations for the MR model. It can be clearly seen that all the statistical measurements of simulated prices are close to the original prices.

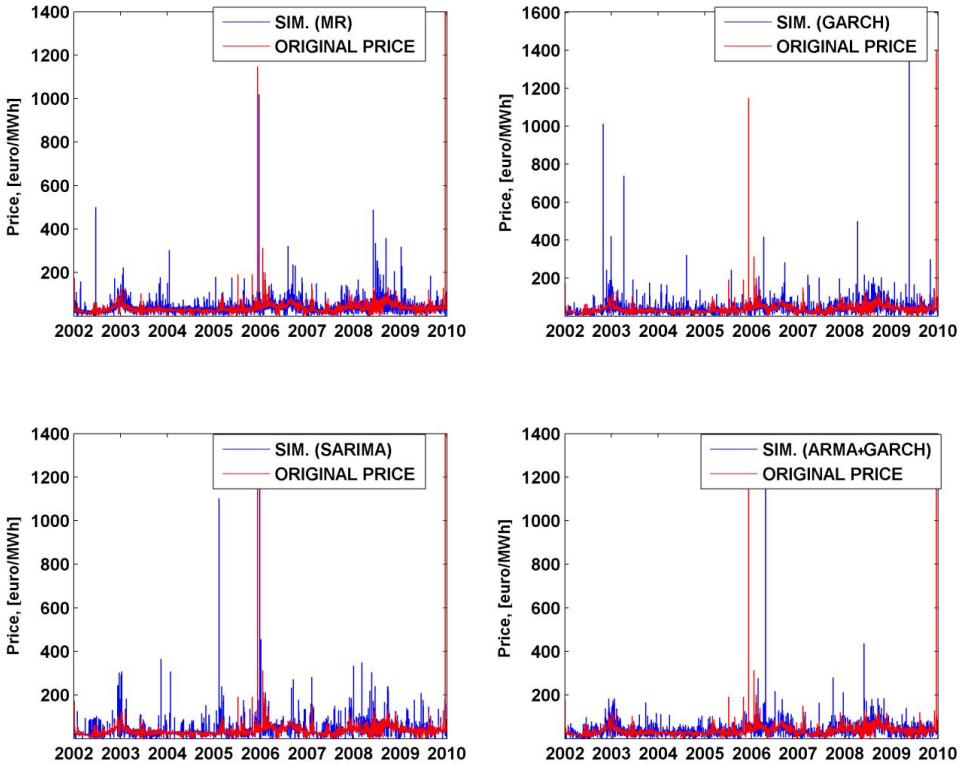

Figure 3.19. Simulated and original Finnish day-ahead electricity prices of the years 2002–2009.

After the in-sample analysis of the model performances, out-of-sample simulations are carried out for the models with regime-switching and preliminary data detrending/ deseasonalizing and without regime-switching (no r/s) and deterending/ deseasonalizing (no seas.). The out-of-sample simulations are run for the period of the first month of the year 2010 and the outcomes are compared with the original prices (see Figure 3.20). The corresponding distributions of the simulated prices with respect to the original prices can be found in Appendix C.

Table 3.8. Basic statistics for original and simulated prices and price spikes in the Finnish day-ahead energy market of the years 2002–2009.

	Number	Mean	Std	Skewness	Kurtosis
Original normal prices	67025	35.53	13.46	1.22	4.80
Simulated normal prices	66989	35.78	14.09	1.20	4.58
Original upper spikes	1456	73.14	67.01	13.87	258.33
Simulated upper spikes	1477	70.52	75.00	15.02	283.04
Original lower spikes	1647	19.05	9.10	0.91	4.43
Simulated lower spikes	1662	17.71	12.26	0.94	4.37

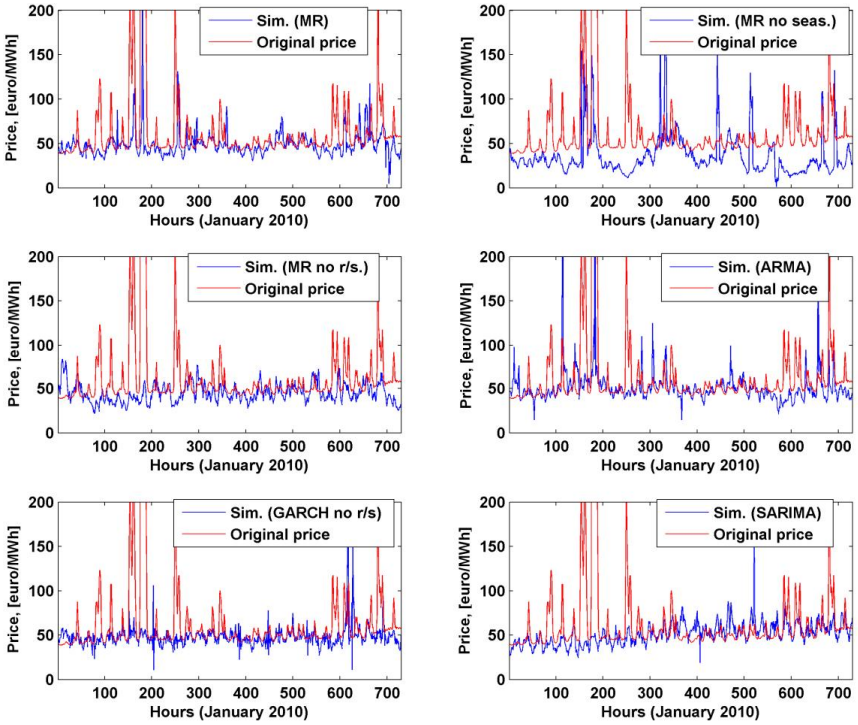

Figure 3.20. Out-of-sample simulated price curves versus the original price curve.

In addition to the graphical comparison of the simulated and historical price paths, MAPE values are calculated for the sorted simulated and real price paths. Table 3.9 shows the expected MAPE values of the out-of-sample analysis when 50 forecasts are carried out.

Table 3.9. Out-of-sample MAPE measures for the different stochastic models for the Finnish day-ahead energy market of the year 2010.

Model	MAPE, [%]
MR	15.57
MR no seas.	38.15
MR no r/s.	24.35
ARMA(2,1)	13.74
GARCH(1,1) no r/s.	16.17
ARMA(2,1)+GARCH(1,1)	12.84
SARIMA(1,1,1)$(1,1,1)_{24}$ no seas.	17.81

3.7.1 **Summary**

A comparison of the results obtained by the models combining regime-switching and decomposition (i.e. detrending/deseasonalizing) techniques with the model results without those techniques showed that the impact of the techniques is very clear. The analysis of the price paths generated by the models without the regime-switching technique makes clear that not only the volatility of the price paths is not well-fitted, but also jumps are not adequately produced. Even the GARCH process, the only method that can handle heteroscedasticity, cannot incorporate jumps with a height that is usually observed in historical prices and generates volatile price paths higher than the historical ones.

The analysis pointed out that a difference filter used within the SARIMA process cannot remove and add deterministic elements accurately for out-of-sample price modeling. Therefore, a separate treatment of the deterministic elements is more effective.

An evaluation of the different models showed that the ARMA/ARMA+GARCH processes enhanced with the regime switching and decomposition techniques outperform other examined processes in fitting the daily and weekly movements and especially the stochastic volatility. These results can be improved by introducing fundamental data (e.g. electricity demand, generation capacity, fuel prices) to the model (e.g. ARMAX) when distinctive structural changes can be captured. Before estimating an ARMAX model, the fundamental data are initially detrended/deseasonalized to have them treated analogically to the deterended/deseasonalized prices.

4 Combination of classical and modern forecasting approaches

The adoption of approaches combining several forecasting models has been advocated in the previous section as a way to improve the forecasting accuracy, as by combining different models, different aspects of the underlying series patterns can be captured. In Section 4.1 the neural network is discussed. A hybrid methodology for the prediction of both normal range electricity market prices and price spikes is presented in Section 4.2.

4.1 NN

Regression models (Nogales et al., 2002), AR models (Fosso et al., 1999), ARIMA models (Contreras et al., 2003), and financial market models, that is, geometrical mean-reverting models (Barlow, 2002) are the classical techniques where an exact model of the system is built and the solution is found by using algorithms that consider the physical phenomena governing the process. These approaches require a lot of information, and the computational costs are very high (Catalão, 2007). Most of the classical models are not able to adequately capture the nonlinearity of the real price behavior. To solve this problem, modern computing techniques have been proposed for electricity price forecasting. The modern computing techniques, namely AI techniques, do not model the system; instead, they find an appropriate mapping between the several inputs and the target variable, usually learned from historical examples, thus being computationally more efficient.

NN model is one of the most popular modern computing techniques implemented for electricity price prediction (Aggarwal et al., 2009). NNs are simple but powerful and flexible tools for forecasting, provided that there are enough data for training, an adequate selection of the input–output samples, an appropriate number of hidden units, and enough computational resources available (Catalão et al., 2007). NNs are able to capture the autocorrelation structure in a time series even if the underlying law governing the series is unknown or too complex to describe. NNs are highly interconnected simple processing units designed to imitate the way the human brain performs a particular task. Each of those units, also called neurons, forms a weighted sum of its inputs, to which a constant term called bias is added. This sum is then passed through a transfer function (e.g. linear, sigmoid, or hyperbolic tangent) (Catalão et al., 2009). Figure 4.1 shows the internal structure of a neuron.

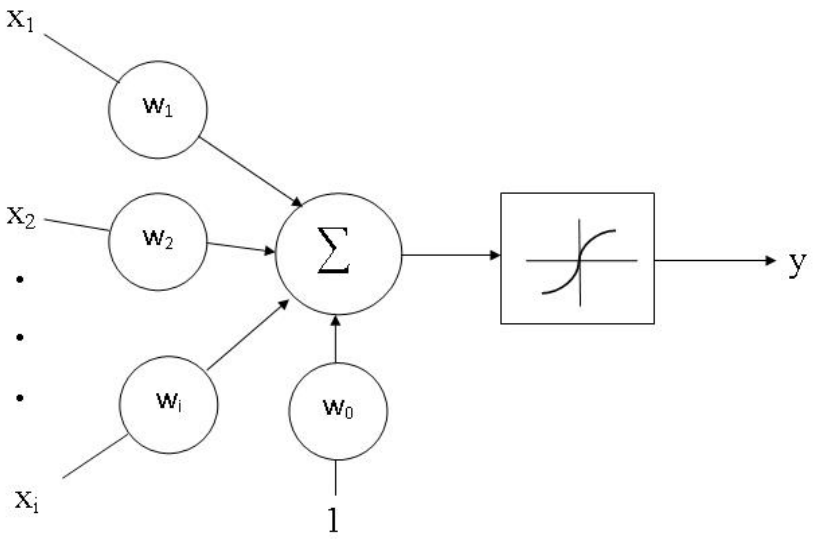

Figure 4.1. Structure of a neuron.

Multilayer perceptrons (MLPs) are the best known and most widely used kind of NN (Aggarwal et al., 2009). Perceptrons are arranged in layers with no connections inside a layer, and each layer is fully connected to the preceding and following layers without loops. The first and last layers are called input and output layers, respectively. Other layers are hidden layers. According to Kolmogorov's theorem, NN can solve a problem by using one hidden layer provided that it has a proper number of hidden neurons (N_h) (Haykin, 1994). Figure 4.2 shows the architecture of a generic three-layered feed-forward NN model that has been most commonly used by researchers to forecast electricity prices (Aggarwal, 2009).

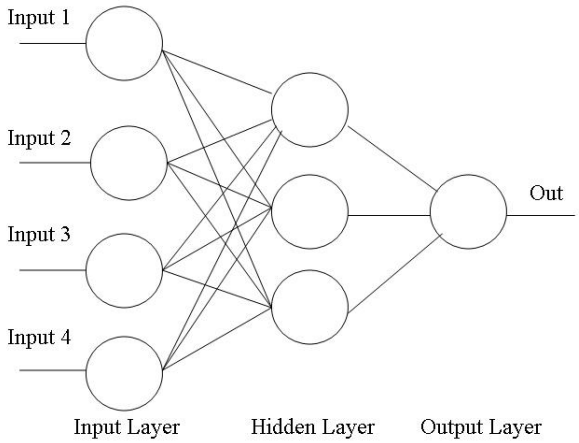

Figure 4.2. Example of a three-layered feed-forward NN model with a single output unit.

The procedure for developing NNs is as follows: data preprocessing, definition of the architecture and parameters, weight initialization, training until the stopping criterion is reached (the number of iterations, the sum of squares of error is lower than a predetermined value), finding the network with the minimum forecasting error on a validation data set, and forecasting future outcomes. The common NN learning algorithm is the backpropagation. It is a steepest descent algorithm minimizing the sum of squared errors by adjusting the weights and biases in each NN's layer.

Three-layered feed-forward NNs with sigmoid and linear transfer functions in the hidden and output layers are implemented within the study. The Levenberg-Marquardt (LM) algorithm, which is an advanced optimization algorithm and more efficient than the usual backpropogation is mainly used in this study for training NNs. General principles of operating the backpropagation and LM algorithms are given in the literature (Yan, 2009).

It should be kept in mind that if there are too few neurons, the network will not be flexible enough to model the data well and, on the other hand, if there are too many neurons, the network may overfit the data. Typically, the number of units in the hidden layer is chosen by trial and error, selecting a few alternatives, and then running simulations to find the one with the best results.

NNs and ARMA models are often compared with mixed conclusions in terms of forecasting capacity. A comparison of the NNs and the ARMA models to forecast commodity prices showed that the NN forecasts were more accurate than the ARMA forecasts (Kohzadi et al., 1996; Catalão et al., 2007). Methodologies that combine NNs and ARMA models have also been proposed to take advantage of the unique strength of each model in linear and nonlinear modeling (Tseng et al., 2002; Zhang, 2003; Wu and Shahidehpour, 2010).

4.2 Hybrid electricity price forecasting model

As mentioned above, linear- and nonlinear-based models have both achieved successes in their own linear or nonlinear domains. However, none of them is a universal model that is suitable for all circumstances. Since it is difficult to thoroughly know the characteristics of the data in a real problem, a hybrid methodology that has both linear and nonlinear modeling capabilities would appear to be a possibly productive strategy for practical use.

Combinations of modern computational intelligence (CI) methods and classical methods or several CI methods have been proposed (Liao, 2007; Wu and Shahidehpour, 2010). Researchers have compared various adaptive and nonadaptive linear and potentially nonlinear models and concluded that the models combining multivariate adaptive linear and nonlinear models outperform other models for many variables (Swanson and White, 1997).

Moreover, while most of the existing hybrid and nonhybrid approaches to forecast electricity prices are reasonably effective for normal range electricity prices, they disregard price spike events. The superiority of separate normal price and price spike forecasting has been advocated in Section 3.7.

A hybrid electricity price forecasting model is proposed where the electricity price time series is analyzed in two parts, that is, normal behavior and spiky behavior. The proposed hybrid model consists of two modules, which are used, respectively, to predict electricity prices within a normal range and price spikes up to one week ahead (Voronin et al., 2013a).

4.2.1 Forecasting strategy

An ARMA-based model is used to catch the linear relationship between the normal range price series and the explanatory variables, a GARCH model is used to unveil the heteroscedastic character of residuals, and a NN model is applied to present the nonlinear impact of the explanatory variables on electricity prices of normal range and to improve predictions obtained from the ARMA-based model. The probability of a price spike occurrence and the height of a price spike are produced by two different forecasting engines. Then, forecasts of normal range prices and price spikes are generated to form an overall price forecast up to one week ahead.

Two previously mentioned competing approaches to handle trend and seasonality presented in price series are considered and implemented to predict prices of normal range. In the first approach, the trend and seasonality terms are directly captured by the forecasting model (i.e., SARIMA). Another approach to handle trend and seasonality in time series is the application of time series decomposition (see Section 3.3.1).

The proposed methodology of the hybrid system can be summarized by the following step-by-step algorithm, shown also in Figure 4.3:

1. The original price time series is statistically divided into a normal range price set and a spikes set by the method introduced in Section 3.2 applying the specific values of *n* and *w*.
2. Both the produced data sets are analyzed and predicted independently through a normal range and price spike modules. The corresponding module is activated according to an output obtained from the price spike occurrence predictor, based on a Gaussian Mixture model (GMM). A GMM based on a Bayesian classifier approximates the probability density function of electricity prices and classifies the given test samples as nonspike or spike. The superiority of a GMM over a Naïve Bayesian classifier is discussed and a mathematical description of a GMM is given in Appendix D (see Section D.1).
3. If the given test sample is classified as a nonspike, the normal range price prediction module is activated.

3.1. Depending on the type of data, that is, raw or decomposed, the SARIMAX or ARMAX model is applied to forecast the normal range prices, and the GARCH model is used to present the heteroscedastic characteristics of the corresponding residuals, resulting in a SARIMAX+GARCH or ARMAX+GARCH model.

3.2. Three-layered NN is applied to present the nonlinear, nonstationary impact of the explanatory variables on electricity prices. The set of inputs for the NN includes both historical and forecasted variables produced by ARMA-based models. Such a strategy aims to improve the price predictions through a combination of linear and nonlinear forecasting techniques.

4. If the given test sample is classified as a spike, the price spike prediction module is activated.

4.1. The magnitude of a price spike is produced by a K-Nearest Neighbor (KNN) model. A KNN model has been previously applied to predict the magnitude of price spikes with promising results (Lu et al., 2005; Zhao et al., 2007a). A mathematical description of a KNN model is given in Appendix D (see Section D.2).

5. The final output is the overall electricity price forecast consisting of a normal range price and price spike forecasts. Evaluation of the hybrid model is implemented separately for normal range prices and price spike modules.

The model is calibrated and evaluated with the Finnish daily day-ahead energy prices of Nord Pool Spot for the period from 1 Jan 2006 to 31 Dec 2009. The use of daily average prices to evaluate the forecasting methodology is motivated by the search for the most appropriate and simple possible model relating day-ahead electricity price characteristics and price spike occurrences to a limited number of exogenous factors. The use of average daily data simplifies the modeling of seasonal components of the series, and the influence of important driving factors are more easily captured in terms of the available data. Moreover, the use of average daily data significantly reduces the

computational costs when a number of different forecasting techniques is tested to select the most accurate approach and the set of external variables.

In the methodology, the whole data are divided into two sets: the training data set and the testing data set. Taking into account data resolution, the daily day-ahead electricity prices of the three years from 1 Jan 2006 to 31 Dec 2008 are used as the initial training data set. The data of one year, from 1 Jan 2009 to 31 Dec 2009, are used as the testing data set. The training interval is shifted and the values of the model parameters are re-estimated in a moving fashion, that is, at each step of forecasting. By using the models obtained, the prices are predicted on the testing data set of a length up to seven days.

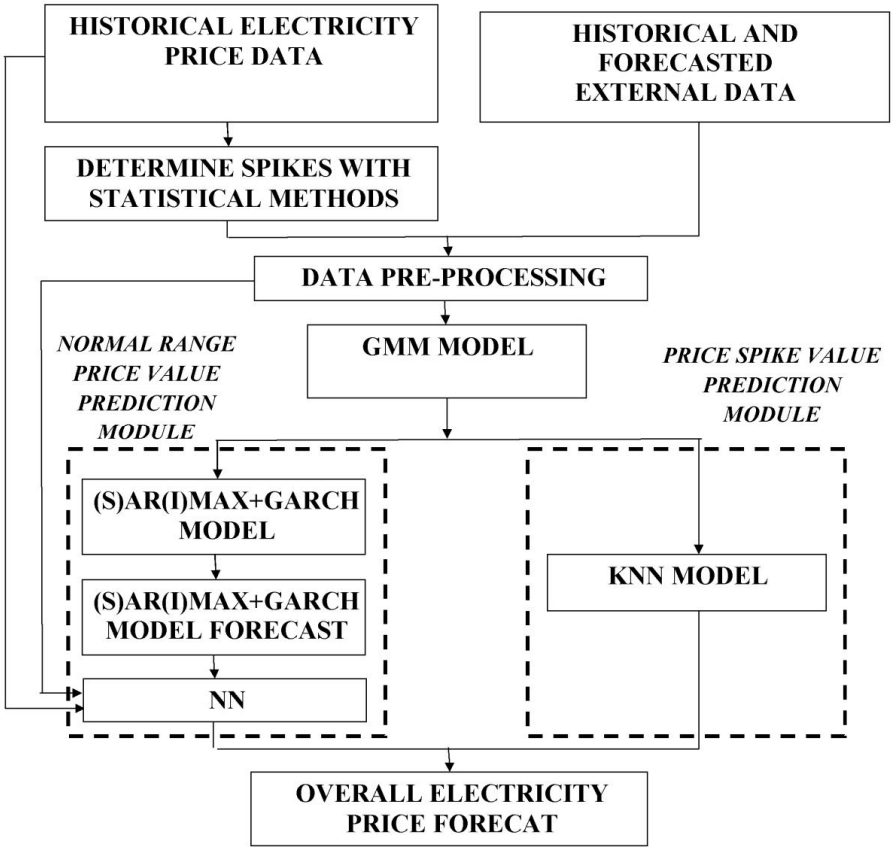

Figure 4.3. Flowchart of the proposed forecasting methodology.

4.2.2 Normal price module

Electricity demand and base generation are united into the nonbase electricity demand and used as an exogenous factor in the SARIMAX/ARMAX+GARCH model of the

normal price module. Figure 4.4 shows the relationship between the nonbase electricity demand and the Finnish daily day-ahead electricity prices over the period from 1 Jan 2006 to 31 Dec 2009.

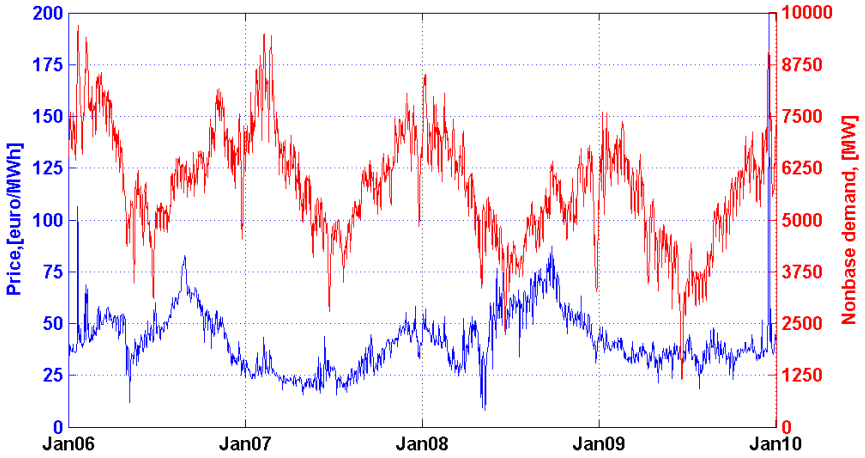

Figure 4.4. Finnish daily day-ahead electricity prices and nonbase electricity demand.

Then, the combined model of the normal price module is implemented on raw or decomposed data as SARIMAX+GARCH+NN or ARMAX+GARCH+NN, respectively. The structures of the implemented SARIMAX/ARMAX+GARCH models and their respective parameter estimates are reported in Appendix D (see Section D.3). When daily day-ahead electricity prices are predicted, the total electricity demand and generation (i.e., internal supply) values are forecasted by the SARIMA model (Taylor, 2006). The values of hydro and nuclear power generation are forecasted by a simple random walk model described in Appendix D (see Section D.4).

The selection of input features fed to the NN is mainly based on past experience in the study. Here, along with the forecasted price and demand values, the historical price and demand data are applied as the inputs for the NN to indicate the trend and weekly periodicity of a given series. Thus, there are six neurons in the input layer of the NN of the combined model to predict the price value on a single test day D: a SARIMAX+GARCH/ARMAX+GARCH model price prediction on day D; a historical price on day D-1; a historical price on day D-7; a historical price on day D-14; a nonbase electricity demand on day D, and a nonbase electricity demand on day D-1.

Besides the proposed combined model, other models implemented to predict normal range prices are examined: a random walk model (as a benchmarking model), an ARMA/SARIMA+GARCH model, an ARMAX/SARIMAX+GARCH model, and an NN (for original and decomposed data). Here, two ARMA-based models without an external variable (ARMA/SARIMA+GARCH) are used to check whether the inclusion

of the nonbase electricity demand in the price forecasting model could result in a significant improvement in the forecasting performance.

4.2.3 Price spike module

Given n = 3 and w = 90 days, the thus defined price spikes are extracted from the original price series, as shown in Figure 4.5.

Table 4.1 shows the basic distribution parameters for prices and spikes. It can be seen from the number of spikes (N_{spike}) that the spikes constitute less than 1.5% of all the daily prices. However, their magnitude and unexpectedness cause them to have a disproportionate significance in the electricity markets. The statistics show that there is zero probability of an electricity price spike during weekends and holidays.

Table 4.1. Basic statistics for normal prices and price spikes over the period 2006-2009.

	Number of observations	Mean	Std	Skewness	Kurtosis	Weekday, N_{spike}	Weekend/ Holiday, N_{spike}
Normal	1436	41.26	13.28	0.55	2.91	—	—
Spikes	25	71.86	41.88	3.26	14.65	25	0

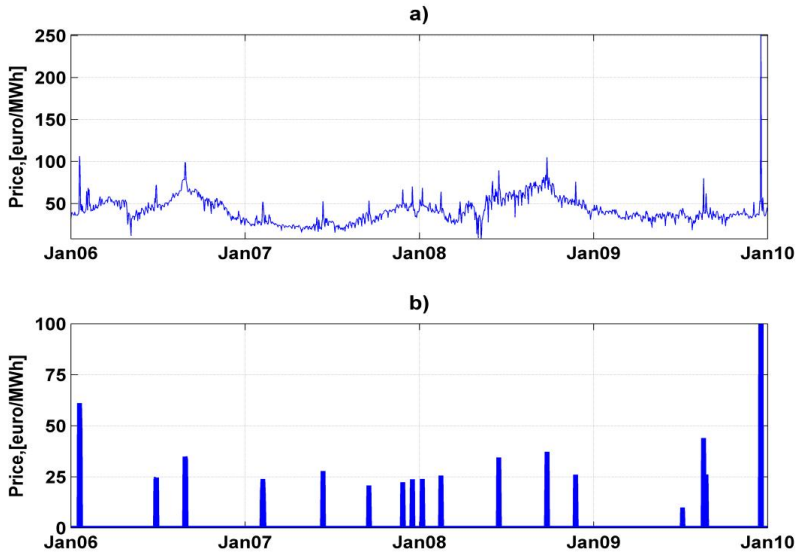

Figure 4.5. a) Original Finnish daily day-ahead electricity prices for the period 1 Jan 2006–31 Dec 2009; b) extracted price spikes.

In an ideal competitive electricity market, price spikes occur only when the demand exceeds supply. Most electricity markets, however, are not ideally competitive. Therefore, price spikes may take place even when the supply completely covers the demand. The set of attributes selected to determine the probability of price spike occurrence and its magnitude are given below:

- *SDI.* This study uses the composite relationship between electricity price, demand, and supply that was proposed in (Lu et al., 2005) and presented as a supply-demand balance index (SDI). The SDI on a single day D is defined in Eq. 4.1:

$$SDI(D) = (Supply(D) - Demand(D)) / Demand(D) \cdot 100\% \quad (4.1)$$

  where *Demand(D)* is the market demand on day D, and *Supply(D)* is the electricity supply on day *D*.

- *Nonbase electricity demand.* The importance of electricity demand for electricity price forecasting was discussed in Section 3.3.2.

- *Temperature.* Atmospheric temperature is chosen as a main indicator of weather extremity in the electricity price spike study. The main electricity consumption areas in Finland are the south and central regions (Statistics Finland, 2012), and the temperature data for the city of Helsinki are used because the geographical location of the city indicates a temperature that is relevant to overall electricity consumption in the country. Temperature data forecasted for the city of Helsinki are available on the Weather Underground web site (Weather Underground, 2012).

- *Elspot capacity-flow difference.* Power transfer constraints for electricity come in the form of a capacity limit on the transmission lines and the transmission losses, which can make it impossible or uneconomical to transfer electricity in certain regions (Lucia and Schwarz, 2000).

  Two regimes of the Finnish electricity system are considered. One of the regimes is the regular regime; the other, the nonregular regime, is the capacity-limited regime and exists when the difference between the total Elspot power flow and the total Elspot capacity to Finland is close to zero. Congestion and thus extreme price changes are more likely to occur when the difference between the total Elspot power flow and the total Elspot capacity is small.

  The total Elspot power flow and the total Elspot capacity to Finland were calculated as a daily sum of the Elspot net exchange and Elspot capacities from Sweden, Norway, and Estonia to Finland, respectively. The power flow data and the generation and demand data for the Finnish electricity system are provided by Fingrid, the company responsible for the high-voltage electricity transmission in Finland (Fingrid, 2013b). The Elspot power flow and capacity data for day D-1 are published by the TSO and are available on day D-2. Therefore, to forecast the flow

and capacity for day D, the flow and capacity data of day D-1 are considered known. The Elspot power flow data have strong seasonal patterns, which can be captured by SARIMA. The Elspot capacity is rather constant during the whole week.

- *Temporal effect.* In addition to the physical factors given above, the day status of the sample needs to be implemented into the forecasting model (similarly to the model presented in Section 3.7). The whole data set was divided into weekdays, weekends, and holidays of different yearly seasons.

The distribution of the prices versus the chosen driving factors is shown in Figure 4.6.

Note that the factors cannot exactly determine the occurrence of spikes. The Gaussian Mixture model predicts spikes by evaluating their occurrence probability. The inputs of the model are not necessarily the determinants of spikes.

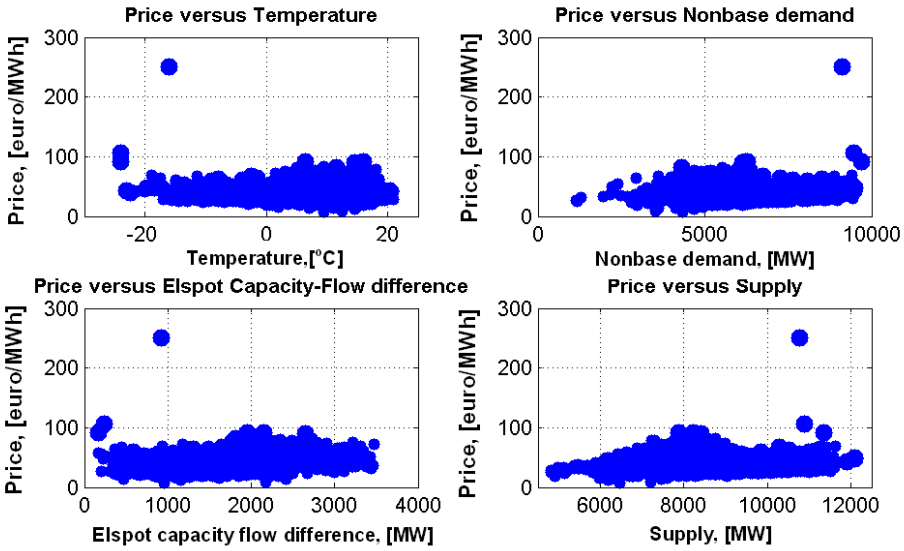

Figure 4.6. Scatter plots of the prices versus potential price spike driving factors.

One modification is implemented within the GMM model. The probability of spike occurrence is calculated for every input vector and then compared with a predetermined threshold denoted as V_0. If the probability is larger than the threshold, a spike is predicted to occur, regardless of whether this probability is larger than the probability of nonspike occurrence. This modification is performed because the price spike prediction problem is a serious imbalanced classification problem (i.e., some classes have many more samples than other classes) (Zhao et al., 2007a). In fact, the probability of spikes is less than the probability of nonspikes on most occasions. Many spikes occur when

their occurrence probabilities are smaller than 50%. Without setting a threshold smaller than 50%, many spikes will be misclassified. The threshold can be determined by historical data. A Bayesian-based classifier considering prior information, that is, prior class probability, tends to be less prone to problems regarding sample class imbalance.

4.2.4 Normal range price forecasting results

The performance of the proposed combined model applying raw and decomposed data is compared with the performance of the previously mentioned seven models predicting normal range prices up to one week ahead. Figure 4.7 summarizes the statistical measures in terms of AMAPE that characterize the prediction accuracy of the different models studied. In addition to Figure 4.7, their MAE, MSE, and AMAPE values corresponding to different forecasting horizons are given in Appendix D (see Section D.5).

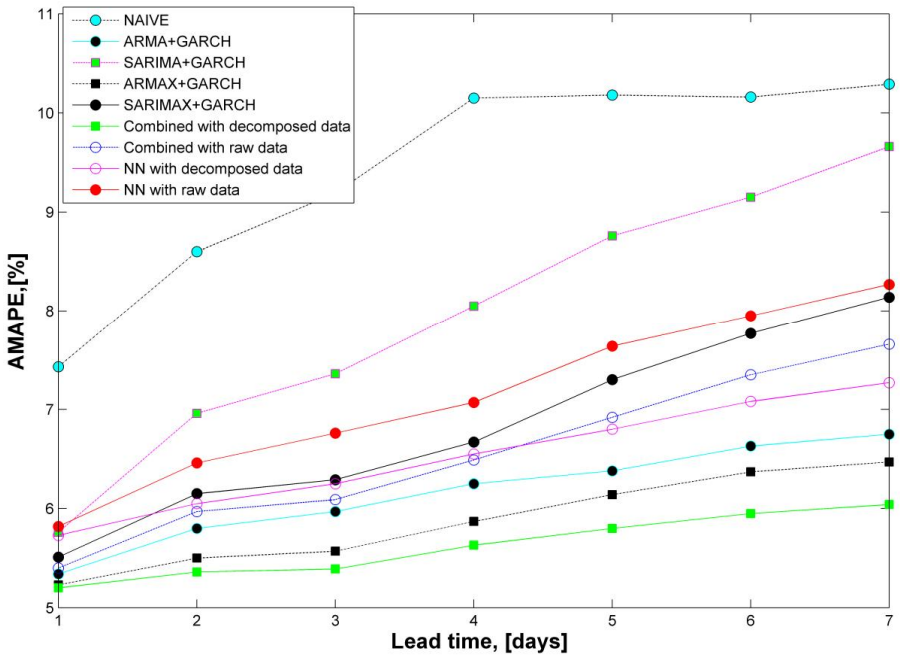

Figure 4.7. AMAPE results plotted against lead time.

Figure 4.7 shows that the naïve benchmark model (random walk model) was substantially outperformed by all the other methods at all lead times. The SARIMA+GARCH model implemented on raw data has the poorest performance of the other models at all lead times, except for a lead time of one day when it shows a very

similar performance to the NNs. The figure, furthermore, shows that the models using decomposed data as an input performed better than the analogical models trained on original data. For the ARMA-based methods, the decomposing preprocessing technique is more effective than the direct entering of trend and seasonal terms into a model. It supports the previous findings (see Section 3.7). The NNs built with detrended and deseasonalized data can produce significantly more accurate forecasts than with the original data. This result suggests that NNs built on raw data are unable to adequately learn seasonality and trend; a finding that refers to previous studies on NNs (Nelson et al., 1999; Zhang and Qi, 2005). It is unsurprising that the performance of the SARIMA+GARCH and ARMA+GARCH models was much improved after the inclusion of exogenous factor information, resulting in SARIMAX+GARCH and ARMAX+GARCH models, respectively. The performance of the NNs relative to the ARMA-based models was worse, which differs from the work in (Zhang and Qi, 2005). A possible explanation could be that the specific characteristics of the initial price time series have become more linear after the transformation of hourly data into daily data and spikes elimination. Moreover, the forecasted values of prices on day D-1 were used as an input to the NNs for the multistep predictions. In a nonlinear model, errors might be spread significantly. It must be noted that if a particular NN fails to produce good results, this does not indicate that NNs in general are poor predictors because a different specification of the NN could have performed better (Taylor, 2006). Problems may arise from the difficult tune-up of the NN algorithms, which need validation of the model by the number of hidden layers, and the number of neurons in the input and hidden layers (Conejo et al., 2005a). Of the remaining models, the combined model applying decomposed data performed considerably better than all the other models. Therefore, the combined model with decomposed data is used as the normal range price prediction module of the hybrid model.

Figure 4.8 focuses more closely on the AMAPE results for lead times of one and seven days for the two best models, that is, ARMAX+GARCH and the combined model with decomposed data, against the day of the week. Both techniques have relatively poor AMAPE results for Sundays and Mondays. This problem may arise from the model algorithm used. The price depends on the time of day and whether it is a weekday or weekend. Weekdays and weekends have substantially different characteristics in terms of electricity price distribution. Some authors divide the whole data set into weekday and weekend data sets to build different models for each data set (Wu and Shahidehpour, 2010; Wang and Ramsay, 1998; Gao et al., 2000) whereas others just report results for models fitted and evaluated on weekday observations only (Lora et al., 2002). In this study a division of data into weekdays and weekends was not performed when normal range prices were predicted. Therefore, it would seem that all the models have been somewhat challenged by being evaluated and tested on data with slightly different intraweek seasonal characteristics. In this context, the robustness of the models becomes important.

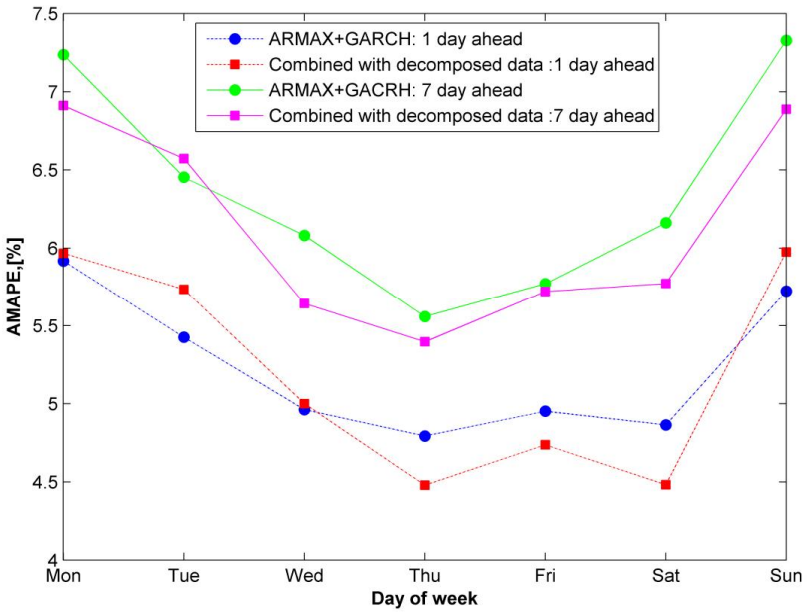

Figure 4.8. AMAPE results for the lead time of one and seven days plotted against the day of week.

4.2.5 Price spike forecasting results

It is important to define reliable measures to assess the performance of the classification model when the occurrence of spikes is predicted. Some classification performance measures have been proposed in (Lu et al., 2005). A standard performance measure of a classification is the estimate of the probability of correct classification:

$$Classifier\ accuracy = (N_{corr\_all} / N) \cdot 100\% \tag{4.2}$$

where N_{corr_all} is the number of correctly classified patterns, and N is the total number of patterns.

This measure is robust for many classification problems but not for the problem under discussion here. Since the data in the problem are extremely unbalanced, the value of the measure will remain high even if all the spikes are misclassified. Thus, other classification performance measures, namely, spike prediction accuracy and spike prediction confidence, are proposed to solve this problem.

Spike prediction accuracy is a ratio of the number of correctly classified spikes N_{corr} to the number of actual spikes N_{spikes} in terms of percentage:

$$Spike\ \ prediction\ accuracy = (N_{corr} / N_{spikes}) \cdot 100\% \tag{4.3}$$

This measure was introduced because the ability to correctly predict spike occurrence is the subject of greatest concern.

Spike prediction confidence is another measure that aims to account for the uncertainties and risks carried within the forecast. Spike prediction confidence is described as

$$Spike\ \ prediction\ confidence = (N_{corr} / N_{as\_spikes}) \cdot 100\% \tag{4.4}$$

where N_{corr} is the number of correctly classified spikes and N_{as_spikes} is the number of observations classified as spikes. As the classifier may misclassify some nonspikes as spikes, this definition is used to assess the percentile in which the classifier makes this kind of a mistake. These measures are used within the thesis to estimate the classification performance of the price spike forecasting models.

Actual electricity price spikes and estimated probabilities of the spike occurrence obtained from the GMM model for a lead time of one day and seven days are shown in Figures 4.9–4.10, respectively. The spike threshold resulting in the best overall performance of the model on a validation set is determined as 38%.

Classifier accuracy, spike prediction accuracy, and spike prediction confidence values decline as the forecasting horizon increases. The spike prediction accuracy and spike prediction confidence values vary between 60–80% and 50–60% respectively, depending on the forecasting horizon (see Table 4.2). Of the five spikes in the testing data set, four of them were predicted by the model for a lead time of one day, and three of five spikes were predicted for a lead time of seven days.

Values of the three closest samples were selected to determine the unknown value of a price spike by the KNN. Table 4.3 shows that most of the error rates of the price spike value forecast are less than 35% for the lead times of one and seven days, with only one case close to 70%. The forecast error for this particular day can easily be understood given that the actual price is much higher than the historical average price spike. In fact, other forecasted price spike values are close enough to the real values to provide useful information for practitioners.

The output values of the spike forecasting model partially indicate the robustness of this model because of the small number of spiky observations in the testing data set. The model robustness and the stability of the model to changes in the training data may be assessed by cross-testing. Here, the aim is to estimate how accurately the predictive model would perform in practice. In cross-testing, the original sample is randomly portioned into training and testing subsamples of a specific size. This parting process is independently repeated T times to yield T portioned data sets, which are treated as independent sets. The cross-testing estimates matrix of the model performance measures

for the testing data set, denoted $\widehat{\Omega}^{*(.)}$, is merely the mean matrix of the T estimates on the individual testing data sets executed by the data parting

$$\widehat{\Omega}^{*(.)} = \frac{1}{T}\sum_{t_p=1}^{T} \widehat{\Omega}^{*(t_p)} \tag{4.5}$$

where $\widehat{\Omega}^{*(t_p)}$ is a matrix of the model performance measures for the testing set on the portioned sample t_p.

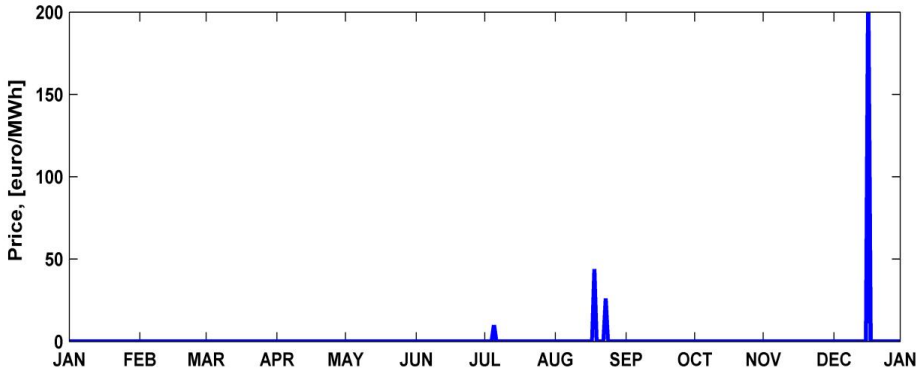

Figure 4.9. Actual electricity price spikes for the period 1 Jan 2009–31 Dec 2009 of the Finnish day-ahead energy market.

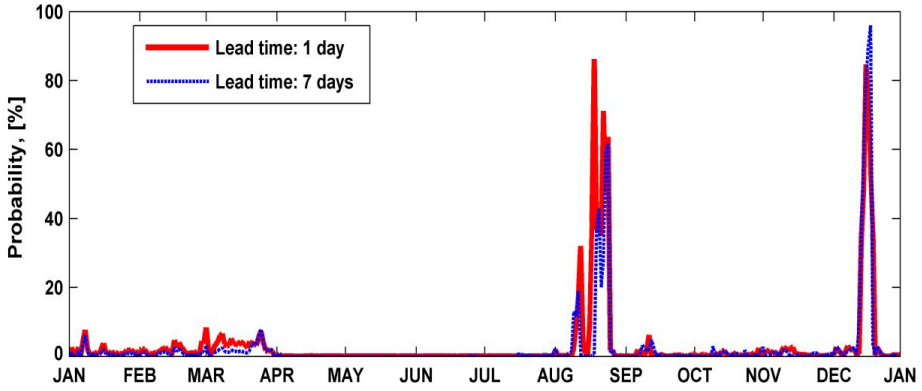

Figure 4.10. Forecasted probability of spikes for a lead time of one day and seven days for the period 1 Jan 2009–31 Dec 2009 of the Finnish day-ahead energy market.

Table 4.2. Accuracy and confidence of the probability model on the testing data for lead times of one and seven days.

Horizon	Classifier accuracy, [%]	Spike accuracy, [%]	Spike confidence, [%]
1	98.63	80.00	57.14
7	98.36	60.00	50.00

Table 4.3. Comparison between the actual and price spike values forecasted by the KNN on the testing data for lead times of one and seven days.

Horizon	Spike number	Forecasted price, [euro/MWh]	Original price, [euro/MWh]	Forecast error, [%]
1	1	38.11	44.45	14.40
	2	59.40	79.12	24.92
	3	51.04	65.30	21.47
	4	65.20	56.35	15.71
	5	88.57	251.04	64.71
				Mean: 28.56
7	1	34.16	44.45	23.15
	2	53.46	79.12	32.43
	3	48.57	65.30	25.62
	4	61.34	56.35	8.85
	5	73.14	251.04	70.86
				Mean: 31.65

At the cross-testing stage, the robustness of the model is checked on testing data sets of a specific size. The whole data set from 1 Jan 2006 to 31 Dec 2009 (see Figure 4.4) is randomly divided into training data sets and testing data sets T times. The model is fitted by using the training data set, and by using the model obtained, the values are predicted seven days ahead on the testing set of specific length. Note, that in this study it was opted not to re-specify or re-estimate the model values at each step of forecasting. This action gives an opportunity to compare the performance of the model using the specific size of the training and testing data sets.

The seemingly not high rate in the values of the spike prediction accuracy and confidence for the probability model and the relatively high error values for the KNN model are mainly due to the very limited number of price spike events in the historical data (1.5% of the whole sample). The values were obtained with insufficient data containing spikes. Many stochastic events causing spikes could not be considered in the model. Tables 4.4 and 4.5 suggest that further training with historical data would improve accuracy and confidence of the models. The spike prediction accuracy using the proposed probability model is above 50% for the testing data set of seven days,

which means that more than 50% of spikes can be predicted. In the light of the fact that price spikes are highly stochastic, the achieved forecast accuracy level is sufficient to provide market participants with an ability to anticipate price spikes.

Table 4.4 and Table 4.5 show the results of the performance measures of the GMM and KNN for testing data sets of a specific size after *T=500* simulations of the cross-testing:

Table 4.4. Accuracy and confidence of the GMM on randomly selected testing data sets of a specific size after T = 500 simulations.

	Length of a testing data set, [days]			
	365	180	30	7
Classifier accuracy, [%]	98.53	97.27	96.61	91.64
Spike prediction accuracy, [%]	41.01	43.45	48.76	57.82
Spike prediction confidence, [%]	29.24	31.34	36.93	45.46

Table 4.5. Error value of the KNN on randomly selected testing data sets of a specific size after T = 500 simulations.

	Length of a testing data set, [days]			
	365	180	30	7
Error value, [%]	33.75	28.17	22.35	18.23

4.2.6 **Overall price prediction**

Integration of the spike occurrence probability and value forecasting results with the normal range price forecasting results obtained from the combined model (ARMAX+GARCH+NN) with decomposed data (the best performing normal range price model) gives a complete electricity price forecast. The forecasted normal range prices, the probability of price spike occurrence, and a complete electricity price forecast for a lead time of one day are shown in Figures 4.11–4.13, respectively. Obviously, without the spike occurrence and spike value predictors, the performance of the normal range price forecasting model deteriorates when spikes occur.

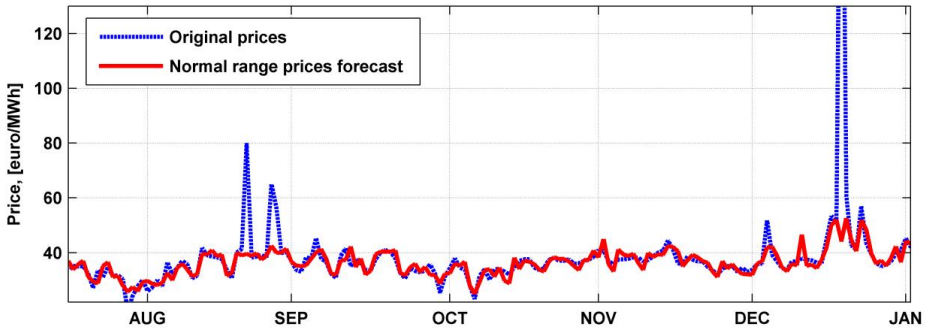

Figure 4.11. Forecasted normal range prices for a lead time of one day for the period Jul 2009–Dec 2009 of the Finnish day-ahead energy market.

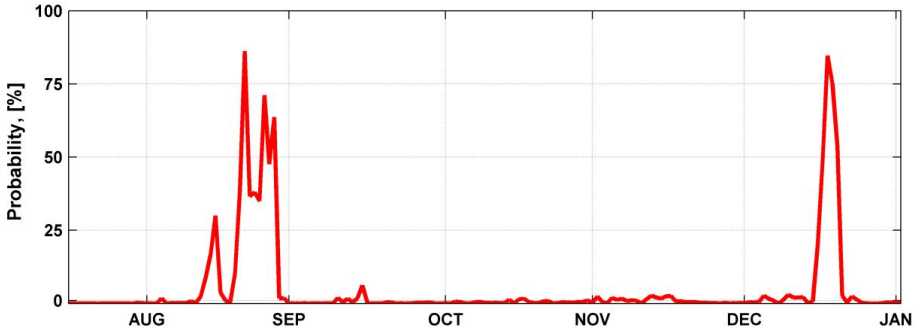

Figure 4.12. Forecasted probability of spike occurrence for a lead time of one day for the period Jul 2009– Dec 2009 of the Finnish day-ahead energy market.

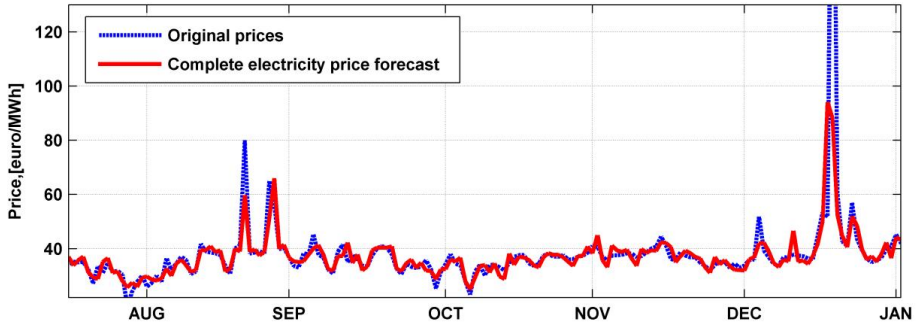

Figure 4.13. Complete electricity price forecast ((ARMAX+GARCH+NN) + (GMM + KNN)) for a lead time of one day for the period Jul 2009–Dec 2009 of the Finnish day-ahead energy market.

4.2.7 Summary

Coupling of the normal price range prediction and the price spike prediction provides valuable information about the electricity market and gives the market participants an ability to manage their risks. Hence, when applied in addition to the normal range price forecast, the proposed price spike forecast method can provide practically useful and reasonably accurate forecasts, enhancing the applicability of price forecasts in the actions of electricity market participants. In the light of the fact that price spikes are highly stochastic, the achieved spike forecast accuracy level is acceptable. Further training with historical data would improve the accuracy.

The forecasting principle implying separate normal price and price spike prediction is utilized in further study presented in Chapter 6.

5 Tuning of the forecasting model parameters

Most of the previous studies have utilized past experience in selecting parameter settings and the input variables for the forecasting model to be applied. However, as each market has characteristics of its own, the choice of the corresponding relevant factors and model parameters is still an open area of research. A two-step feature selection algorithm is introduced in Section 5.1. The search procedure to tune the parameter settings and obtain the best inputs for a particular forecasting model is proposed and implemented for different forecasting engines in Section 5.2. Section 5.3 presents a hybrid forecasting method for simultaneous prediction of price and demand in the day-ahead energy market.

5.1 Feature selection

Feature selection is a process commonly used in forecasting model learning, wherein a subset of features available from data is selected for the application of a learning algorithm. Among different factors that may have an impact on the prices, the most informative ones have to be identified. In previous works (see Chapter 4), past experience in the study of different markets was utilized to select input features for a forecasting model. These selected inputs were based on daily and weekly periodicity and trend of the electricity price signal, and there was no reasonable proof. Obviously, this selection method is not efficient because of the complex time-dependent behavior of electricity prices and the large number of effective input features.

A set of factors that affect prices in the Finnish day-ahead energy market can be very extensive. Thus, there may be irrelevant and redundant inputs in the set of inputs, which can mislead the forecast engine. It is necessary to refine the initial set of potential inputs

such that a subset of the most effective inputs is selected for the forecast engine (Amjady and Keynia, 2010). An analytical method that can select a minimum set of the most effective input features for day-ahead price prediction is very valuable.

To get the subset of the most effective input features, the relevance-redundancy feature selection algorithm is used. The ability to filter out redundant information from the set of candidate features is the benefit of this procedure versus a simple calculation of a relevance value between the target and explanatory variables (Yu and Liu, 2004).

In the relevance-redundancy feature selection algorithm, $SET_1 = \{x_1,x_2,\ldots,x_k\}$ is supposed as a set of candidate inputs. In the first step (i.e., relevance filtration), a relevance value between each candidate input $x_i \in SET_1$ and the target y (continuous or binary) is computed as $RV(x_i,y)$. If an absolute relevance value between the candidate input and the target is greater than a prespecified value V_1, this candidate feature is retained for further processing; otherwise, it is filtered out

$$abs(RV(x_i,y))>V_1,\ 1\leq i \leq k \tag{5.1}$$

In the second step (i.e., redundancy filtration), the set of the retained candidate inputs is supposed as $SET_2 \subset SET_1$. For any two retained candidate inputs $x_a, x_b \in SET_2$, a relevance value between those candidate inputs is computed as $RV(x_a, x_b)$ and supposed as the redundancy measure. If the absolute redundancy measure between any two candidate inputs (x_a and x_b) is smaller than a prespecified value V_2, both inputs are retained; otherwise, only the input with the largest relevance value with respect to the target ($RV(x_a,y)$ or $RV(x_b,y)$) is retained. For instance, if for $x_a, x_b \in SET_2$,

$$abs(RV(x_a,x_b))>V_2,\ 1\leq a,b\leq k,\ a\neq b \tag{5.2}$$

The redundancy filtering process is repeated for all candidate inputs of SET_2, until no redundancy measure becomes greater than V_2, then the subset of candidate inputs $SET_3 \subset SET_2$ that has passed the redundancy filter is finally selected as the optimal inputs by the proposed relevance-redundancy feature selection algorithm.

Several feature selection techniques are applied in the study to calculate a relevance value between given variables in the proposed feature selection algorithm. The first considered feature selection technique often used for a day-ahead electricity price forecasting is a simple linear correlation (Moghram and Rahman, 1989; Rodriguez and Anders, 2004; Amjady and Keynia, 2008; Vahidinasab et al., 2008). Linear correlation analysis is widely used for the feature selection; however, it is a linear technique and, therefore, often cannot consider the nonlinearities of the original price signal. Thus, to validate the weakness of the linear based techniques, sequential feature selection based on linear regression is considered as another linear based technique and put among the competing feature selection approaches (Rückstieß, 2011). Then, a group of feature selection techniques based on nonlinear approaches is considered. These techniques are mutual information (MI) criterion, Relief, and KNN (see Appendixes D and E). The

above-mentioned feature selection techniques and their combinations are examined to find the best approach resulting in the highest accuracy when prices are predicted.

5.2 Proposed search procedure to tune the model parameters

As mentioned above, efficiency of a forecasting model is highly dependent on the correct setting of its adjustable parameters and selected inputs. The following search procedure aims at finding the optimal setting parameters of a forecasting model including threshold values (V_1 and V_2) for the proposed relevance-redundancy feature selection algorithm.

For instance, the proposed search procedure has three main adjustable parameters when a three-layered NN is used as a forecasting engine: N_h of the NN and two threshold parameters V_1 and V_2 for the relevance and redundancy filters, respectively. Here, an iterative search procedure is carried out, which can automatically adjust the above-mentioned parameters with a minimum reliance on the heuristics. The procedure is outlined below and shown also in Figure 5.1:

1) Set initial values of N_h, V_1 and V_2.

2) By the selected inputs, training samples are constructed. The NN is trained and the corresponding validation error (e.g. AMAPE) is evaluated and stored. Here, the validation set is extracted from the training sample and kept hidden to the model during its training period and is being used to examine the model predictability.

   For the sake of having an adequate validation set (accurate representative of a forecast horizon), a day before a forecast day is used as a validation set. However, if a selected "day before" has a status (nonworking/working day) different from the status of a forecasting day, the historically last day having the status same as the forecasting day is used as a validation set. This modification is made to distinguish the working and nonworking day patterns within a forecasting model.

3) Each adjustable parameter is varied in turn at a neighborhood around its previously selected value, while two remaining parameters are kept constant. A fixed radius of neighborhood (e.g. ± 25 % of the previously selected value) is considered in the local search. For instance, if N_h is varied with the unit step and its previously selected value is 8, then the NN is trained with $N_h = \{6,7,8,9,10\}$. Therefore, for each value of the varied parameter in the neighborhood, training of the NN is repeated and validation error is evaluated and stored. The value of the varied parameter resulting in the least validation error is selected and fixed.

   When only the first cycle of the variation is executed (i.e. all adjustable parameters are varied once), this cycle is repeated again. This modification is made to avoid a local minimum trap in the search procedure. Therefore, if the procedure misses the optimum solution in one cycle, it may find the optimum point in the next cycle.

4) If the selected values of N_h, V_1, and V_2 obtained after two consequent cycles are the same, the iterative search procedure is terminated. Otherwise, go back to step 3.

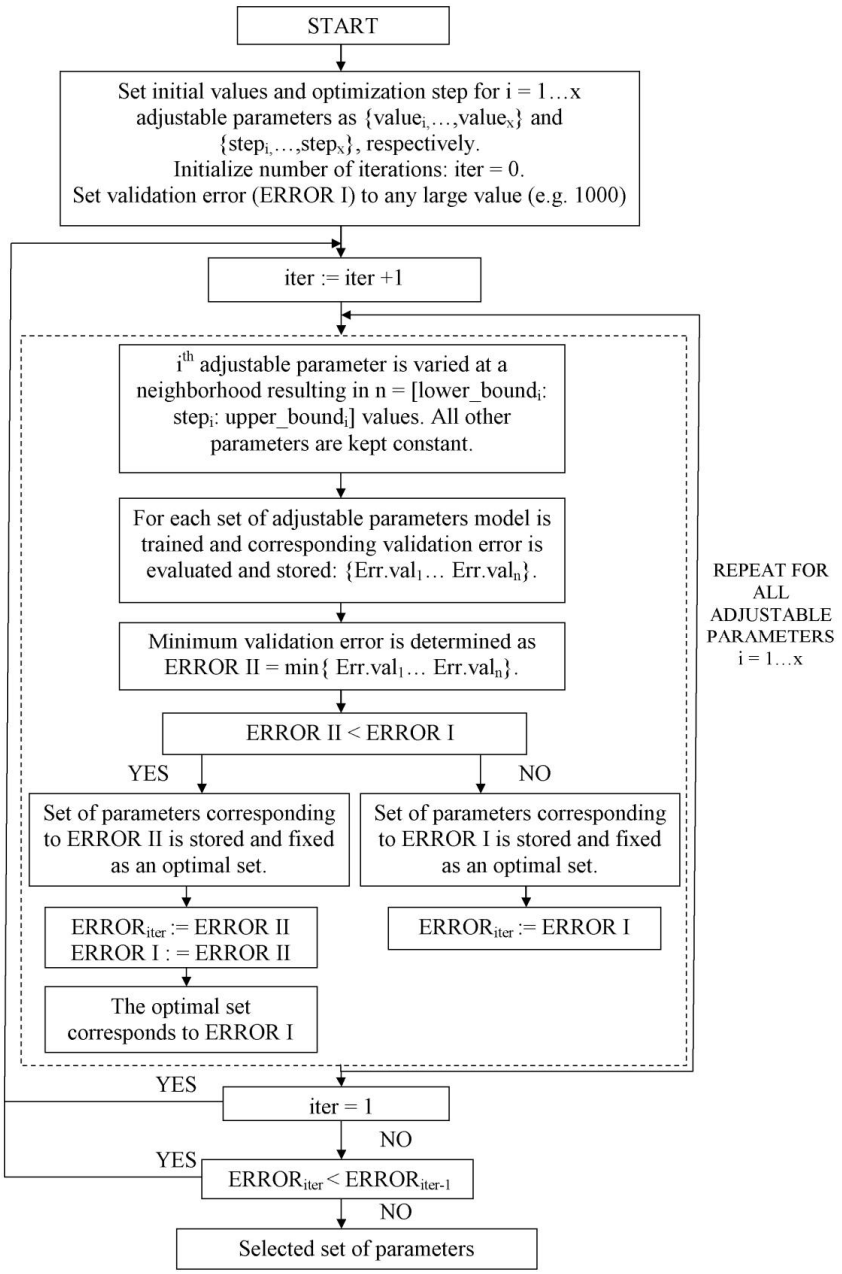

Figure 5.1. Search procedure.

5.2.1 Tuning NN parameters

The proposed search procedure is implemented within a forecasting model to obtain the optimal input set and tune the model parameters to predict prices in the Finnish day-ahead energy market on a single day.

To take into account the short-run trend and the seasonal patterns (daily and weekly) of electricity price time series, historical price data of about one week before a forecast day are considered to construct a set of candidate inputs. If the period of the study is extended further, the results are not affected seriously, that is, the relation of the current price with the price of much more than one week ago is very small (Vahidinasab et al., 2008). Finally, there is a certain number of candidate inputs for target price at hour *h* ($price_h$) of the forecasting day: {$price_{h-1}$,…, $price_{h-200}$}.

Now, the training period must be selected. If the functional relationships of a signal vary slowly with time, a long history of the signal can be considered as the training period resulting in a large number of training samples. However, as mentioned in Section 3.4.2, the market conditions evolve with time and, hence, the use of a long training period for the price forecasting model may result in significant inaccuracies. On the other hand, if the training period is selected to be very short, a forecasting model cannot derive functional relationships of prices because of the small number of training samples. In Section 3.4.2 the period of 60 days has been selected to train the forecasting model. The last 50 days have been proposed as the training period for the electricity day-ahead price forecasting by other researchers (Conejo et al., 2005a; Conejo et al., 2005b; Vahidinasab et al., 2008; Shafie-khan, 2011). The training period of 50 days is used here to train the NN. Thus, the historical prices of the Finnish day-ahead energy market over the period from 8 Sep 2008 to 5 Nov 2008 are used to predict prices on a single day, 6 Nov 2008. Validation data set is extracted from the training period as a day before the forecasting day, that is, 5 Nov 2008 (see Figure 5.2).

Linear correlation is used as a relevance measure within the proposed relevance-redundancy feature selection algorithm. Within the proposed search procedure, the NN is applied having a different number of neurons (N_h:{1,2,…,40}) in the hidden layer, while the remaining parameters (V_1, V_2) are kept constant (see Figure 5.3). Here, a wide neighborhood range of N_h is used for more clarity. For the neurons numbered from 1 to 4, the AMAPE for the validation set has the smallest values. For the neurons numbered higher than 5, the behavior of the NN is unstable. The proposed number of neurons to be used is selected as three.

Similarly, the values of V_1 and V_2 are varied in turn. In Figure 5.4, the set of V_1 is {0.40…0.80} and the set of V_2 is {0.60…0.95}. The number of neurons is kept constant. It can be seen that the best result is given for $V_1 = 0.61$ and $V_2 = 0.85$ when the AMAPE on the validation set is equal to 2.45%.

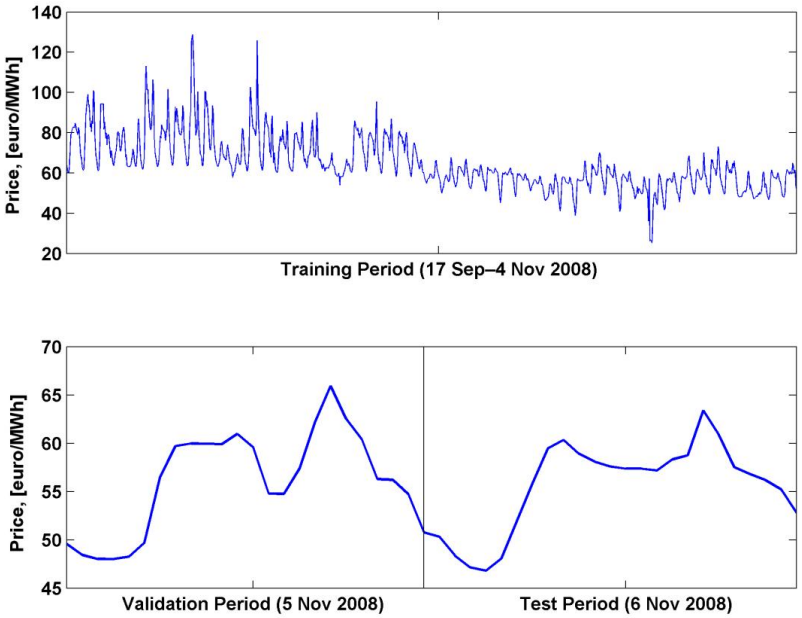

Figure 5.2. Training (top), validation, and test (bottom) periods for the NN.

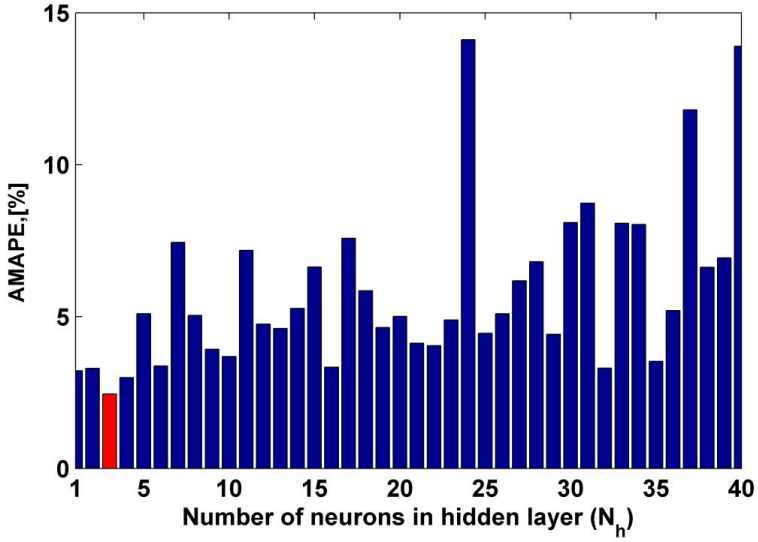

Figure 5.3. Validation error (AMAPE) versus N_h. V_1 and V_2 are kept constant.

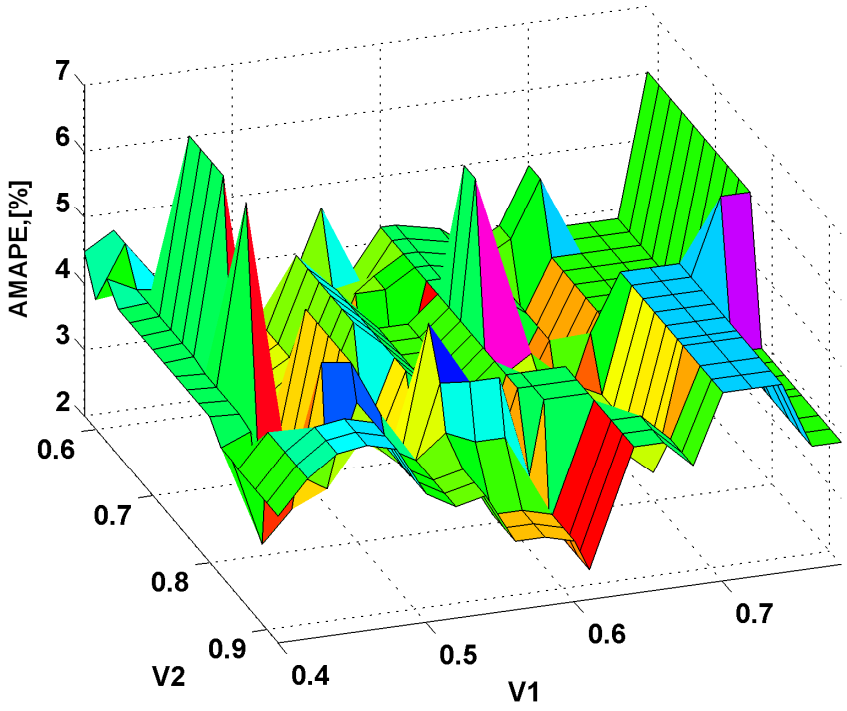

Figure 5.4. Validation error (AMAPE) versus V_1 and V_2. N_h is kept constant.

For a better illustration of the relevance-redundancy filtering process, its sample results are presented in Appendix F (see Figure F.1 and Table F.1-2).

The finally retained inputs are $price_{h-1}$, $price_{h-3}$, $price_{h-24}$, $price_{h-26}$, $price_{h-48}$, $price_{h-72}$, $price_{h-96}$, $price_{h-144}$, $price_{h-168}$, and $price_{h-192}$. These retained inputs indicate a short-run trend ($price_{h-1}$), daily periodicity ($price_{h-24}$), and weekly periodicity ($price_{h-168}$).

The original and predicted price curves obtained from the forecasting model are presented for a single test day, 6 Nov 2008, of the Finnish day-ahead energy market in Figure 5.5. As can be seen, the forecast curve acceptably follows the actual one. The corresponding AMAPE is 2.02%.

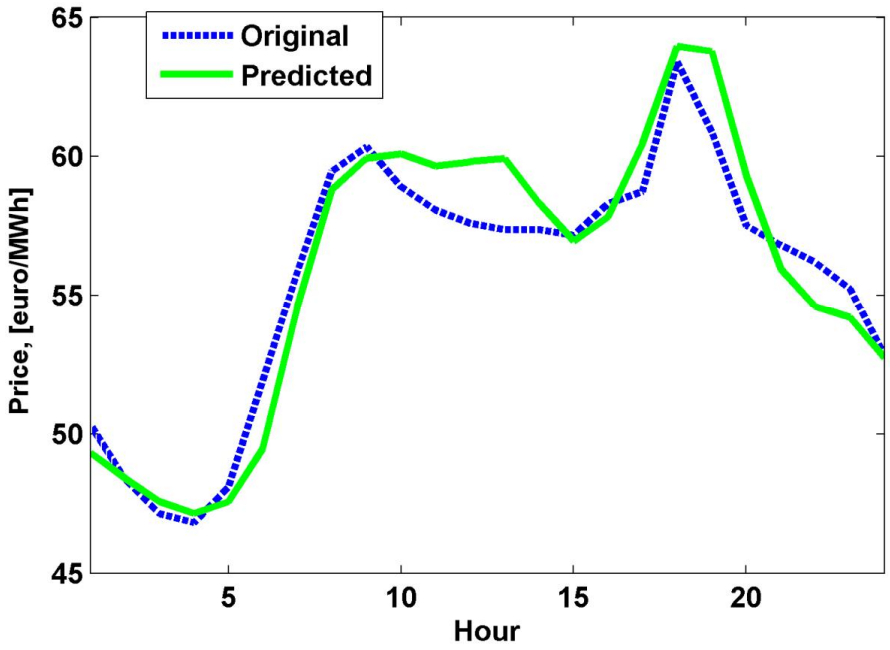

Figure 5.5. Original and predicted prices for a single test day 6 Nov 2008.

5.2.2 Linear and nonlinear feature selection techniques

The forecasting performance of different forecasting engines applying different feature selection techniques is examined in this section. Four test weeks corresponding to the four seasons of the years 2009–2010 from the Finnish day-ahead energy market are considered. The four weeks, 15 Feb to 21 Feb, 15 May to 21 May, 15 Aug to 21 Aug, 15 Nov to 21 Nov (months 2, 5, 8 and 11), are approximately in the middle week of each season. Therefore, in the context of this study, it is assumed that these week results are reasonably accurate for a study spanning one whole year. This way, representative results for the whole certain year are provided (Conejo et al., 2005b; Catalão et al., 2007; Amjady and Keynia, 2008; Vahidinasab et al., 2008; Shafie-khah et al., 2011). Three different forecast engines merged with five feature selection techniques and their combinations are examined. The selected forecasting engines are three-layered NN, Relevance Vector Machine (RVM), and Random Forest (RF) regression (see Appendix G, Section G.1 and G.2). These engines are chosen because they have been previously used in several other applications with promising results (Zhang and Qi, 2005; Catalão et al., 2007; Nelson et al., 1999; Meng et al., 2009; Niu et al., 2009; Mori and Awata, 2007; Breiman, 1984). For the sake of a fair comparison, historical price data up to 200 hours before a forecast hour are considered to construct a set of candidate inputs, and

the last 50 days before a forecasting day are used as a training period for each forecasting engine.

Heuristic method is among the examined feature selection techniques. It implies selection of specific inputs for a forecasting engine with a criterion based on past experience in the study (Contreras et al., 2003; Nogales, 2002). In this case, previous price values indicating trend (lags 1–2), daily (lags 24, 48, 72), and weekly seasonality (lag 168) are directly passed to a forecasting engine.

AMAPE values of the NN model producing price forecasts for the years 2009–2010 are presented in Tables 5.1–5.2. The respective AMAPE values for the RVM and RF are given in Appendix G (see Section G.3).

The following observations can be seen from Tables 5.1–5.2:

- The feature selection techniques consisting of both relevance and redundancy filters outperform the feature selection techniques having only a relevance filter for all the considered forecasting engines.
- The models applying nonlinear feature selection techniques (i.e. MI, Relief, KNN) perform better than the models merged with linear feature selection techniques (i.e. linear correlation, sequential feature selection based linear regression) when the prices of the year 2010 are predicted. The performance of these models is close enough for the price data of the year 2009, since the price series of the year 2009 has a relatively lower volatility with respect to the prices of the year 2010. Therefore, the superiority of the feature selection based on a nonlinear technique becomes obvious when it is implemented on a highly volatile data, and an ability to capture the nonlinear behavior of given prices is required.
- The heuristic approach indicates the worst performance among the examined feature selection techniques.

The above-mentioned statements are justified for the RVM and RF forecasting engines (see Appendix G, Section G.3).

Historical electricity demand data (up to 200 hours before a forecast hour) are put into a set of candidate inputs and fed to the feature selection engine of the NN to model and predict the price data of the four weeks. Inclusion of historical electricity demand data in a set of candidate inputs results in a significant improvement in the forecasting performance of the NN (see Tables 5.3–5.4).

Table 5.1. AMAPE (%) for the price forecast of the four weeks of the Finnish day-ahead energy market in the year 2009 produced by the NN employing different feature selection techniques. Input data: historical prices.

	Feature selection technique				
Test week	Linear Corr./ —	Relief/ —	MI/ —	KNN/ —	Sequential selection
Fall	5.63	5.12	5.42	5.85	5.72
Summer	10.65	11.04	10.34	10.61	11.75
Spring	6.10	5.57	6.05	5.57	8.68
Winter	6.81	7.06	6.45	6.94	7.69
Average	7.30	7.19	7.07	7.24	8.46
	Linear Corr./ Linear Corr.	MI/ MI	Relief/ Linear Corr.	Relief/ MI	Heuristic
Fall	4.75	4.90	4.63	4.63	5.59
Summer	9.43	9.11	10.55	10.42	10.86
Spring	5.45	5.57	5.73	5.08	8.34
Winter	4.70	4.83	5.19	5.10	7.30
Average	6.08	6.10	6.52	6.30	8.02

Table 5.2. AMAPE (%) for the price forecast of the four weeks of the Finnish day-ahead energy market in the year 2010 produced by the NN employing different feature selection techniques. Input data: historical prices.

	Feature selection technique				
Test week	Linear Corr./ —	Relief/ —	MI/ —	KNN/ —	Sequential selection
Fall	3.33	3.68	3.28	2.91	6.16
Summer	6.53	5.97	6.31	5.48	5.49
Spring	9.01	8.88	8.73	8.47	7.67
Winter	21.74	18.72	17.96	19.88	19.81
Average	10.15	9.31	9.07	9.19	9.78
	Linear Corr. / Linear Corr.	MI/ MI	Relief/ Linear Corr.	Relief/ MI	Heuristic
Fall	2.42	3.01	3.37	2.92	5.18
Summer	5.51	4.64	5.32	5.35	7.73
Spring	8.13	7.59	6.99	7.89	8.86
Winter	15.32	14.54	17.87	15.22	21.54
Average	7.85	7.45	8.38	7.65	10.83

Table 5.3. AMAPE (%) for the price forecast of the four weeks of the Finnish day-ahead energy market in the year 2009 produced by the NN employing different feature selection techniques. Input data: historical prices and demand.

	Feature selection technique				
Test week	Linear Corr./ —	Relief/ —	MI/ —	KNN/ —	Sequential selection
Fall	5.43	5.16	5.25	5.63	5.46
Summer	9.99	10.47	9.67	10.03	10.93
Spring	5.64	7.68	5.79	6.05	6.15
Winter	6.05	4.60	6.24	5.23	4.88
Average	6.78	6.98	6.74	6.74	7.11
	Linear Corr./ Linear Corr.	MI/ MI	Relief / Linear Corr.	Relief/ MI	Heuristic
Fall	4.64	4.52	4.98	4.65	5.51
Summer	9.58	9.69	10.02	9.79	10.48
Spring	5.31	5.30	6.14	5.98	7.37
Winter	4.58	5.01	4.98	5.01	5.99
Average	6.03	6.13	6.53	6.33	7.34

Table 5.4. AMAPE (%) for the price forecast of the four weeks of the Finnish day-ahead energy market in the year 2010 produced by the NN employing different feature selection techniques. Input data: historical prices and demand.

	Feature selection technique				
Test week	Linear Corr./ —	Relief/ —	MI/ —	KNN/ —	Sequential selection
Fall	3.36	3.30	3.03	3.79	3.44
Summer	6.53	5.98	5.29	5.60	5.81
Spring	6.77	6.35	6.84	7.01	10.57
Winter	20.21	17.31	15.77	16.05	16.84
Average	9.22	8.24	7.73	8.12	9.17
	Linear Corr./ Linear Corr.	MI/ MI	Relief / Linear Corr.	Relief/ MI	Heuristic
Fall	2.77	2.34	3.01	2.58	3.09
Summer	5.15	3.25	4.09	3.47	6.27
Spring	6.47	6.20	7.66	6.40	8.37
Winter	15.54	13.57	15.10	14.55	22.38
Average	7.48	6.34	7.47	6.75	10.03

5.3 Simultaneous forecasting electricity prices and demand

It is very useful for a market participant to be able to predict electricity demand and prices simultaneously because demand and prices are intertwined activities in electricity markets (Aggarwal, 2009). In many studies, it is assumed that actual values of electricity demand are known when day-ahead electricity prices are forecasted (Wu and Shahidehpour, 2010; Amjady and Keynia, 2008; Nogales et al., 2002), and therefore, the effect on the price forecast accuracy of uncertainty in the demand forecast has been ignored.

In the experiments given in Section 5.2.2, actual values of electricity demand were used to forecast day-ahead electricity prices. However, such a strategy is not suitable for a real-life problem as realized electricity demand values are not known for a day-ahead market at the moment when the prices are predicted. Some authors have used the projected demand of TSO of the concerned electricity market as an input variable, while a few have predicted the demand first and then used it as an input variable for the price-forecasting model (Georgolakis, 2006; Mandal et al., 2006).

In deregulated markets, the system demand may be significantly affected by electricity prices when consumers are encouraged to use less energy during peak hours. In other words, variability in the electricity price can influence the energy-use patterns of the consumers. As mentioned above, actual values of electricity demand were used when day-ahead prices were predicted. Moreover, the same assumption was used for electricity demand forecasting, when the electricity demand was dependent on actual values of prices rather than their predicted values (Yun et al., 2008; Niu et al., 2009).

The methodology presented in this section uses a hybrid wavelet transform combined with SARIMA and a three-layered NN to implement simultaneous demand and price forecasting processes in the day-ahead energy market. The proposed methodology is better adapted to actual conditions of an energy market since the forecast features for price and demand are not assumed known values but are predicted by the model, thus accounting for the interactions of the demand and price forecast processes (Voronin and Partanen, 2013b). The forecasting performance evaluation applied actual hourly data of the four weeks in the Finnish day-ahead energy market of the year 2009, corresponding to four seasons, as in the previous section.

5.3.1 Wavelet transform

When using classical statistical techniques, a stationary process is assumed for the data. For electricity demand and price time series, the assumption of stationarity usually has to be rejected. One of the ways to capture localized trending in the series is to apply models with time-varying parameters (Granger, 2008). Another way to deal with nonstationarity is the use of mathematical transformations of an initial series. In many cases, information that cannot be readily seen in the time domain can be obtained in the frequency domain. Fourier transform (FT) is probably the most popular transform and is

used in many different areas, including many branches of engineering. However, no time information is available in the Fourier transformed signal, in other words, it is not clear where the time specific spectral components appear. The short-time Fourier transform (STFT) gives time information by dividing the signal into small enough segments so that these segments of the signal can be assumed to be stationary. For this purpose, a window function is chosen. The width of this window must be equal to the segment of the signal where its stationarity is valid. Depending on the window length, STFT gives a poor time resolution and a good frequency resolution, or vice versa. The wavelet transform (WT) was developed as an alternative approach to STFT to overcome the resolution problem (Olkkonen, 2011). Implicitly, wavelets have a window that automatically adapts itself to give an appropriate resolution. The basic concept in the wavelet analysis begins with the selection of a proper wavelet (mother wavelet) and an analysis of its translated and dilated versions (Galli, 1996). A wavelet can be defined as a function $\psi(h)$ with a zero mean.

$$\int_{-\infty}^{+\infty} \psi(h)dh = 0 \tag{5.3}$$

A signal can be decomposed into many series of wavelets:

$$\psi_{a^*,b^*}(h) = \frac{1}{\sqrt{a^*}}\psi\left(\frac{h - b^*}{a^*}\right)dh \tag{5.4}$$

where the scale parameter a^* controls the spread (dilation) of the wavelet and the translation factor b^* determines its central position. Thus, the continuous WT w_{a^*,b^*} of the signal $f(h)$ with respect to a wavelet ψ_{a^*,b^*} is given by:

$$w_{a^*,b^*} = \frac{1}{\sqrt{a^*}}\int_{-\infty}^{+\infty} f(h)\psi\left(\frac{h - b^*}{a^*}\right)dh \tag{5.5}$$

A $w_{a,b}$ coefficient represents how well the original signal $f(h)$ and the scaled/translated mother wavelet match. Since the continuous WT is achieved by continuously scaling and translating the mother wavelet, substantial redundant information is generated. Therefore, instead of doing that, the mother wavelet can be scaled and translated using certain scales and positions usually based on powers of two (Conejo et al., 2005b; Reis and Alves, 2005). This scheme is more efficient and as accurate as the continuous WT. It is known as the discrete WT:

$$w_{l,m} = 2^{-(l/2)}\sum_{h=0}^{H-1} f(h)\psi\left(\frac{h - m \cdot 2^l}{2^l}\right) \tag{5.6}$$

where H is the length of the signal f. The scaling and translation parameters are functions of the integer variables l and m ($a=2^l$ and $b = m \cdot 2^l$), h is the discrete time index, and $w_{l,m}$ is the decomposition coefficient corresponding to l and m.

An efficient algorithm to implement the discrete WT using filters has been developed in (Mallat, 1989). Multiresolution via Mallat's algorithm is a procedure to obtain approximations (e.g. A1) and details (e.g. D1) from a given signal f. In the reconstruction stage, these components can be assembled back into the original signal f' (see Figure 5.6).

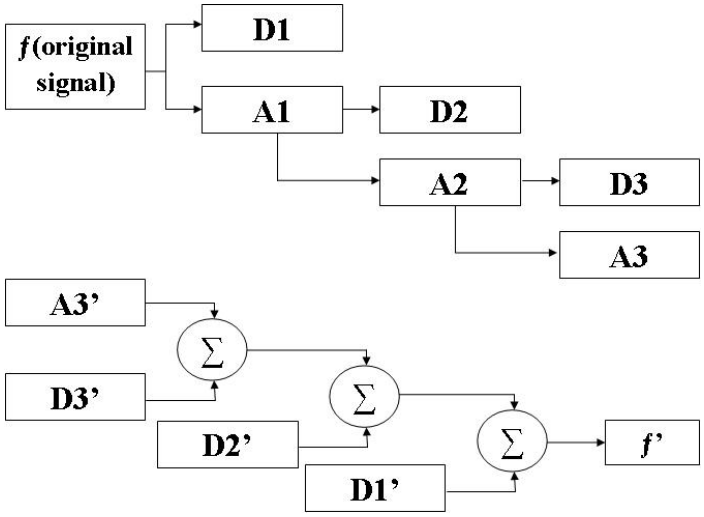

Figure 5.6. Multilevel decomposition (top) and reconstruction (bottom) processes.

Multilevel wavelet decomposition is applied to data preprocessing and considered as an alternative to the previously used time series decomposition technique (see Section 3.3.1). Depending on the selected resolution level, the time series signal is decomposed into a set of wavelet domain components. This set of components presents a better behavior (more stable variance and no outliers) than the original price series. Unlike classical time series decomposition, where deterministic patterns are projected to the future and used as forecasted values, the obtained wavelet components are more accurately predicted by the corresponding model.

Hereafter, a Daubechies wavelet of order 5 is used as the mother wavelet to transform the price and demand series into several wavelet subseries. This wavelet offers an appropriate trade-off between wavelength and smoothness, resulting in an appropriate behavior for the price and demand forecast. Similar wavelets have been considered in previous studies (Conejo et al., 2005b; Tan et al., 2010).Three decomposition levels are considered, since this describes the price series in a more thorough and meaningful way (Conejo, et al., 2005b). Thus, each of the original price and demand series is

decomposed and reconstructed into one approximation subseries (general trend component) and three detail subseries (high-frequency components).

5.3.2 Forecasting time framework

The time framework to simultaneously forecast electricity prices and demand in the day-ahead energy market of Nord Pool Spot is illustrated in Figure 5.7 and explained below. As mentioned above, the market day-ahead electricity price forecast for day D is required on day D-1. Actual day-ahead price data up to 24 hours of day D-1 are published by the TSO and available on day D-2. However, actual demand data for day D-1 are not available on day D-2.

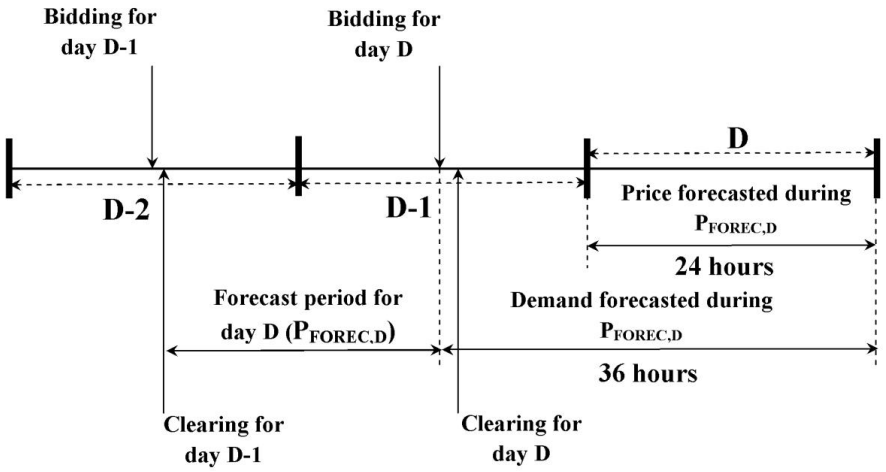

Figure 5.7. Time framework to forecast market prices and demand in the Finnish day-ahead energy market.

Therefore, when bidding for day D (hour 12 of day D-1), day-ahead price data up to hour 24 of day D-1 are considered known while demand data are available only up to hour 12 of day D-1. As a result, the actual forecasts of market day-ahead prices and demand for day D can take place between the clearing hour for day D-1 of day D-2 and the bidding hour for day D of day D-1. At least a 36 hours ahead (12 hours of day D-1 plus 24 hours of day D) demand forecast is required to predict prices 24 hours ahead when bidding for day D.

5.3.3 Forecasting strategy

WT deals with nonstationarity by decomposing the price and demand series into less volatile components. A linear SARIMA and a three-layered NN are combined to capture different aspects of the underlying linear and nonlinear patterns of the wavelet

subseries. The SARIMA model incorporates the cyclicality of the series, which clearly exhibits hourly and weekly patterns and produces initial day-ahead forecasts for all wavelet subseries of demand and prices. The proposed relevance-redundancy feature selection algorithm is performed for the feature selection of each wavelet subseries. The NN uses the selected inputs to forecast the demand and prices of the next hours.

The proposed simultaneous forecast strategy can be summarized by the following step-by-step algorithm, shown also in Figure 5.8:

1) Electricity price and demand series are decomposed by WT into approximation subseries (A3) and three detail subseries (D3, D2, D1).
2) WT+SARIMA models are built to forecast the future values of the price and demand wavelet subseries.
3) The set of candidate inputs for each subseries is constructed, including lagged and predicted features of both the wavelet and time domains. Although the wavelet components are obtained by the decomposition of the price and demand signals, the past values of the original price and demand series are considered among the candidate inputs of each wavelet component, since it is still possible that some characteristics of the price and demand signals are better highlighted in the original time domain (Amjady, 2008). Taking into account the short-run trend, and daily and weekly periodicity of the electricity and demand time series, their lagged values up to about one week are considered among the candidate inputs. Finally, the candidate inputs for each subseries of demand and price include lagged values of these subseries, original price or demand lagged up to 200 hours before a forecast hour, and price and demand values of these subseries forecasted by the WT+SARIMA model. For the sake of clarity, prices and demand wavelet components predicted by the WT+SARIMA model are additionally indexed as "SARIMA". For instance, the approximation price wavelet component value predicted by the WT+SARIMA at hour *h* is denoted $A3_{SARIMA\_price,h}$. There are the 602 candidate inputs to predict the approximation price wavelet subseries at hour *h* ($A3_{price,h}$):{$A3_{SARIMA\_price,h}$, $A3_{price,h-1}$,…, $A3_{price,h-200}$, $price_{h-1}$,…, $price_{h-200}$, $A3_{SARIMA\_demand,h}$, $A3_{demand,h-1}$,…, $A3_{demand,h-200}$}.
4) An iterative search procedure introduced in Section 5.2 is carried out. The procedure automatically adjusts V_1 and V_2 of the relevance-redundancy feature selection algorithm and N_h of the NNs for each subseries in order to minimize the forecasting error on a validation data set.
5) Given adjusted V_1 and V_2 values, the inputs are selected. With the selected N_h, the NNs are trained by their respective training samples and separately predict the price and demand subseries of the next hours.

In multistep ahead prediction, the predicted price and demand values of the current step are used to determine their values in the next step up to hour 24 of the forecasting day.

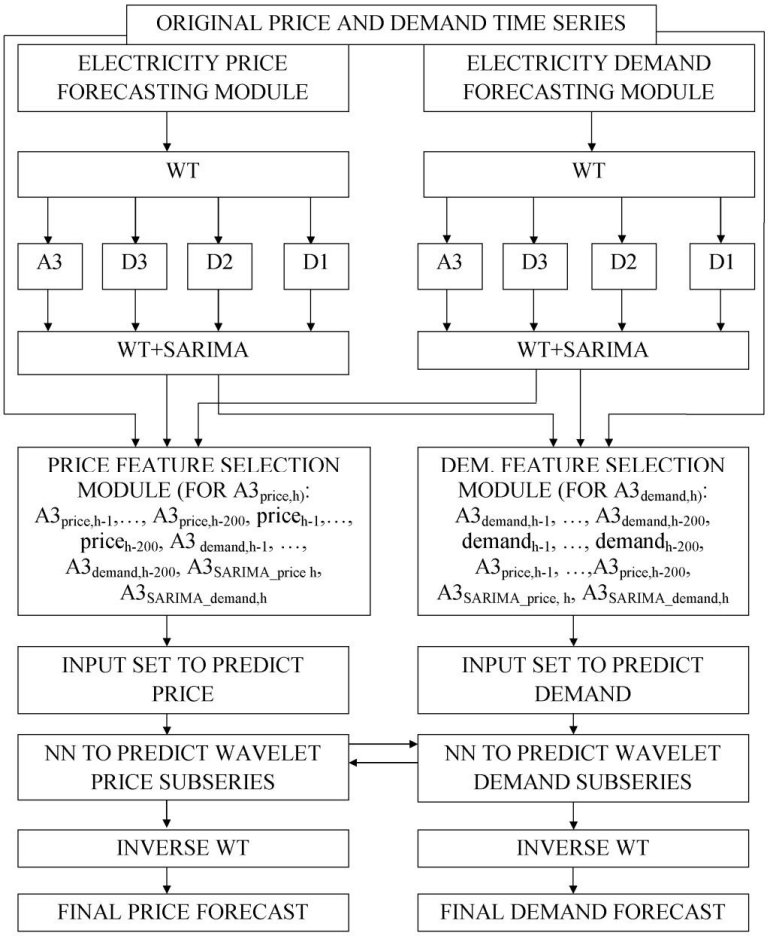

Figure 5.8. Flowchart of the proposed forecasting methodology.

5.3.4 **Training phase**

To commence the training phase of the proposed forecast strategy, the training periods of the WT+SARIMA and NN models are first determined. Intervals of 50 days (1200 h) for both WT+SARIMA and NN are considered resulting in two consequent equal training periods.

First, a day denoted by D is considered in the training period of the NN (second training interval) and the values of price wavelet subseries for this day are assumed unknown. In the case of demand forecasting, the values of the demand wavelet subseries for day D plus 12 after bidding hours of day D-1 are assumed unknown. The WT+SARIMA models are trained by the historical data of the 50 days preceding hour 1 of day D or

bidding hour of day D-1 and predict hourly price and demand subseries of day D, respectively. To improve the performance of the WT+SARIMA forecast process for each day of the second training period (D = 1,…,50), the WT+SARIMA models are trained in a moving fashion, that is, by the immediately previous 50-day period. This process is repeated until price and demand subseries forecasts of the WT+SARIMA models are obtained for all days of the training period of the NN model. The selected samples of the validation period (i.e., the 24 hours before the forecast day for the price forecasting and the 36 hours before the forecast horizon of the demand forecasting) are removed from the training set of the NN model. Then, the NN model is trained by the remaining selected samples, and the hourly demand and price values of the forecast day are predicted.

The plots of price, demand, and their wavelet components A3, D3, D2, and D1 in one of the four time intervals used in the study are shown in Figs. 5.9–5.11, respectively.

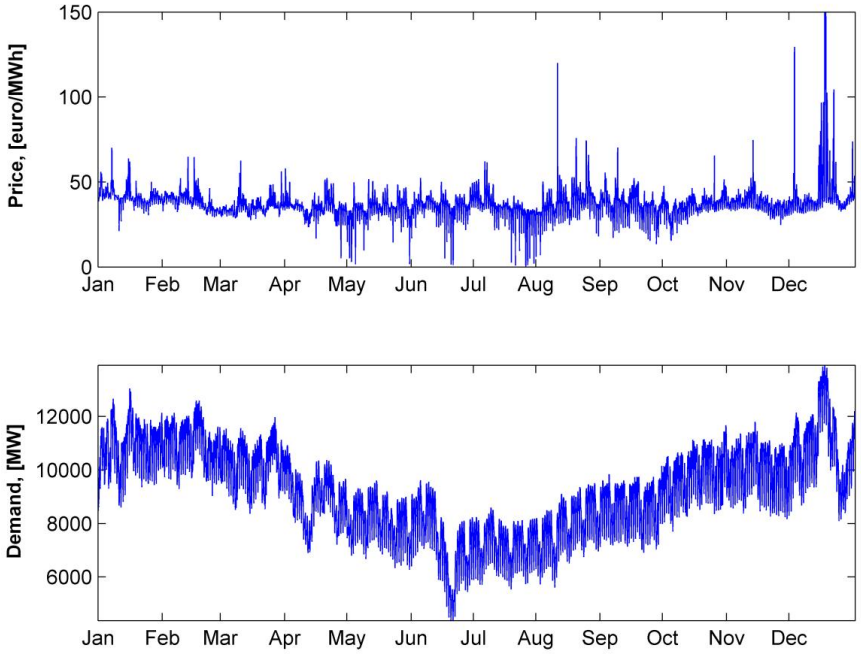

Figure 5.9. Plot of price (top) and demand (bottom) for the Finnish day-ahead energy market of the year 2009.

For each day-ahead forecast over the test week, the overall time interval consists of a forecast day, the first training period for the WT+SARIMA, and the second training period (with a validation interval) for the NN (see Figures 5.10–5.11).

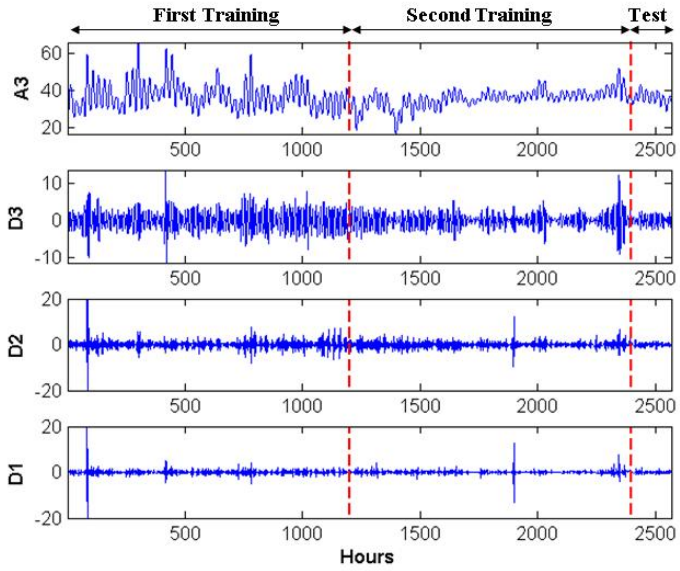

Figure 5.10. Plot of the price components A3, D3, D2, and D1 for the first forecast day of the fall test week (from 15 Nov to 21 Nov 2009).

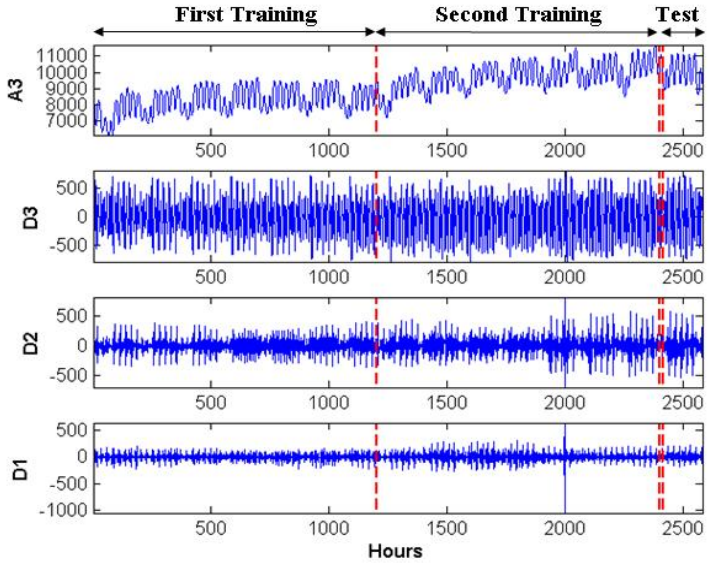

Figure 5.11. Plot of the demand wavelet components A3, D3, D2, and D1 for the first forecast day of the fall test week (from 15 Nov to 21 Nov 2009).

5.3.5 **Numerical results**

The results obtained of the relevance-redundancy linear correlation based feature selection algorithm for price and demand forecasting for the first fall test day, that is, 15 Nov 2009 are presented in Tables 5.5–5.6. The values selected for N_h, V_1, and V_2 of each wavelet component for the first day of the fall test week are shown in the second, third, and fourth columns of Tables 5.5–5.6. From the obtained results, the forecast features (for both price and demand) produced by the WT+SARIMA model are always among the selected features for the A3 and D3 subseries of price and demand. For the approximation components of price and demand, inputs indicating the effect of short-run trend ($A3_{price,h-1}$, $A3_{demand,h-1}$), daily periodicity ($A3_{price,h-25}$, $A3_{demand,h-23}$), and weekly periodicity ($A3_{price,h-167}$, $A3_{demand,h-167}$, $price_{h-169}$, $demand_{h-167}$) can also be seen from the obtained results.

Table 5.5–5.6 also show that the dependency of the price and demand wavelet components on the exogenous variables decreases from A3 to D1.

Table 5.5. Inputs selected by the two-step feature selection analysis for the four wavelet price components (the first day of the fall test week).

Variable	N_h	V_1	V_2	Selected candidates
$A3_{price,h}$	6	0.63	0.83	$A3_{SARIMA\_price,h}$, $A3_{SARIMA\_demand,h}$, $A3_{price,h-1}$, $A3_{price,h-2}$, $A3_{price,h-6}$, $A3_{price,h-21}$, $A3_{price,h-25}$, $A3_{price,h-165}$, $A3_{price,h-167}$, $A3_{price,h-169}$, $A3_{price,h-170}$, $A3_{demand,h-4}$, $A3_{demand,h-22}$, $A3_{demand,h-142}$, $A3_{demand,h-146}$, $A3_{demand,h-167}$, $A3_{demand,h-170}$, $A3_{demand,h-172}$, $A3_{demand,h-194}$, $price_{h-169}$, $price_{h-170}$
$D3_{price,h}$	9	0.59	0.84	$D3_{SARIMA\_price,h}$, $D3_{SARIMA\_demand,h}$, $D3_{price,h-1}$, $D3_{price,h-23}$, $D3_{price,h-144}$, $D3_{price,h-167}$, $D3_{price,h-168}$, $D3_{price,h-169}$, $D3_{price,h-192}$, $D3_{demand,h-144}$, $D3_{demand,h-169}$
$D2_{price,h}$	5	0.47	0.75	$D2_{SARIMA\_price,h}$, $D2_{price,h-1}$, $D2_{price,h-144}$, $D2_{price,h-168}$, $D2_{price,h-192}$
$D1_{price,h}$	11	0.16	0.85	$D1_{SARIMA\_price,h}$, $D1_{price,h-3}$, $D1_{price,h-24}$, $D1_{price,h-48}$, $D1_{price,h-72}$, $D1_{price,h-96}$, $D1_{price,h-120}$

The results obtained of the relevance-redundancy correlation based feature selection for the approximation subseries of price and demand for the first day of the fall test week, that is, 15 Nov 2009 can be found in Appendix H (see Section H.1).

Table 5.6. Inputs selected by the two-step feature selection analysis for the four wavelet demand components (the first day of the fall test week).

Variable	N_h	V_1	V_2	Selected candidates
$A3_{demand,h}$	11	0.53	0.63	$A3_{SARIMA\_demand,h}$, $A3_{SARIMA\_price,h}$, $A3_{demand,\ h-1}$, $A3_{demand,h-3}$, $A3_{demand,h-4}$, $A3_{demand,h-20}$, $A3_{demand,h-48}$, $A3_{demand,h-94}$, $A3_{demand,h-140}$, $A3_{demand,h-167}$, $A3_{price,h-1}$, $A3_{price,h-3}$, $A3_{price,h-142}$, $A3_{price,h-144}$, $A3_{price,h-145}$, $A3_{price,h-167}$, $demand_{h-146}$, $demand_{h-163}$, $demand_{h-167}$
$D3_{demand,h}$	9	0.56	0.70	$D3_{SARIMA\_demand,h}$, $D3_{SARIMA\_price,h}$, $D3_{demand,h-13}$, $D3_{demand,h-37}$, $D3_{demand,h-71}$, $D3_{demand,h-97}$, $D3_{demand,h-108}$, $D3_{demand,h-121}$, $D3_{demand,h-168}$, $D3_{price,h-23}$, $D3_{price,h-47}$, $D3_{price,h-71}$, $D3_{price,h-156}$, $D3_{price,h-179}$
$D2_{demand,h}$	9	0.71	0.80	$D2_{SARIMA\_demand,h}$, $D2_{\ demand,h-36}$, $D2_{\ demand,h-48}$, $D2_{\ demand,h-120}$, $D2_{\ demand,h-156}$, $D2_{\ demand,h-180}$
$D1_{demand,h}$	5	0.68	0.80	$D1_{SARIMA\_demand,h}$, $D1_{demand,h-1}$, $D1_{demand,h-24}$

In order to illustrate graphically the accuracy of the price and demand forecasts of the proposed strategy, the forecasts and actual signals for the four test weeks of the Finnish day-ahead energy market of the year 2009 are shown in Figures 5.12–5.15. As can be seen, the forecast curves acceptably follow the actual curves of both prices and demand for all the four test weeks.

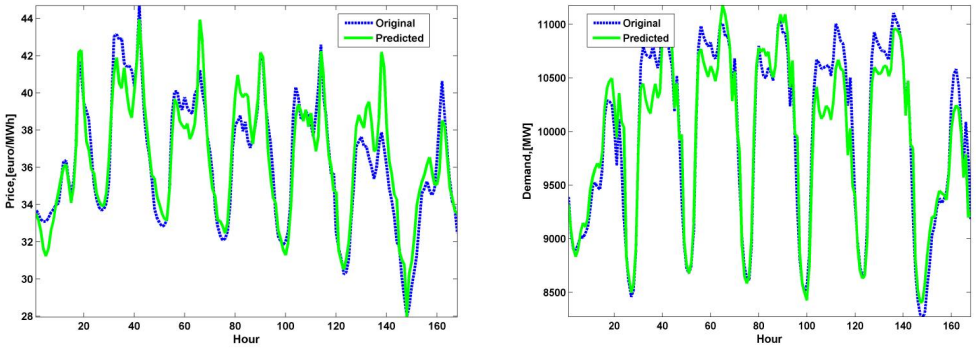

Figure 5.12. Original and predicted price (left) and demand (right) curves for the fall test week of the Finnish day-ahead energy market of the year 2009.

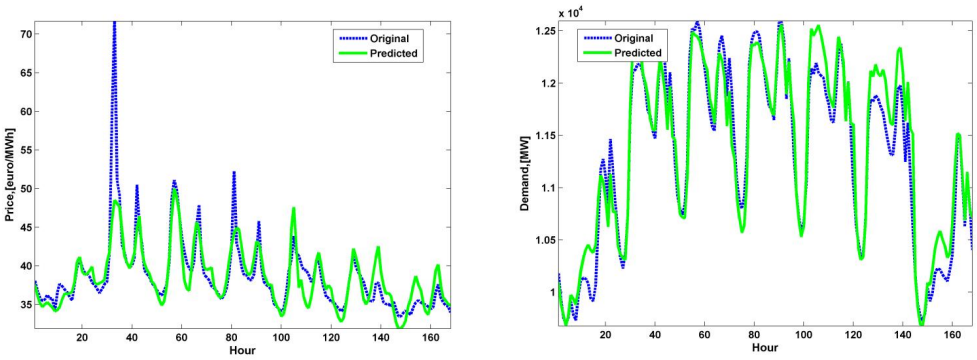

Figure 5.13. Original and predicted price (left) and demand (right) curves for the winter test week of the Finnish day-ahead energy market of the year 2009.

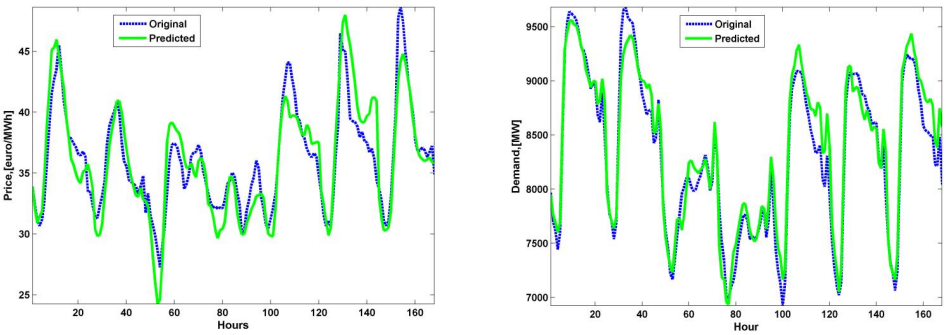

Figure 5.14. Original and predicted price (left) and demand (right) curves for the spring test week of the Finnish day-ahead energy market of the year 2009.

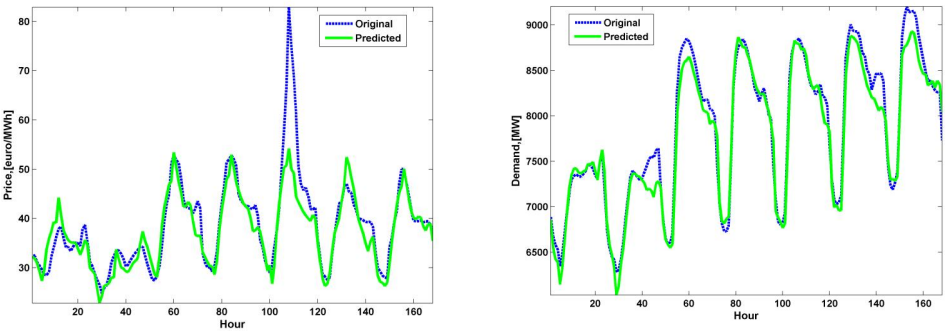

Figure 5.15. Original and predicted price (left) and demand (right) curves for the summer test week of the Finnish day-ahead energy market of the year 2009.

Only a few studies have considered price and demand forecasting in the Finnish day-ahead energy market, and it was not possible to find price and demand forecast methods considering the above-mentioned four test weeks in the literature. Therefore, the proposed method is compared with some of the most recent price and demand forecast techniques applied to case studies of energy markets of other countries: SARIMA (Taylor, 2006; Contreras et al., 2003; Nogales et al., 2002), WT+SARIMA (Conejo et al., 2005b; Tan, 2010), NNs with different training algorithms (Taylor, 2006, Bhattacharyya and Thanh, 2003; Cavallaro, 2005; Nasr et al., 2001; Mandal et al., 2007; Szkuta, 1999; He and Bo, 2009), and WT+NN (Shafie-khah , 2011, Chen et al.,2010).

AMAPE values of the proposed method and five other forecast techniques for the four weeks corresponding to the four seasons of the year 2009 in the Finnish day-ahead energy market are presented in Table 5.7. Particle Swarm Optimization (PSO) is another learning algorithm for NNs. As can be seen from the tables, the NN(PSO) model results are close to those of NN(LM). A detailed mathematical description of the algorithm to train the NN by the PSO method can be found in (He and Bo, 2009). In the WT+NN(LM) model, WT is used to decompose the price and demand series into less volatile components; separate NNs with the LM learning algorithm are applied for each component. For a fair comparison, NN(LM), NN(PSO), and WT+NN(LM) have a feature selection analysis based on the proposed two-step feature selection. The target variables of the feature selection technique for the NN(LM) and the NN(PSO) models are the original price and demand signals, respectively, while the feature selection technique is executed for each wavelet component in the WT+NN(LM) model and in the proposed method. It also should be noted that in the set of candidate inputs of the alternative models examined, no variables are predicted beforehand by the models.

Table 5.7. AMAPE in percent (%) for the price/demand forecast of the four test weeks of the Finnish day-ahead energy market in the year 2009.

Test week	SARIMA	WT+ SARIMA	NN(LM)	NN(PSO)	WT+ NN (LM)	Proposed method
Winter	5.19/1.60	4.27/1.55	4.70/2.45	5.25/3.09	5.16/1.81	3.93/1.19
Spring	5.76/3.34	4.69/2.31	5.45/2.57	6.01/3.35	4.85/2.36	4.17/1.98
Summer	13.08/2.08	7.49/1.65	9.43/3.42	11.05/3.99	9.13/2.22	6.81/1.89
Fall	5.83/1.93	3.28/1.76	4.75/3.29	5.87/3.97	4.30/2.14	3.01/2.09
Aver.	7.47/2.24	4.93/1.82	6.08/2.93	7.05/3.60	5.86/2.13	4.48/1.79

As seen from Table 5.7, on the basis of the AMAPE values, the proposed strategy outperforms the other examined methods in all four test weeks. Table 5.7 shows that for the demand forecast, the WT+SARIMA model has lower AMAPE values than the proposed strategy in the summer and fall test weeks. However, the average of the AMAPE values of the proposed strategy is lower than that of all other techniques (indicated in the last row of Table 5.7). The accuracy improvement of the proposed method for price prediction with respect to SARIMA, WT+SARIMA, NN(LM),

NN(PSO), and WT+NN(LM) in terms of average AMAPE is 40.03% (1-4.48/7.47), 9.13% (1-4.48/4.93), 26.32% (1-4.48/6.08), 36.45% (1-4.48/7.05), and 23.55% (1-4.48/5.86), respectively. The corresponding improvement in the average AMAPEs for demand forecasting is 20.09% (1-1.79/2.24), 1.65% (1-1.79/1.82), 38.91% (1-1.79/2.93), 50.27% (1-1.79/3.60), and 15.96% (1-1.79/2.13).

From the results presented in Table 5.7, it can be seen that the use of WT decomposition results in an improvement in the model accuracy. These improvements for SARIMA in comparison with WT+SARIMA in terms of average AMAPE are 34.00% (1-4.93/7.47) and 18.75% (1-1.82/2.24) for the price and demand forecasts, respectively. For NN(LM) in comparison with WT+NN(LM) these values are 3.62% (1-5.86/6.08) and 27.30% (1-2.13/2.93) for price and demand forecasts, respectively.

The results for both price and demand forecasts also confirm the efficiency of the hybrid methodology with linear and nonlinear modeling capabilities. Furthermore, it should be noted that the results given in Table 5.7 show that the performance of models based only on a nonlinear framework was worse compared with the ARMA-based models. A possible explanation could be that the certain characteristics of the initial demand time are more linear than those of the price time series.

To demonstrate the efficiency of the proposed methodology over a longer period, a detailed representation of the performance of the price and demand forecast strategy for all the weeks of 2009 is shown in Appendix H (see Section H.2).

The running time required for the setup of the proposed simultaneous price and demand forecasting strategy including the training and prediction phases of WT+SARIMA to forecast price and demand for each day of the second training interval (50 days), the relevance-redundancy feature selection processes, the tuning of the adjustable model parameters, the training of the NNs, and the generation of price and demand forecasts for the first forecasting day is about 11 h 40 min on the personal computer with an Intel Core i5 2.40 GHz processor and 3.24 GB RAM. For the next forecast days, the total computation time for the training of the proposed strategy and the generation of price and demand forecasts 24 and 36 hours ahead, respectively, is about one hour since the price and demand predictions generated by WT+SARIMA become available. Therefore, the running time of the proposed strategy is considered sensible (except for the first forecast day) for day-ahead energy market operation. The overall average running times for SARIMA, WT+SARIMA, NN(LM), NN(PSO), WT+NN(LM) to generate a price or demand prediction for the forecast day are about 3 min, 7 min, 8 min, 10 min, and 23 min measured on the same hardware. All computer codes are provided by the MATLAB and R software packages. As can be seen, the running time to set up the competitive methods is lower than the setup of the proposed strategy. However, the prediction accuracy is a crucial concern for a forecasting method (as far as the computation time is reasonable).

5.3.6 Summary

The methodology consisting of SARIMA and NN frameworks is able to explain intermittent high volatility in prices by incorporating the effect of demand pressure. Moreover, frequency components obtained by WT are separately predicted. Such a strategy was supposed to improve an overall forecasting performance and, in particular, spikes in the series since there is a high correlation between price spikes and high-frequency wavelet components of the price signal spectrum. The proposed methodology generally outperforms other alternative forecasting methods because of its ability to capture different essential features of the given time series and incorporate interactions between demand and price forecasting processes being better adapted to the actual conditions of the energy market.

The methodology can produce acceptable results over a longer period of a calendar year. However, the methodology that typically predicts normal price behavior fairly well does not capture anomalous behavior (when prices increase rapidly and unexpectedly) to the full. The drawback of the proposed methodology can be clearly observed in Figures 5.13 and 5.15, and additionally, in Figure 5.16 where the predicted and actual prices of the selected weeks of the year 2009 are presented.

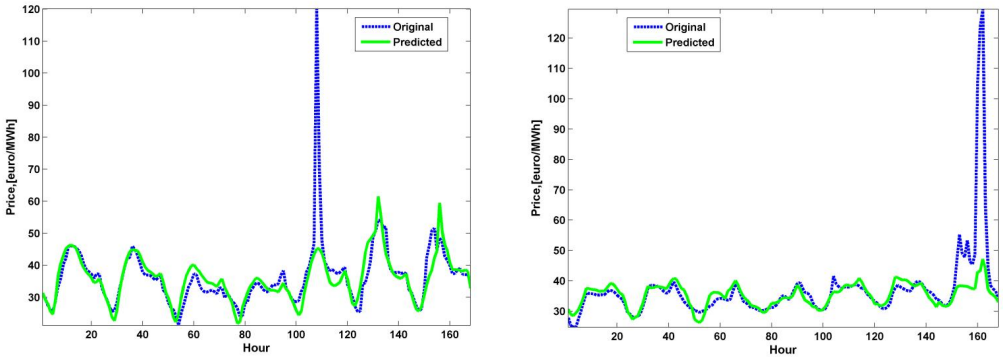

Figure 5.16. Original and predicted prices of week 32 (left) and week 49 (right) of the Finnish day-ahead energy market of the year 2009.

In the light of the findings obtained in Chapters 4–5, an approach separately predicting normal price behavior and price spikes becomes more preferable because of its main ability to use different forecasting engines (for normal prices and price spikes). Such a strategy provides an opportunity to train forecasting models more effectively while the nonseparate methodology should learn the behavior of both normal prices and price spikes. A forecasting methodology for separate treatment of hourly normal prices and price spikes in the day-ahead energy market is extended in the further study.

6 Iterative day-ahead price prediction with separate normal range price and price spike forecasting frameworks

This chapter introduces the day-ahead electricity price forecasting methodology based on an iterative strategy implemented as a combination of two modules separately applied to normal price and price spike prediction. The methodology is intended to capture all essential features of electricity price series, and it produces forecasts of not only normal range prices of high accuracy but of also price spikes. The methodology is built on the findings made within the study and implemented as a combination of different forecasting techniques.

6.1 Description of the forecasting methodology

Similarly to the hybrid model presented in Section 4.2, the new proposed methodology consists of two modules to separately predict normal range prices and price spikes. The normal price module is a mixture of WT, linear SARIMA and nonlinear NN. In the price spike module, the probability of a price spike occurrence is produced by a compound classifier in which three single classification techniques are used jointly to make a decision. Combined with the spike value prediction technique (KNN model), the output from the price spike module aims at providing a comprehensive price spike forecast. The best inputs and optimal parameter settings for forecasting engines of both modules are chosen by the proposed relevance-redundancy feature selection algorithm and the search procedure. The overall electricity price forecast is formed as combined normal price and spike forecasts.

6.2 Electricity price spike extraction

Given n = 3 and w = 6 months (4380 hours), the thus defined price spikes are extracted from the original price series of the Finnish energy market of Nord Pool over the period 1 Jan 2009 – 31 Dec 2010 (see Figure 6.1).

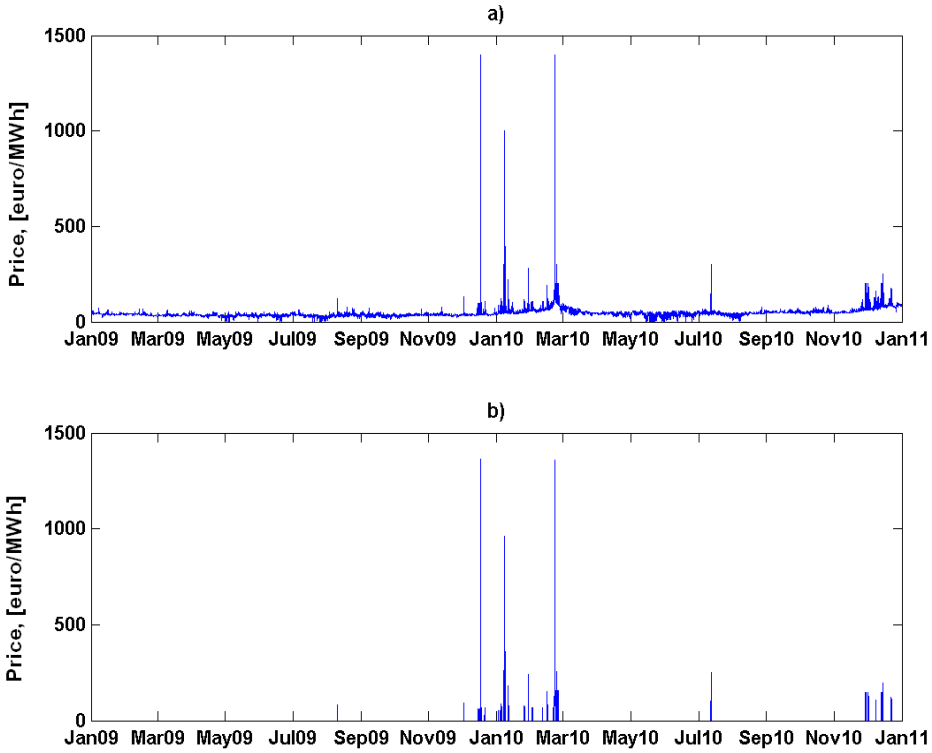

Figure 6.1. a) Original Finnish day-ahead prices over the period 1 Jan 2009–31 Dec 2010; b) extracted price spikes.

Table 6.1 shows the basic distribution parameters for normal prices and spikes in the Finnish energy market of the years 2009–2010.

Table 6.1. Basic statistics for normal prices and price spikes over the period 2009-2010

	Number of observations	Mean	Std	Skewness	Kurtosis
Normal prices	17324	44.62	15.77	9.65	1.88
Spikes	196	240.98	256.68	13.61	3.30

6.3 Compound classifier

As previously, the problem of the price spike occurrence prediction is stated as a classification problem that can be solved by a pattern recognition framework. The results of the experimental assessment of the different classification designs are supposed to be the basis for choosing one of the classifiers as a final solution to the problem. It had been observed that even if one of the designs did yield the best performance, the sets of patterns misclassified by the different classifiers would not necessarily overlap. This suggests that different classifier designs potentially offer complementary information about the patterns to be classified, which could be harnessed to improve the performance of the selected classifier. The idea behind the use of the compound classifier presented in this study is to avoid reliance on a single classifier. Instead, a set of proposed classifiers is used for decision making by combining their individual opinions to derive a consensus decision. Various classifier combination schemes have been devised, and it has been experimentally demonstrated that some of them consistently outperform a single best classifier (Kittler et al., 1998).

To enhance the accuracy and confidence of the price spike occurrence prediction, the compound classifier composed of three different single classifiers is used within the proposed forecasting methodology. The majority vote rule is applied to get an overall output (spike/non-spike) from the compound classifier (Kittler et al., 1998). Here, the votes received from the individual classifiers are simply counted. The class that receives the largest number of votes is then selected as the consensus (majority) decision.

The three individual classifiers used together in the compound classifier are a probabilistic neural network (PNN), RF, and RVM. These methods are chosen because they provide probabilistic output (probability of class membership e.g. probability of spike occurrence). The methods have been previously applied to several other applications with promising results (Amjady and Keynia, 2010; Meng et al., 2009; Huang et al., 2012). Description of the PNN is given in Appendix I (see Section I.1). The output from each classifier is modified by introducing a predetermined probability threshold value V_0 as described in Section 4.2.3.

6.4 Construction of the candidate input set

6.4.1 Price spike forecasting: probability of spike occurrence

The set of candidate inputs selected for the compound classifier is given below:

Historical prices of both time and wavelet domains

Historical prices of original range, approximation price wavelet component ($A3_{price}$), and detail price wavelet component of the highest frequency ($D1_{price}$), all lagged up to 200 hours before a forecast hour, are among the candidate inputs for the compound classifier. Here, a high relation of a spiky price series with the original price series and the above-mentioned price wavelet components is assumed (Amjady and Keynia, 2010).

Electricity demand and supply

Besides the historical price data, electricity demand and supply are among the set of candidate inputs for the price spike forecasting since the relations of these variables are known to drive the movement in the price spikes to a large extent (Zhao et al., 2007b). Therefore, total electricity generation (i.e. internal supply) and electricity demand in Finland, both lagged up to 200 hours before a forecast hour, are selected as candidate inputs for the compound classifier. As mentioned above, reliable forecasts of the demand and supply are highly required. To decrease the overall computational costs within the proposed methodology, it was opted not to simultaneously predict electricity demand/supply and prices as proposed in Section 5.3. Instead, the WT+SARIMA approach, proposed as one of the competitive approaches in Section 5.3 for electricity demand forecasting, is implemented here as a side forecasting model to separately predict demand and supply.

Temporal effect

Hourly (hour_index), daily (day_index), and seasonally (season_index) indices are considered as parameters to indicate the temporal effect when price spikes are predicted.

The SARIMA model is used as a model producing an initial forecast for the compound classifier and provides preliminary day-ahead forecasts for all price wavelet subseries.

Finally, the set of candidate inputs for both the compound classifier and the KNN model of the proposed method includes both historical and forecasted features of both wavelet and time domains. For instance, the 1008 candidate inputs to predict price spike occurrence at hour h are $\{price_{SARIMA,h}, price_{h-1}\ldots,price_{h-200}, demand_{h},\ldots,demand_{h-200}, supply_{h},.., supply_{h-200}, A3_{SARIMA\_price,h}, A3_{price,h-1},\ldots,A3_{price,h-200}, D1_{SARIMA\_price,h}, D1_{price,h-1},\ldots, D1_{price,h-200}, hour\_index_{h}, day\_index_{h}, season\_index_{h}\}$.

It should be noted that within the methodology, the effect of weather variation is incorporated in the electricity demand unlike in the hybrid model presented in Section 4.2 where the temperature data were directly passed to the spike forecasting module. Moreover, the Elspot capacity-flow difference variable is not considered among the candidate inputs to relate price spikes as additional experiments have indicated inefficiency of this variable within the short-term forecasting (i.e. low relation to spike occurrence).

6.4.2 Price spike forecasting: spike magnitude

If the forecast sample is classified by the compound classifier as a spike, the price spike module is activated. The target set to train the KNN model is formed by the price spike samples extracted from the original training price series. The KNN model uses the set of candidate inputs similar to the one utilized for the compound classifier.

6.4.3 Normal range price forecasting

If the forecast sample is classified as a non-spike, the normal price module is activated. All electricity price spikes are extracted from the original training price series and replaced by the corresponding mean price value to form new normal price series. The set of candidate inputs for the model to predict normal prices (i.e. NN), is similar to the one used within the model to predict prices presented in Section 5.3. The SARIMA model produces preliminary day-ahead forecasts for all subseries of the normal price series. The 602 candidate inputs to predict an approximation normal price wavelet component at hour *h* ($A3_{price,h}$) are {$A3_{SARIMA\_price}$, $A3_{price,h-1}$,…, $A3_{price,h-200}$, $A3_{demand,h}$,…, $A3_{demand,h-200}$, $price_{h-1}$,…,$price_{h-200}$}.

6.5 Forecasting strategy

The proposed forecast strategy can be summarized by the following step-by-step algorithm shown also in Figure 6.2:

1) An electricity price time series is decomposed by the wavelet transform into one approximation subseries ($A3_{price}$) and three detailed subseries ($D3_{price}$, $D2_{price}$, $D1_{price}$).
2) WT+SARIMA's are built to produce an initial forecast to predict the future values of the price wavelet subseries 24 hours ahead.
3) The compound classifier is activated.
   3.1. The set of candidate inputs for each classification approach of the compound classifier is constructed.
   3.2. Values of the thresholds (V_0, V_1, V_2) plus corresponding value of a classifier parameter are fine-tuned for each single classifier on the validation data set by the proposed search procedure. Spread values for the Gaussian radial basis

function (RBF) (σ_{RVM}, σ_{PNN}) and the number of trees (N_{tree}) are the specific tuned parameters for the RVM, PNN, and RF, respectively.

3.3. With the selected values, each classification approach of the compound classifier is trained and predicts the price spike occurrence possibility 24 hours ahead.

3.4. A final output from the compound classifier is formed as an overall output from all three single classifiers in a majority voting scheme.

4) For all test samples forecasted by the compound classifier as nonspikes, the normal price prediction module is activated.

4.1. All spike samples are extracted from the original training price series. The new adjusted normal price series is decomposed into four wavelet components.

4.2. WT+SARIMA's are built to forecast the future values of the normal price wavelet subseries 24 hours ahead.

4.3. The set of candidate inputs to predict each normal price wavelet subseries by the NN model is constructed.

4.4. The threshold values (V_1, V_2) and N_h of NNs to predict each normal price wavelet subseries are fine-tuned on the validation data set by the proposed search procedure.

4.5. With the selected values, NNs are trained and predict the normal price wavelet subseries 24 hours ahead.

5) For all test samples forecasted as spikes, the price spike module is activated.

5.1. All spike samples extracted from the original training price series are formed into price spike series used as targets to train the KNN model.

5.2. The set of candidate inputs to predict spike value by the KNN model is constructed.

5.3. The threshold values (V_1, V_2) and the number of neighbor samples (K) for the KNN approach are fine-tuned on the validation data set by the search procedure.

5.4. With the selected values, the KNN model is trained and predicts the price spike value of the test sample.

6) The overall electricity price forecast is formed as a joint output from the normal price and price spike modules.

7) The overall price forecast replaces the predictions produced by the initial forecasting model for the current forecast day (step 2), since it is expected that electricity prices predicted by the separate forecasting frameworks have more accuracy and thus have more relevance with actual values of price. After replacement, the forecasting cycle is repeated as shown in Figure 6.2 until no difference in the overall electricity price forecast output of two successive iteration steps is observed.

6.6 Training and validation phases

The training periods for the forecasting models of the normal price and price spike modules are different. As previously proposed, a 50-day period preceding the forecast

day to train NNs of the normal price module is considered. There are only a few price spike samples in the whole data set (see Table 6.1). Unlike normal price prediction, in order to get a sufficient number of spike samples to train the models of the price spike module, a longer price series period is required. Hence, 365 days preceding the forecast day are considered for the price spike prediction (the compound classifier and the KNN model).

Since the forecasting models of the normal price and price spike modules have the inputs preliminarily predicted by other models (i.e. WT+SARIMA), their training periods are extended to comprise two consecutive training periods: the moving training period for the preliminary model and the training period of the main model. Then, to predict normal prices or price spikes, a day denoted by D is considered in the corresponding second training period.

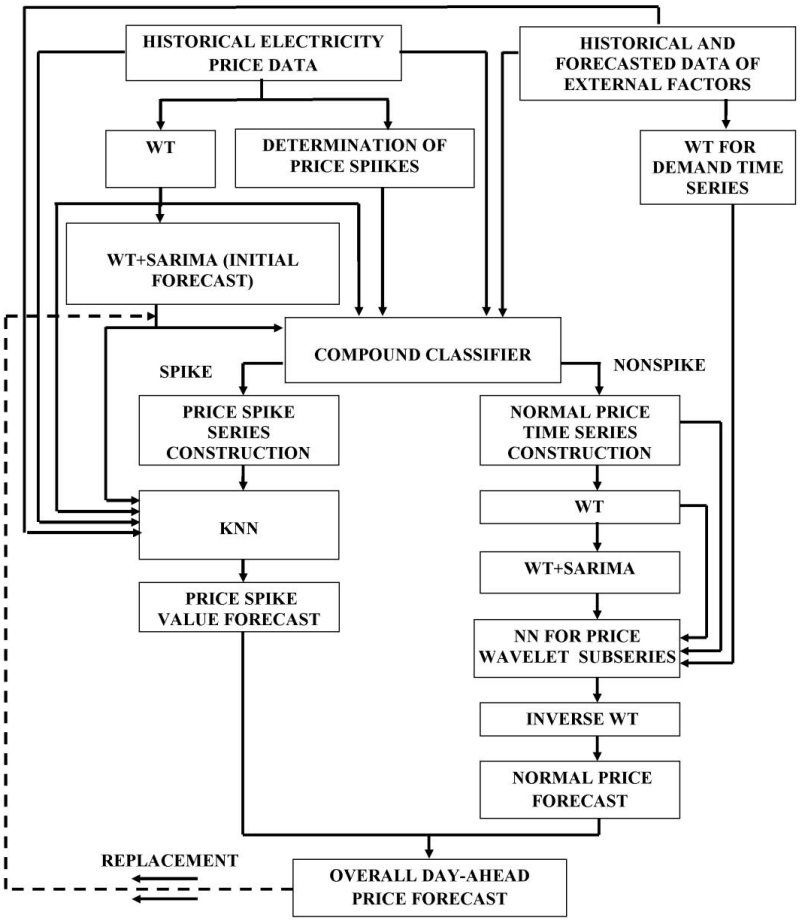

Figure 6.2. Flowchart of the proposed forecasting methodology.

The values of prices for day D are assumed unknown. The preliminary WT+SARIMA models are trained by the historical data of the 50 days preceding hour 1 of day D and predict the price wavelet subseries of day D. To improve the performance of the WT+SARIMA forecast process for each day of the second training period (D = 1,..,50 for NNs or D = 1,…,365 for the price spike module), the WT+SARIMA models are trained by the immediately previous 50-day period. This process is repeated until forecasts from the WT+SARIMA models are obtained for all days of the corresponding second training period (see Figure 6.3).

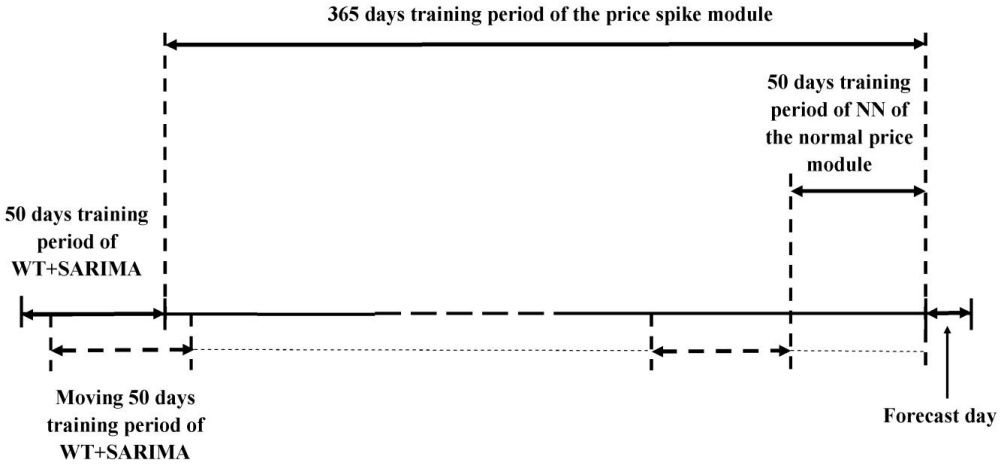

Figure 6.3. Historical data required for the training of the normal price and price spike modules.

The 24 hours price data before the forecast day are removed from the training set of the NNs of the normal price module and used as the validation set. Then, the NNs are trained by the remaining training samples. Adjusted parameters are fine-tuned on the validation set.

For the price spike module, all tuned parameters of the respective approaches are fine-tuned by a 10-fold cross-validation technique applied to the whole training data set (Arlot and Celisse, 2010).

6.7 Numerical results

For examination of the proposed method, the actual hourly data of the Finnish day-ahead energy market are considered. The electricity price, demand, and supply historical data during November 2008–December 2009 are used to establish the training data sample set. The data over the period 1 Jan 2010– 31 Dec 2010 are used as the test set.

The results obtained of the two-step MI based feature selection algorithm for the compound classifier, the KNN model, and the NN model to predict prices in the Finnish

day-ahead energy market for each hour of a single day, 5 Jan 2010, are presented in Tables 6.2–6.3.

Since electricity price spikes have a very volatile stochastic nature with respect to the normal price time series, the regular and periodic behavior of the price spikes are not so obvious. As can be seen in Table 6.2, no features related to the periodic behavior are obtained by the feature selection algorithm.

Table 6.2. Inputs selected by the two-step feature selection analysis for the three classification approaches of the compound classifier and the KNN model.

Engine	V_0	V_1	V_2	Parameter	Selected candidates
RVM	0.43	0.46	0.64	$\sigma_{RVM} = 0.13$	$A3_{SARIMA\_price,h}$, $A3_{price,h-1}$, $A3_{price,h-2}$, $A3_{price,h-4}$, $A3_{price,h-5}$, $A3_{price,h-6}$, $A3_{price,h-7}$, $D1_{SARIMA\_price,h}$, $D1_{price,h-1}$, $D1_{price,h-2}$, $D1_{price,h-3}$,$price_{h-1}$, $price_{h-2}$, $price_{h-3}$, $price_{h-4}$, $demand_h$, $demand_{h-2}$, $demand_{h-46}$, $demand_{h-72}$,$supply_h$, $hour\_index_h$, day_index_h,
PNN	0.47	0.50	0.78	$\sigma_{PNN} = 0.03$	$A3_{SARIMA\_price,h}$, $A3_{price,h-1}$, $A3_{price,h-2}$, $A3_{price,h-3}$, $A3_{price,h-4}$, $A3_{price,h-5}$, $A3_{price,h-6}$, $A3_{price,h-22}$, $D1_{price\_arima,h}$, $D1_{price,h-1}$, $D1_{price,h-2}$, $D1_{price,h-3}$, $D1_{price,h-4}$, $D1_{price,h-5}$, $price_{h-2}$, $price_{h-3}$, $price_{h-4}$, $demand_{h-2}$, $demand_{h-21}$, $demand_{h-22}$, $supply_{h-2}$, $hour\_index_h$, day_index_h, $season\_index_h$
RF	0.42	0.48	0.61	$N_{tree} = 100$	$A3_{SARIMA\_price,h}$, $A3_{price,h-1}$, $A3_{price,h-2}$, $A3_{price,h-4}$, $A3_{price,h-5}$, $A3_{price,h-6}$, $A3_{price,h-7}$, $D1_{price,h-1}$, $D1_{price,h-2}$, $D1_{price,h-3}$, $price_{h-1}$, $price_{h-2}$, $price_{h-3}$, $price_{h-4}$, $price_{h-5}$, $demand_h$, $demand_{h-4}$, $demand_{h-19}$, $demand_{h-69}$, $demand_{h-73}$, $hour\_index_h$, day_index_h
KNN		0.45	0.56	K = 3	$A3_{SARIMA\_price,h}$, $A3_{price,h-2}$, $A3_{price,h-9}$, $A3_{price,h-15}$, $A3_{price,\ h-21}$, $D1_{price,\ h-2}$, $D1_{price,h-5}$, $D1_{price,\ h-7}$, $D1_{price,\ h-8}$, $D1_{price,\ h-16}$, $price_{h-1}$, $price_{h-3}$, $price_{h-7}$, $price_{h-52}$, $demand_{h-190}$, $hour\_index_h$, day_index_h, $season\_index_h$

The variables of the short-run trend ($A3_{price,h-1}$, $D3_{price,h-2}$), daily periodicity ($A3_{price,h-25}$, $D3_{price,h-24}$), and weekly periodicity ($A3_{price,h-169}$, $A3_{demand,h-169}$) are among the selected input features to forecast normal price wavelet components (see Table 6.3). The dependency of the normal price wavelet components on the exogenous variables decreases from $A3_{price}$ to $D1_{price}$.

The overall accuracy of the proposed method is compared with some of the most popular price forecast techniques applied to case studies of energy markets of other

countries: SARIMA (Contreras et al., 2003; Nogales et al., 2002; Taylor et al., 2006); WT+SARIMA (Conejo et al., 2005b; Tan et al., 2010); NN (Zhang and Qi, 2005; Taylor et al., 2006), and WT+NN (Safie-khan et al., 2011). Additionally, WT+SARIMA+NN, which has not been found in the literature is among competitive techniques.

Table 6.3. Inputs selected by the two-step feature selection analysis for the normal price wavelet components.

Variable	N_h	V_1	V_2	Selected candidates
$A3_{price,h}$	4	0.52	0.71	$A3_{SARIMA\_price,h}$, $A3_{price,h-1}$, $A3_{price,h-3}$, $A3_{price,h-4}$, $A3_{price,h-16}$, $A3_{price,h-21}$, $A3_{price,h-25}$, $A3_{price,h-72}$, $A3_{price,h-97}$, $A3_{price,h-121}$,$A3_{price,h-144}$, $A3_{price,h-169}$, $A3_{demand,h-8}$, $A3_{demand,h-10}$, $A3_{demand,h-11}$, $A3_{demand,h-42}$, $A3_{demand,h-91}$, $A3_{demand,h-98}$, $A3_{demand,h-141}$, $A3_{demand,h-169}$, $price_{h-72}$, $price_{h-95}$, $price_{h-97}$, $price_{h-120}$
$D3_{price,h}$	7	0.47	0.81	$D3_{SARIMA\_price,h}$, $D3_{price,h-1}$, $D3_{price,h-2}$, $D3_{price,h-11}$, $D3_{price,h-24}$, $D3_{price,h-48}$, $D3_{price,h-60}$, $D3_{price,h-96}$, $D3_{demand,h-12}$, $D3_{demand,h-47}$, $D3_{demand,h-71}$, $D3_{demand,h-143}$
$D2_{price,h}$	4	0.41	0.74	$D2_{SARIMA\_price,h}$, $D2_{price,h-1}$, $D2_{price,h-7}$, $D2_{price,h-8}$, $D2_{price,h-24}$
$D1_{price,h}$	6	0.15	0.85	$D1_{SARIMA\_price,h}$, $D1_{price,h-6}$, $D1_{price,h-24}$, $D1_{price,h-30}$, $D1_{price,h-48}$, $D1_{price,h-72}$, $D1_{price,h-94}$, $D1_{price,h-120}$, $D1_{price,h-157}$

To demonstrate the efficiency of the proposed methodology, the results obtained for the Finnish day-ahead energy market in the year 2010 are shown in Table 6.4 with the corresponding results obtained from five other competing prediction techniques.

Table 6.4. AMAPE in percent (%) obtained by different prediction techniques for price forecasts in the Finnish energy market of the year 2010.

	SARIMA	WT + SARIMA	NN	WT+NN	WT+ SARIMA+NN	Proposed method
Normal price	10.53	7.53	8.17	8.01	7.18	5.89
Price spikes	55.76	40.51	46.33	44.22	37.72	32.91
Overall	14.93	10.03	12.52	11.98	9.66	8.08

For a fair comparison, NN, WT+NN, and WT+SARIMA+NN have historical and forecasted demand data among the candidate inputs. A feature selection analysis based on the proposed relevance-redundancy filtration is made for all the examined models. The adjustable parameters of the competing models are fine-tuned on the basis of the proposed search procedure. It should be noted that among the alternative examined

models, only the WT+SARIMA+NN model has preliminarily predicted prices in its set of candidate inputs; that is, the NN part of the WT+SARIMA+NN model uses predictions from SARIMA as the candidate input.

As seen from Table 6.4, the AMAPE values corresponding to the proposed strategy are lower than the values obtained from other examined methods. The accuracy improvement of the proposed method with respect to SARIMA, WT+SARIMA, NN, WT+NN, and WT+SARIMA+NN in terms of AMAPE is 45.88% (1-8.08/14.93), 19.44% (1-8.08/10.03), 35.46%(1-8.08/12.52), 32.55% (1-8.08/11.98), and 16.36%(1-8.08/9.66), respectively. It can also be seen that the use of WT results in an improvement in the model accuracy. This improvement in SARIMA in comparison with WT+SARIMA in terms of AMAPE is 32.82% (1-10.03/14.93). For NN in comparison with WT+NN, this value is 4.31% (1-11.98/12.52). The results also confirm the efficiency of the hybrid methodology with linear and nonlinear modeling capabilities (WT+NN versus WT+SARIMA+NN) where the improvement is 19.37% (1-9.66/11.98).

It is expected that the implementation of the proposed iteration strategy increases the accuracy of the overall price prediction. Detailed results of the proposed iteration strategy for the four test weeks of the Finnish day-ahead energy market of the year 2010 are shown in Table 6.5. These test weeks are related to dates 1–7 Jan 2010, 8–14 Jan 2010, 29 Jan–4 Feb 2010, and 5–11 Feb 2010, respectively, and indicate periods of high volatility in the price series. Iteration 0 in Table 6.5 represents the results obtained from the initial forecasting model (i.e., the WT+SARIMA model).

Table 6.5. Accuracy of the proposed iteration procedure in terms of AMAPE (%) for the four test weeks in the Finnish day-ahead energy market of the year 2010.

Iteration number	Week1	Week2	Week5	Week7
0	17.46	37.27	13.49	10.87
1	12.56	26.24	7.96	6.93
2	9.50	25.16	7.24	6.81
3	9.41			6.59

As seen from Table 6.5, the iteration procedure converges in at most of the three cycles, and the prediction error for the four test weeks at the end of the iterative forecast process with respect to Iteration 1 is improved by 13% on average.

In addition, the performance of the proposed compound classifier is compared with each single classifier of the compound classifier and other techniques recently used for price spike occurrence prediction: Naïve Bayesian (Zhao et al., 2007a), SVM (Zhao et al., 2007a), PNN (Amjady and Keynia, 2010), RVM (Meng et al., 2009), and RF (Huang et al., 2012). The total number of the price spike samples extracted from the test period is 182. N_{spikes}, N_{corr}, and N_{as_spikes} for the Finnish day-ahead energy market of the year 2010

are presented in the second, third, and fourth columns of Table 6.6, respectively. Spike prediction accuracy and confidence are given in the fifth and sixth columns of Table 6.6. For a fair comparison, the candidate input sets of all alternative classifiers are similar to the set of candidate inputs of the compound classifier. Optimal settings are selected and the candidate input set is refined for each examined classifier on the basis of the proposed search procedure. All preliminarily predicted price variables that are among the input sets of each competing classifier are predicted by the WT+SARIMA model.

To justify the proposed iteration strategy particularly for the price spike occurrence prediction, N_{corr} and N_{as_spikes}, the classifier accuracy and confidence measures obtained from the compound classifier on the final iteration step of the proposed methodology are shown in the seventh, eighth, ninth, and tenth columns of Table 6.6, respectively.

Table 6.6. N_{corr} and N_{as_spikes}, classifier accuracy and confidence for price spike classification.

		WT+SARIMA as the initial forecasting model				Final iteration step of the proposed methodology			
Engine	N_{spikes}	N_{corr}	N_{as_spikes}	Acc., [%]	Conf., [%]	N_{corr}	N_{as_spikes}	Acc., [%]	Conf., [%]
Bayes		124	247	68.13	50.20				
SVM		120	174	65.93	68.79				
PNN		112	155	61.54	72.26	147	161	80.77	91.30
RVM	182	119	168	65.38	70.83	163	190	89.56	85.79
RF		122	166	67.03	71.08	152	179	83.51	84.92
Comp-d		122	152	67.03	77.63	162	174	89.01	93.10

The results given in Table 6.6 indicate that the use of the iteration strategy results in a notable accuracy improvement of the price spike occurrence prediction. Table 6.6 also shows that the compound classifier performs better than all single classifiers. Only the RVM has a slightly better spike prediction accuracy than the compound classifier, while the compound classifier has a considerably better spike prediction confidence than the RVM.

Further, the set of test price spike samples are divided according to their original price value intervals. Large price spikes with price values varying between 300 and 1500 euro/MWh constitute around 15% of all the spike samples. Because of their values and stochastic character, such spikes are extremely important for market participants. The results obtained from each classifier and the compound classifier itself on the final iteration step of the proposed methodology for the Finnish day-ahead energy market of the year 2010 are shown in Table 6.7. All the classifiers presented in Table 6.7 are able to correctly discriminate all the large spike samples over the test period. The price

prediction accuracy of the examined classifiers varies in prediction of price spike samples between 85 and 300 euro/MWh.

Table 6.7. Results obtained from the compound classifier for different price spike intervals in the Finnish day-ahead energy market of the year 2010.

Original price interval, [euro/MWh]	N_{spikes}	N_{corr}			
		PNN	RVM	RF	Compound
85 -150	66	50	55	53	55
150 -300	87	68	79	70	78
300 -500	18	18	18	18	18
500 -1000	1	1	1	1	1
1000 -1500	10	10	10	10	10
TOTAL	182	147	163	152	162

For a more detailed representation of the performance of the proposed overall price forecast strategy and separately for the price spike occurrence on the whole test year, their results for all the weeks of the 2010 are shown in Table 6.8. There are six measures given for all test weeks of the Finnish day-ahead energy market of the year 2010: the overall AMAPE, N_{spikes}, N_{cor}, N_{as_spikes}, the classifier accuracy and confidence of the compound classifier.

Table 6.8. Results obtained from the proposed forecasting methodology for each week of the year 2010.

Week	1	2	3	4
AMAPE, [%]	9.41	25.16	6.31	5.75
$N_{spikes}/N_{corr}/N_{as\_spikes}$	23/21/22	22/22/22	0/0/0	9/8/8
Acc./Conf, [%]	91.30/95.45	100/100	-	88.89/100
Week	5	6	7	8
AMAPE, [%]	7.24	4.51	6.59	30.75
$N_{spikes}/N_{corr}/N_{as\_spikes}$	7/7/7	1/0/0	5/5/5	44/39/39
Acc./Conf, [%]	100/100	0/-	100/100	77.27/97.14
Week	9	10	11	12
AMAPE, [%]	6.49	6.11	4.76	3.55
$N_{spikes}/N_{corr}/N_{as\_spikes}$	2/2/2	0/0/0	0/0/1	0/0/1
Acc./Conf, [%]	100/100	-	-/0	-/0
Week	13	14	15	16
AMAPE, [%]	2.84	2.99	3.31	5.07
$N_{spikes}/N_{corr}/N_{as\_spikes}$	0/0/0	0/0/0	0/0/0	0/0/0
Acc./Conf, [%]	-	-	-	-
Week	17	18	19	20
AMAPE, [%]	6.11	6.88	7.43	15.51

$N_{spikes}/N_{corr}/N_{as\_spikes}$	0/0/0	0/0/0	0/0/0	0/0/0
Acc./Conf, [%]	-	-	-	-
Week	21	22	23	24
AMAPE, [%]	6.35	8.03	7.23	6.46
$N_{spikes}/N_{corr}/N_{as\_spikes}$	0/0/0	0/0/0	0/0/0	0/0/0
Acc./Conf, [%]	-	-	-	-
Week	25	26	27	28
AMAPE, [%]	6.15	7.23	4.26	12.82
$N_{spikes}/N_{corr}/N_{as\_spikes}$	0/0/0	0/0/0	0/0/0	9/7/7
Acc./Conf, [%]	-	-	-	77.78/100
Week	29	30	31	32
AMAPE, [%]	4.56	5.38	7.58	6.34
$N_{spikes}/N_{corr}/N_{as\_spikes}$	0/0/2	0/0/0	0/0/0	0/0/0
Acc./Conf, [%]	-/0	-	-	-
Week	33	34	35	36
AMAPE, [%]	3.06	3.14	4.99	2.19
$N_{spikes}/N_{corr}/N_{as\_spikes}$	0/0/0	0/0/0	0/0/0	0/0/0
Acc./Conf, [%]	-	-	-	-
Week	37	38	39	40
AMAPE, [%]	3.64	2.65	3.64	2.43
$N_{spikes}/N_{corr}/N_{as\_spikes}$	0/0/0	0/0/0	0/0/0	0/0/0
Acc./Conf, [%]	-	-	-	-
Week	41	42	43	44
AMAPE, [%]	3.83	4.09	5.91	3.14
$N_{spikes}/N_{corr}/N_{as\_spikes}$	0/0/0	0/0/0	0/0/1	0/0/0
Acc./Conf, [%]	-	-	-	-
Week	45	46	47	48
AMAPE, [%]	2.26	2.83	4.03	17.12
$N_{spikes}/N_{corr}/N_{as\_spikes}$	0/0/0	0/0/0	0/0/0	12/7/8
Acc./Conf, [%]	-	-	-	58.33/87.50
Week	49	50	51	52
AMAPE, [%]	8.37	11.02	5.70	4.22
$N_{spikes}/N_{corr}/N_{as\_spikes}$	3/3/3	35/33/34	9/8/11	0/0/1
Acc./Conf, [%]	100/100	82.86/96.67	88.89/72.73	-/0

As can be seen from Table 6.8, the price forecasts of the weeks related to a winter season (December–February), that is, the weeks 1–8 and 48–52 of the year 2010, have a relatively higher prediction error compared with the price forecasts related to other seasons. It is unsurprising that the performance of the proposed forecasting methodology is worse during the winter season because of the extreme price volatility reflected in price spikes, which is caused by a number of complex factors and which takes place during periods of market stress. These stressed market situations are generally associated with extreme meteorological events and unusually high demand. However, in the light of the fact that the occurrence of price spikes typical in the winter

period is predicted by the proposed methodology with high confidence, the achieved overall forecast accuracy level is fairly good and provides market participants with an ability to analyze price spike probabilities and thus manage their risks.

In order to graphically illustrate performance of the proposed forecasting methodology, the prediction performance and actual signals for the four test weeks of the year 2010, corresponding to the four seasons are shown in Figure 6.4.

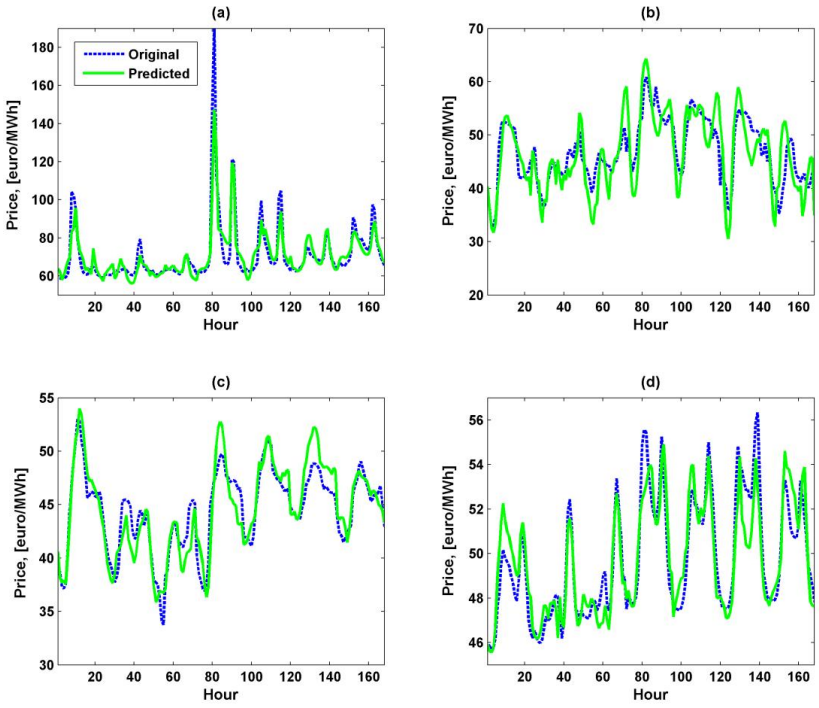

Figure 6.4. Original and predicted prices for the four test weeks of the Finnish day-ahead energy market of the year 2010: (a) Winter week; (b) Spring week; (c) Summer week; (d) Fall week.

The four weeks, a winter week (12 Feb to 18 Feb), a spring week (14 May to 20 May), a summer week (13 Aug to 19 Aug), a fall week (12 Nov to 18 Nov), were considered representative for a study spanning one whole year. All the forecast price curves acceptably follow the actual ones. The proposed methodology based on a hybrid iterative strategy is able to capture the essential features of the given price time series: nonconstant mean, cyclicality, exhibiting daily and weekly patterns, major volatility, and significant outliers.

Additionally, to emphasize the ability of the proposed methodology to capture spikes in the price series, Figure 6.5 presents forecasting results from the proposed methodology for the four selected spiky weeks (weeks 1, 2, 5, and 28 in Table 6.8). The forecasting performance of the competing approaches for these weeks is shown in Appendix I (see Section I.2).

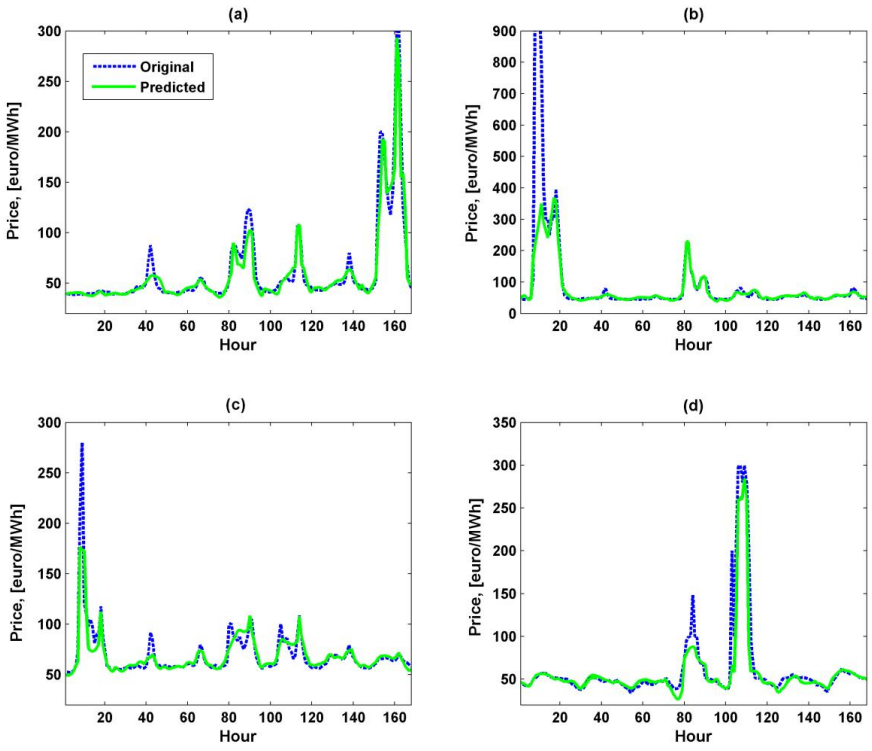

Figure 6.5. Original and predicted prices for the four weeks with prominent spikes of the Finnish day-ahead energy market of the year 2010: (a) Week 1; (b) Week 2; (c) Week 5; (d) Week 28.

It should be noted that many other exogenous variables can be considered in candidate input sets for feature selection, such as fuel costs and some meteorological information, but this is a topic for future research. Moreover, there is a clear need for a more accurate method for price spike value prediction.

The total running time to set up the proposed separate forecasting strategy including its normal price module, price spike module, and iterative prediction process for the first forecast day is about 42 hours since price predictions produced by the initial forecasting model are required over the period up to 365 days. Similarly, in the previously proposed

forecasting strategy, simultaneously predicting price and demand (see Chapter 5), the running time of the training and prediction procedures for the next forecast days after the first one is significantly lower (about 50 min) and considered suitable for day-ahead energy market operation. All the competitive nonseparate forecasting approaches examined for price prediction have lower computation costs than the proposed separate forecasting strategy but are outperformed by the proposed strategy in terms of forecasting accuracy. The PNN and RVM classifiers of the compound classifier have relatively lower computational costs than the alternative back-propagation NN and SVM, respectively. The training process of the PNN is carried out through one run of each training sample unlike the back-propagation algorithm. The RVM is faster than the SVM in decision speed, as the RVM has a much sparser structure (the number of relevant vectors versus the number of support vectors). The computation times to set up the proposed and competitive forecasting strategies are measured on a hardware including Intel Core i5 2.40 GHz processor and 3.24 GB RAM. All computer codes are provided by the MATLAB and R software packages.

6.8 Summary

The proposed methodology is able to capture high volatility of prices to distinctly distinguish normal prices and price spikes when the overall price path is forecasted. By providing such ability, the proposed methodology significantly outperforms all other competing approaches examined in the study. Thus, the proposed methodology can be applied to the entire Nordic market and deregulated markets in other countries to provide extensive and useful information for the participants of the energy market, who have limited and uncertain information for price prediction.

7 Conclusions

7.1 Summary and conclusions

The main purpose of this thesis was to present a model able to predict not only day-ahead electricity prices within the normal range with a high degree of accuracy but also price spikes. The structure of a case market, which is selected to be the day-ahead Nordic energy market (Nord Pool Spot) and, particularly, the Finnish day-ahead energy market, is studied in detail, and then, a set of potential explanatory factors that may influence the price behavior in the Nordic electricity market are stated.

A wide range of market data from the Nord Pool Spot over the period from 1 Jan 1999 to 31 Dec 2012 were investigated and statistically analyzed. The existing seasonal patterns and the remaining stochastic component were extracted with the help of a decomposition technique for further analysis.

Various classical and more elaborate modern approaches were developed to relate the electricity market price behavior in the Finnish day-ahead energy market. A linear multiple moving regression model was examined with different lengths of training periods to predict day-ahead prices. Residuals obtained from the regression model fit were prone to outliers and presented nonconstant mean level and high spikes over the testing period.

Next, the Box-Jenkins models were presented to relate the electricity price behavior by altering the given series to make it stationary. It was shown that the Box-Jenkins models were unable to estimate high volatility and spike clustering presented in the original price series. A difference filter used within the Box-Jenkins model was not able to remove and add deterministic elements of original price series accurately for out-of-sample modeling. Therefore, separate treatment of the deterministic elements through a decomposition model was proposed as more effective technique.

MR financial stochastic models based on Ornstein-Uhlenbeck approach, even when optimally calibrated with the real day-ahead electricity market prices, were not able to capture the statistical characteristics of the real series.

Following the aforementioned conclusions, spike samples were extracted from the corresponding data set to avoid the undesirable effect of those samples on the parameter estimates in the models. Further, it was suggested to implement a separate prediction of adjusted price series (when spikes were extracted) and price spikes with the use of different forecasting engines.

The Box-Jenkins and MR models enhanced with the regime switching approach were presented as common tools to model day-ahead electricity prices. The regime-switching model for the spike process was applied to working and nonworking days of different seasons and the transition probabilities were determined for each case, as the number of

spikes and the length of spike groups could differ for different day types. The impact of the regime-switching was very clear when comparing the results of the models combining the regime-switching with the corresponding results obtained from the models without the regime-switching.

Then, an electricity price forecasting model implemented as a combination of classical and modern forecasting approaches was presented to analyze electricity price time series in two parts; normal behavior and spiky behavior. The Box-Jenkins models were combined with the NN to capture linear and nonlinear relationships between the normal range price series and the selected explanatory variables. The probability of a price spike occurrence and the value of a price spike were produced by two different forecasting engines (GMM and KNN). Selection of explanatory variables and model parameters to predict both normal range prices and price spikes was based on heuristics. It has been shown that when applied in addition to the normal range price forecast, the proposed price spike forecast method could provide practically useful and reasonably accurate forecasts, thereby enhancing the applicability of the price forecasts in the actions of electricity market participants. However, the feature selection process based on heuristics and past experience was found inappropriate for accurate price forecasting since each electricity market has characteristics of its own.

The two-step feature selection algorithm was proposed to refine the set of candidate inputs such that a subset of the most effective inputs was selected for the forecast engine. Besides the simple calculation of the corresponding relevance value between the target and candidate inputs, the problem of the presence of excessive information in the set of candidate inputs was solved by adding the filter removing redundant information from the set of candidate inputs. Several linear and non-linear statistical techniques were used to calculate the relevance value between the given variables within the proposed feature selection algorithm. The performance of different forecasting engines given by different feature selection techniques was examined. It was shown that the feature selection techniques consisting of both relevance and redundancy filters outperformed the feature selection techniques having only a relevance filter for all the considered forecasting engines. Moreover, the models applying feature selection techniques that were able to consider the nonlinearities of the price signal performed significantly better than those models merged with linear feature selection techniques. The heuristic approach that was among the examined feature selection techniques indicated the worst performance for all the considered forecasting engines.

The importance of the correct parameter setting for any forecasting model was addressed in the thesis. Hence, the search procedure implemented to find the optimal setting model parameters including the thresholds for the proposed two-stage feature selection algorithm was proposed.

The study showed that besides past price values, inclusion of fundamental data (e.g. electricity demand) in a forecasting model could significantly improve a forecasting performance. In some past studies, actual values of fundamental data were used to

forecast prices. However, such a strategy was shown to be not suitable for a real-life problem as realized values of fundamental data were not known for a day-ahead market at the moment when the prices were predicted. The thesis addressed the mutual effect of price and demand in the deregulated markets where not only the system demand influences the prices but the system demand may be significantly affected by the electricity prices when consumers are encouraged to use less energy during peak hours. To incorporate the above-mentioned mutual effect in the forecasting model, a methodology was proposed that simultaneously predicted the electricity demand and prices in the day-ahead energy market. The proposed methodology was better adapted to actual conditions in an energy market since the forecast features for price and demand were not assumed known values but were predicted by the model, thus accounting for the interactions of the demand and price forecast processes. A WT approach was applied within the proposed methodology to data pre-processing and considered as an alternative to the previously applied time series decomposition technique. The corresponding frequency components obtained by WT were separately predicted by the combined SARIMA and NN. Such a strategy was supposed to improve the overall forecasting performance and, in particular, spikes in the series without a separate normal and spiky price prediction since there was a high correlation between the price spikes and high-frequency wavelet components of the price signal spectrum. The best input set and model parameter settings were selected by the proposed search procedure and the relevance-redundancy feature selection algorithm. The methodology outcomes showed that the proposed forecasting strategy was able to achieve more accurate predictions than separate frameworks recently used for the case study of electricity price forecasting. However, the proposed methodology still suffered from its inability to capture anomalous price behavior (when prices increased rapidly and unexpectedly) to the full.

In the light of the findings obtained along the study, separate normal price and price spike forecasting was found to be the most preferable approach owing to its main ability to train forecasting models more effectively while the nonseparate methodology should learn the behavior of both normal prices and price spikes.

The novel methodology based on an iterative forecasting strategy, implemented as a combination of two modules separately applied to normal price and price spike prediction, was presented. The normal price module employed the previously applied forecasting technique that was a mixture of WT, linear SARIMA and nonlinear NN. As previously, the price spike module was a combination of the spike probability and the spike value forecasting models. However, the spike probability model was implemented on basis of three single classification techniques used jointly to make a final decision. Selection of the best input set and parameter settings for the models employed within the methodology was implemented by the proposed search procedure and the relevance-redundancy feature selection algorithm. It has been shown that the use of the proposed iteration strategy significantly increases the accuracy of the overall price prediction and, in particular, price spikes. The methodology was examined for price prediction and resulted in the best forecasting performance when it was compared with some of the

most popular and recently used frameworks for price prediction and separately for price spike occurrence prediction.

The proposed methodology was suggested to be applied to the entire Nordic market and deregulated markets in other countries since it was able to capture different distinct features of the given price time series and incorporate the iteration strategy separately to predict normal prices and price spikes.

Prior information on market price fluctuations is a crucial concern for market participants. An energy producer will benefit from having such information, as it allows developing an optimal production schedule subject to the producer's marginal costs. The generated price forecast can be used for the development of an optimal short-term operation portfolio. Development of an optimal short-term operation portfolio of a single market customer, that is, a CHP power plant, the given price forecast was considered in Appendix J. In this case, two price forecasts of different accuracies were examined and the associated economic impacts were studied.

7.2 Contributions

The focus of this thesis is on developing the most accurate day-ahead energy market price forecasting model that is able to predict not only normal prices but also price spikes and that is feasible from the market participants' viewpoint. The following are the main contributions of the thesis:

1. A survey of current linear and nonlinear forecasting methodologies and their applications to price prediction in the day-ahead energy market of Finland is presented.
2. Two-step feature selection algorithm is applied to refine the set of candidate inputs for a forecasting model by extracting nonrelevant and redundant information from the set of candidate inputs.
3. The search procedure to obtain an optimal set of input features and adjustable parameter settings for a forecasting model is proposed.
4. The forecasting methodology separately predicting normal prices and price spikes is claimed as the most accurate approach to relate actual price behavior on a day-ahead energy market.
5. The model simultaneously predicting electricity demand and prices on a day-ahead energy market is implemented. It is better adapted to actual conditions of an energy market as forecast features for price and demand are not assumed known values but are predicted within the model thus accounting for the interactions of the demand and price forecast processes.
6. The novel forecasting methodology based on iterative strategy implemented as a combination of two modules separately applied to normal price and price spike prediction is proposed and implemented. The two modules are built on a hybrid approach and implemented as a mixture of different models. A set of experiments have shown that the proposed methodology is able to capture

essential features of the real price series and significantly outperforms the most popular and recently used frameworks for price prediction in a day-ahead energy market.

7. The application of a price forecast with different levels of accuracy is examined to obtain an optimal short-term operation scheduling of a single market customer.

7.3 Suggestions for future work

Based on the research work presented and discussed in this thesis, further research may be pursued on the following subjects:

- Construction of an interface between different pieces of forecasting software packages used to implement the proposed forecasting methodology. It could make the forecasting model more easy and practical to use by researchers and software users.
- The effect of other variables (besides electricity demand/supply) when integrated into the proposed price forecasting methodology is a topic of future research. These include fuel costs, regulatory constraints etc.
- Development of a more accurate method for price spike value prediction. The possible methods that could be based on NNs or RVM regression approaches can be considered in the future work.
- Study on the application of price forecasts for short-term operation scheduling of actual market participants. Investigate the energy costs sensitivity to price forecast accuracy across different market participants.
- Application of the proposed forecasting methodology to after-spot energy markets (Elbas market).
- Investigation of the effect of market power and supplier bid behavior on the market price formation.

References

Aggarwal, S., Saini, L., and Kumar, A. (2009), "Electricity price forecasting in deregulated markets: A review and evaluation," *International Journal of Electrical Power & Energy Systems*, 31(1), pp. 13–22.

Ali, J., Khan, R., Ahmad, N., and Maqsood, I. (2012), "Random Forests and Decision Trees," *IJCSI International Journal of Computer Science*, 9(5), pp. 272–278.

Amjady, N. (2006), "Day ahead price forecasting of electricity markets by a new fuzzy neural network," *IEEE Trans. Power Syst*., 21(2), pp. 887–896.

Amjady, N. and Hemmati, M. (2009), "Day-ahead price forecasting of electricity markets by a hybrid intelligent system," *European Transactions on Electric Power*, 19(1), pp. 89–102.

Amjady, N. and Keynia, F. (2008), "Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method," *Int. J. Electr. Power Energy Syst.*, 30 (9), pp. 533–546.

Amjady, N. and Keynia, F. (2010), "Electricity market price spike analysis by a hybrid data model and feature selection technique," *Electr. Power Syst. Research*, 80(3), pp. 318–327.

Arlot, S. and Celisse, A. (2010), "A survey of cross-validation procedures for model selection," *Statistics. Surveys*, 4, pp. 40–79.

Arnold, L. (1974), *Stochastic Differential Equations: Theory and Applications*, London: Wiley.

Bajpai, P. and Singh, S. (2004), "Bidding and gaming in electricity market: an overview and key issues," in *International conference on power systems*, Kathmandu, Nepal, pp. 571–576.

Barlow, M. (2002), "A diffusion model for electricity prices," *Math Finance*, 12(4), pp. 287–298.

Barroso, L., Cavalcanti, T., Giesbertz, P. and Purchala, K. (2005), "Classification of electricity market models worldwide," in: *CIGRE Task Force C5.2.1.*, available at http://raceadmv3.nuca.ie.ufrj.br/BuscaRace/Docs/lanbarroso15.pdf

Barz, G. and Johnson, B. (1998), "Modeling the prices of commodities that are costly to store: the case of electricity," *Proceedings of the Chicago Risk Management Conference*, Chicago, IL.

Bastian, J., Zhu, J., Banunarayanan, V. and Mukerji, R. (1999), "Forecasting energy prices in a competitive market," *IEEE Computer Applicat. Power*, 12, pp. 40–45.

Battiti, R. (1994), "Using Mutual Information for Selecting Features in Supervised Neural Net Learning," *IEEE Trans. Neural Networks*, 5(4), pp. 537–550.

Becker, R., Hurn, S. and Pavlov, V. (2007), "Modelling Spikes in Electricity Prices," *The Economic Record*, 82(263), pp. 371–382.

Bhattacharyya, S.C. and Thanh, L.T. (2003), "Short-term electric load forecasting using an artificial neural network: case of Northern Vietnam," *International Journal of Energy Research*, 28(5), pp. 463–472.

Bierbrauer, S.T., Trück, S. and Weron, R. (2004), "Modeling electricity prices with regime-switching models," *Lecture Notes in Computer Science*, Springer, Berlin, pp. 859–867.

Bollerslev, T. (1986), "Generalized autoregressive conditional heteroscedasticity," *Journal of Econometrics*, 31, pp. 307–327.

Box, G.E.P. and Jenkins, G.M. (1970), *Time Series Analysis: Forecasting and Control*, San Fransisco, Holden-Day.

Breiman, L. (1984), *Classification and regression trees*, *Wadsworth* statistics/ probability series, Chapman & Hall.

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

Carrion, M., Philpott, A. B., Conejo, A. J. and Arroyo, J. M. (2007), "A stochastic programming approach to electric energy procurement for large consumers," *IEEE Trans. Power Syst*., 22(2), pp. 744–754.

Cartea, A. and Figueroa, M. (2005), "Pricing in electricity markets: A mean reverting jump diffusion model with seasonality," *Applied Mathematical Finance*, 12(4), pp. 313–335.

Catalão, J.P.S., Mariano, S.J.P.S., Mendes, V.M.F. and Ferreira, L.A.F.M. (2007), "Short-term electricity prices forecasting in a competitive market: A neural network approach," *Elect. Power Syst. Res.*, 77, pp. 1297–1304.

Catalão, J.P.S., Pousinho, H.M.I. and Mendes, V.M.F. (2009), "An artificial neural network approach for short-term wind power forecasting in Portugal," in: *15th International Conference on Intelligent System Applications to Power Systems*, pp.1–5.

Cavallaro, F. (2005), "Electric load analysis using an artificial neural network," *International Journal of Energy Research*, 29, pp. 377–392.

Chandarasupsang, T., Galloway, S., Burt, G., McDonald, J. and Siewierski, T. (2007), "Bidding behaviour and electricity market simulation," *European Transactions on Electrical Power*, 17, pp. 333–346.

Chen, Y., Luh, P.B., Guan, C., Zhao, Y., Michel, L.D., Coolbeth, M.A., Friedland, P.B. and Rourke, S.J. (2010), "Short-term Load Forecasting: Similar Day-based Wavelet Neural Networks," *IEEE Transaction on Power Systems*, 25(1), pp. 322–330.

Conejo, A.J., Contreras, J., Espínola, R. and Plazas, M.A. (2005a), "Forecasting electricity prices for a day-ahead pool-based electricity energy market," *Int. J. Forecasting*, 21(3), pp. 435–462.

Conejo, A.J., Plazas, M.A., Espinola, R. and Molina, A.B. (2005b), "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," *IEEE Trans Power Syst.*, 20, pp. 1035–1042.

Contreras, J., Espínola, R., Nogales, F.J. and Conejo, A.J. (2003), "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, 18(3), pp. 1014–1020.

Cover, T.M. and Thomas, J.A. (1991), "Entropy, relative entropy and mutual information," *Elements of Information Theory*, John Wiley & Sons, pp. 12–49.

Duda, R.O., Hart, P.E. and Stork, D.G. (2001), *Pattern Classification*, New York, Wiley-Interscience.

Engle, R.F. (1987), "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, 50, pp. 987–1007.

Fingrid (2013a), *Balance service, Regulation power market*, Accessed 1 May 2013, available at http://www.fingrid.fi/.

Fingrid (2013b), *Electricity market: cross-border transmission*, Accessed 1 May 2013, available at http://www.fingrid.fi.

Fosso, O.B., Gjelsvik, A., Haugstad, A., Mo, B. and Wangensteen, I. (1999), "Generation scheduling in a deregulated system. The Norwegian case," *IEEE Trans. Power Syst.*, 14 (1), pp. 75–80.

Fu, Y. and Li, Z. (2006), "Different models and properties on LMP calculations," in: *Proc. IEEE Power Eng. Soc. General Meeting*, Montreal, Canada.

Galli, W. (1996), "Exploring the power of wavelet analysis," *IEEE Comput. Appl. Power*, 9, pp. 37–41.

Gao, F., Cao, X. and Papalexopoulous, A. (2000), "Forecasting power market clearing price and quantity using a neural network method," in *Proc. IEEE Power Eng. Soc. Summer* Meet., 4, Seattle,WA, pp. 2183–2188.

Garcia, R.C., Contreras, J., van Akkeren, M. and Garcia, J.B.C. (2005), "A GARCH forecasting model to predict day-ahead electricity prices," *IEEE Trans. Power Syst*., 20(2), pp. 867–874.

Georgilakis, P.S. (2006), "Market clearing price forecasting in deregulated electricity markets using adaptively trained neural networks," *Lecture notes in artificial intelligence LNAI* 3955. Berlin, Heidelberg: Springer-Verlag, pp. 56–66.

Gibson, R. and Schwartz, E.S. (1990), "Stochastic convenience yield and the pricing of oil contingent claims," *J. Finance*, 45 (3), pp. 959–976.

Granger, C.W.J. (2008), "Non-Linear Models: Where Do We Go Next – Time Varying Parameter Models?" *Studies in Nonlinear Dynamics and Econometrics*, 12(1).

Guan, X., Pepyne, D. and Ho, Y.C. (2001), "Gaming and price spikes in electric power markets," *IEEE Trans. Power Syst*., 16, pp. 402–408.

Haykin, S. (1994), *Neural Networks: A Comprehensive Foundation*, Englewood Cliffs, NJ. Prentice-Hall.

He, Y.G. and Bo, L. (2009), "Price forecasting based on PSO train BP neural network," in: *Power and Energy Engineering Conference*, pp. 1–4.

Hippert, H.S., Pedreira, C.E. and Souza, R.C. (2001), "Neural Networks for Short-Term Load Forecasting: A Review and Evaluation," *IEEE Transaction on Power Systems*, 16(1), pp. 44–55.

Hirsch, G. (2009), "Pricing of hourly exerciseable electricity swing options using different price processes," *Journal of Energy Markets*, 2(2), pp. 3–46.

Hsu, W., Liu, C., Chang, F. and Chen, S. (2012), "Cancer classification: Mutual information, target network and strategies of therapy," *Journal of Clinical Bioinformatics*, pp. 2–16.

Huang, D., Zareipour, H., Rosehart, W.D. and Amjady, N. (2012), "Data mining for electricity price classification and the application to demand-side management," IEEE *Trans. on Smart Grid*, 3(2), pp. 808–817.

International Energy Agency (2012), *Statistics and Balances*, Accessed 21 May 2013, available at http://www.iea.org/stats/index.asp.

Jabłońska, M. (2008), "Analysis of outliers in electricity spot market price with example of New England and New Zealand markets," *Master's Thesis Lappeenranta University of Technology*, Finland.

Jabłońska, M., Nampala, H. and Kauranne, T. (2011), "The multiple-mean-reversion jump-diffusion model for Nordic electricity spot prices," *The journal of energy* market, 4(2), pp 3-25.

Jabłońska, M., Viljainen, S., Partanen, J. and Kauranne, T. (2012), "The impact of emissions trading on electricity spot market price behavior," *International Journal of Energy Sector Management*, 6(3), pp. 343–364.

Karandikar, R.G., Deshpande, N.R., Khaparde, S.A. and Kulkarni, S.V. (2009), "Modelling Volatility Clustering in Electricity Price Return Series for Forecasting Value at Risk," *European Transactions on Electrical Power*, 19, pp. 15–38.

Karatzas, I. and Shreve, S.E. (2000), *Brownian Motion and Stochastic Calculus*, Springer, New York.

Keles, D., Genoese, M., Most, D. and Fichtner, W. (2012), "Comparison of extended mean-reversion and time series models for electricity spot price simulation considering negative prices," *Energy Economics*, 34, pp. 1012-1032.

Kira, K. and Rendell, L.A. (1992), "The feature selection problem: traditional methods and new algorithm," in: *Proceedings of AAAI'92*.

Kittler, J., Hatef, M., Duin, R.P. and Matas, J.G. (1998), "On combining classifiers," *IEEE Trans. Patt. Anal. Mach. Int.*, 20 (3), pp. 226–239.

Kohzadi, N., Boyd, M.S., Kermanshahi, B. and Kaastra, I. (1996), "A comparison of artificial neural network and time series models for forecasting commodity prices," *Neurocomputing*,10 (2), pp.169–181.

Kononenko, I. (1994), "Estimating attributes: analysis and extensions of Relief," in: *L. De Raedt and F. Bergadano (eds.): Machine Learning: ECML-94*, Springer, Verlag, pp. 171–182.

Kwak, N. and Choi, C.-H. (2002), "Input Feature Selection for Classification Problems," *IEEE Trans. Neural Networks*, 13(1), pp. 143–159.

Lari-Lavassani, A., Sadeghi, A.A. and Ware, A. (2001), "Mean reverting models for energy option pricing," Accessed 1 April 2013, available at http:// math.ucalgary.ca/ sites/finance.math.ucalgary.ca/files/u22/LavassaniSadeghiWare2001.pdf.

Li, T., Shahidehpour, M. and Li, Z. (2007), "Risk-constrained bidding strategy with stochastic unit commitment," *IEEE Trans. Power Syst.*, 22(1), pp. 449–458.

Liao, G.C. (2007), "A novel particle swarm optimization approach combined with fuzzy neural networks for short-term load forecasting," in: *IEEE Power Engineering Society General Meeting*.

Lora, A.T., Santos Jesus, R., Santos Jose, R., Exposito, A.G. and Ramos, J.L.M. (2002), "A comparison of two techniques for next-day electricity price forecasting," *Lecture notes in computer science*, Berlin, Heidelberg: Springer Verlag, pp. 384–390.

Lu, X., Dong, Z.Y. and Li, X. (2005), "Electricity Market Price Spike Forecast with Data Mining Techniques," *Electric Power Systems Research*, 73(1), pp.19–29.

Lucia, J. and Schwartz, E. (2000), "Electricity prices and power derivatives: Evidence from the Nordic Power Exchange," *UC Los Angeles: Finance*. Retrieved from: http://www.escholarship.org/uc/item/12w8v7jj

Mallat, S. (1989), "A theory for multiresolution signal decomposition-the wavelet representation," *IEEE Trans Pattern Anal. Mach. Intell.*, 11, pp. 674–693.

Mandal, P., Senjyu, T. and Funabashi, T. (2006), "Neural networks approach to forecast several hour ahead electricity prices and loads in deregulated market," *Energy Convers. Management*, 47, pp. 2128–2142.

Mandal P., Senjyu, T., Urasaki, N., Funabashi, T. and Srivastava, A.K. (2007), "A novel approach to forecast electricity price for PJM using neural network and similar days method," *IEEE Trans. Power Syst.*, 22(4), pp. 2058–2065.

Meng, K., Dong, Z., Wang, H. and Wang, Y. (2009), "Comparisons of Machine Learning Methods for Electricity Regional Reference Price Forecasting," *Advances in NN – ISNN 2009*. Lect. Notes in Comp. Science 5551**,** pp. 827–835.

Mori, H. and Awata, A. (2007), "Data Mining of Electricity Price Forecasting with Regression Tree and Normalized Radial Basis Function Network," in: *Proc. of 2007 IEEE International Conference on Systems, Man and Cybernetics (SMC2007)*, pp.3743–3748.

Möst, D. and Keles, D. (2010), "A survey of stochastic modelling approaches for liberalized electricity markets," *European Journal of Operational Research*, 207(2), pp. 543–556.

Mtunya, A. (2010), "Modeling electricity spot price time series using colored noise forces," *Master's Thesis*, University of Dar es Salaam, Tanzania.

Naeem, M. (2009), "A comparison of electricity spot prices simulation using ARMA-GARCH and mean reverting models," *Master's Thesis*, Lappeenranta University of Technology, Finland.

Nasr, G.E., Badr, E.A. and Younes, M.R. (2001), "Neural networks in forecasting electrical energy consumption: univariate and multivariate approaches," *International Journal of Energy Research*, 26(1), pp. 67–78.

Nelson, M., Hill, T., Remus, W. and O'Connor, M. (1999), "Time series forecasting using neural networks: should the data be deseasonalized first?" *Journal of Forecasting*, 18(5), 359–367.

Niu, L., Zhao, J. and Liu, M. (2009), "Application of relevance vector regression model based on sparse Bayesian learning to long-term electricity demand forecasting," in: *Proc. of the IEEE International Conference on Mechatronics and Automation*, Changchun, China, 2363–2367.

Nogales, F.J., Contreras, J., Conejo, A.J. and Espínola, R. (2002), "Forecasting next-day electricity prices by time series models," *IEEE Trans. Power Syst*., 17, pp. 342–348.

Nord Pool Spot (2013a), *The power market – how does it work*, Accessed 1 May 2013, available at http://www.nordpoolspot.com/How-does-it-work/.

Nord Pool Spot (2013b), *History*, Accessed 1 May 2013, available at http:// www.nordpoolspot.com/About-us/History/.

Nord Pool Spot (2013c), *Power System Overview*, Accessed 1 May 2013, available at http://www.nordpoolspot.com/Market-data1/Maps/Power-System-Overview/Power-System-Map/

Nord Pool Spot (2013d), *Market data*, Accessed 1 May 2013, available at http:// nordpoolspot.com/Market-data1/.

Olkkonen, H. (2011), *Discrete Wavelet Transforms-Biomedical applications*, Publisher: InTech., DOI: 10.5772/1818

Plazas, M. A., Conejo, A. J., and Prieto, F. J. (2005), "Multimarket optimal bid-ding for a power producer," *IEEE Trans. Power Syst*., 20(4), pp. 2041–2050.

Peng, H., Long, F. and Ding, C. (2005), "Feature selection based on mutual information: Criteria of max-dependency, max-relevance and min-redundancy," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 27(8), pp. 1226–1238.

Provost, F. and Domingos, P. (2000), "Well-trained PETs: Improving probability estimation trees," *Technical Report IS-00-04*. Stern School of Busines., New York University.

Rambharat, B.R., Brockwell, A.E. and Seppi, D.J. (2005), "A threshold autoregressive model for wholesale electricity prices," *Journal of the Royal Statistical Society Series C*, 54(2), pp. 287–300.

Reynolds, D. (2005), *Gaussian Mixture Models*, MIT Lincoln Laboratory.

Robinson, T.A. (2000), "Electricity pool prices: a case study in nonlinear time-series modeling," *Applied Economics*, 32(5), pp. 527–532.

Robnik-Šikonja, M. and Kononenko, I. (2003), "Theoretical and Empirical Analysis of ReliefF and RReliefF," *Machine Learning Journal*, 53, pp. 23–69.

Reis, A.J.R. and Alves da Silva, A.P. (2005), "Feature extraction via multiresolution analysis for short-term load forecasting," *IEEE Trans Power Syst*., 20, pp.189–198.

Rodriguez, C.P. and Anders, G.J. (2004), "Energy price forecasting in the Ontario competitive power system market," *IEEE Trans Power Syst.*, 19(3), pp. 366–374.

Rückstieß, T., Osendorfer, C. and van der Smagt, P.K. (2011), "Sequential Feature Selection for Classification," *Lecture Notes in Computer Science* ,7106, pp. 132–141.

Sadeh, J., Mashhadi, H.R. and Latifi, M.A. (2009), "A risk-based approach for bidding strategy in an electricity pay-as-bid auction," *European Transactions on Electrical Power*,19(1), pp. 39–55.

Shafie-khah, M., Moghaddam, M.P. and Sheikh-El-Eslami, M.K. (2011), "Price forecasting of day-ahead electricity markets using a hybrid forecast method," *Energy Convers. Management*, 52(5), pp. 2165–2169.

Specht, D.F. (1988), "Probabilistic neural networks for classification mapping, or associative memory," *IEEE International Conference on Neural Networks*, 1, pp. 525–532.

Statistics Finland (2013), *Energy Statistics*, Accessed 1 May 2013, available at http://www.stat.fi.

Swanson, N.R. and White, H.A. (1997), "Model selection approach to real-time macroeconomic forecasting using linear models and artificial neural networks," *The Review of Economics and Statistics*, MIT Press, 79(4), pp. 540–550.

Szkuta, B.R., Sanabria, L.A. and Dillon, T.S. (1999), "Electricity price short-term forecasting using artificial neural networks," *IEEE Trans. Power Syst*., 14 (3), pp. 851–857.

Tan, Z., Zhang, J., Wang, J. and Xu, J. (2010), "Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models," *Applied Energy*, 87(11), pp. 3606–3610.

Taylor, J.W., de Menezes, L.M. and McSharry, P.E. (2006), "A comparison of univariate methods for forecasting electricity demand up to a day ahead," *Int. J. Forecasting*, 22, pp. 1–16.

Theodoridis, S. and Koutroumbas, K. (2010), *An Introduction to Pattern Recognition: A MATLAB Approach*, Elsevier Inc., pp. 79–84.

Thomas, M.C. and Joy A.T. (1991), *Elements of Information Theory*, John Wiley & Sons, Inc., pp.12–23.

Tipping, M.E. (2001), "Sparse Bayesian Learning and Relevance Vector Machine," *J. of Mach. Learn. Res.*, 1, pp. 211–244.

Tseng, F.M., Yu, H.C. and Tzeng, G.H. (2002), "Combining neural network model with seasonal time series ARIMA model," *Technological Forecasting and Social Change*, 69, pp. 71–87.

Uhlenbeck, G.E. and Ornstein, L.S. (1930), "On the theory of Brownian motion," *Phys. Rev.*, 36, pp. 823–841.

Vahidinasab, V., Jadid, S. and Kazemi, A. (2008), "Day-ahead price forecasting in restructured power systems using artificial neural networks," *Electric Power Systems Research*, 78(8), pp. 1332–1342.

Vapnik, V. (1995), *The Nature of Statistical Learning Theory*, Springer, New York.

Voronin, S., Partanen, J. and Kauranne, T. (2013a), "A hybrid electricity price forecasting model for the Nordic electricity spot market," *International Transactions on Electrical Energy Systems.* Published online, DOI: 10.1002/etep.1734.

Voronin S. and Partanen J. (2013b), "Forecasting electricity price and demand using a hybrid approach based on wavelet transform, ARIMA and neural networks," *International Journal of Energy Research*, Published online, DOI: 10.1002/er.3067

Wang, A. and Ramsay, B. (1998), "A neural network based estimator for electricity spot pricing with particular reference to weekend and public holidays," *Neurocomputing*, 23, pp. 47–57.

Weather Underground (2013), *Historical Weather*, Accessed 1 May 2013, available at http://www.wunderground.com.

Weron, R. (2006), "Modeling and forecasting electricity loads and prices: A statistical approach," Wiley finance series, Wiley, 396, ISBN 13 978-0-470-05753-7 (HB).

Willis, H. (2002), *Spatial Electric Load Forecasting. Second edition, Revised and Expanded.* Marcel Dekker, Inc., pp. 124–128.

Wolak, F.A. (1997), "Market design and price behavior in restructured electricity markets," *Power*, pp. 79–142.

Wu, L. and Shahidehpour, M. (2010), "A hybrid model for day-ahead price forecasting," *IEEE Trans. Power Syst*., 25(3), pp. 1519–1530.

Yamin, H.Y., Shahidehpour, S.M. and Li, Z. (2004), "Adaptive short-term electricity price forecasting using artificial neural networks in the restructured power markets," *Electrical Power Energy Syst*.,26, pp. 571–581.

Yan, X. (2009), "Electricity market clearing price forecasting in a deregulated electricity market," *Master thesis*, University of Saskatchewan.

Yu L. and Liu, H. (2004), "Efficient feature selection via analysis of relevance and redundancy," *Machine Learning Research*, 5, pp.1205–1224.

Yun, Z., Quan, Z., Caixin, S., Shaolan, L., Yuming, L. and Yang, S. (2008), "RBF neural network and ANFIS-based short-term load forecasting approach in real-time price environment," *IEEE Trans. Power Syst*., 23 (3), pp. 853–858.

Zareipour, H., Canizares, C.A. and Bhattacharya, K. (2010), "Economic impact of electricity market price forecasting errors: a demand – side analysis," *IEEE Trans. Power Syst*., 25, pp. 254–262.

Zhang, G.P. (2003), "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, 50, pp. 159–175.

Zhang, G.P. and Qi, M. (2005), "Neural network forecasting for seasonal and trend time series," *European Journal of Operational Research*, 160, pp. 501–514.

Zhao, J.H., Dong, Z.Y., Li, X. and Wong, K.P. (2007a), "A framework for electricity price spike analysis with advanced data mining methods," *IEEE Trans. Power Syst*., 22(1), pp. 376–385.

Zhao, J.H., Dong, Z.Y. and Li, X. (2007b), "Electricity market price spike forecasting and decision making," *IET Gen., Trans., Distrib*., 1(4), pp. 647–654.

Appendix A: ML estimation

ML estimation is a method for estimating the parameters of a statistical model. The ML method views the parameters as quantities whose values are fixed but unknown. The best estimate of their value is defined to be the one that maximizes the probability of obtaining the samples actually observed. Suppose there are n data sets x_1, x_2,…, x_n with the samples in x_i having been drawn independently according to the probability law $p(x_1, x_2, ..., x_n | \theta)$. Such samples are independent and identically distributed (i.i.d.) random variables. The probability law $p(x_1, x_2, ..., x_n | \theta)$ is assumed to have a known parametric form. Because the samples are drawn independently, one can obtain

$$p(x_1, x_2, ..., x_n | \theta) = \prod_{k=1}^{n} p(x_k | \theta) \tag{A.1}$$

$p(x_1, x_2, ..., x_n | \theta)$ is called the likelihood of θ with respect to the set of samples. The ML estimate of θ is, by definition, the value θ that maximizes $p(x_1, x_2, ..., x_n | \theta)$. Therefore, the estimate corresponds to the value θ that best agrees with or supports the actually observed training samples.

It is usually easier to work with the logarithm of the likelihood than with the likelihood itself. Function $l(\theta | x_1, x_2, ..., x_n)$ is defined as the log-likelihood function:

$$l(\theta | x_1, x_2, ..., x_n) = \ln p(x_1, x_2, ..., x_n | \theta) \tag{A.2}$$

thus

$$l(\theta | x_1, x_2, ..., x_n) = \sum_{k=1}^{n} \ln p(x_k | \theta) \tag{A.3}$$

and

$$\nabla_\theta l = \sum_{k=1}^{n} \nabla_\theta \ln p(x_k | \theta) . \tag{A.4}$$

Thus, a set of necessary conditions for the ML estimate for θ can be obtained from the set of equations

$$\nabla_\theta l = 0 \tag{A.5}$$

Appendix B: Parameter estimations of SARIMA+GARCH

All the coefficients are statistically significant at the 5% level.

Table B.1. Parameter coefficients of the SARIMA and SARIMA+GARCH models estimated for original and adjusted price series for the Finnish day ahead energy market from 16 Sep 2009 to 14 Nov 2009.

	Original data		Adjusted data	
Model parame-ters	SARIMA(1,1,1) $((1,7),1,1)_{24}$	SARIMA(1,1,0) $((1,7),1,1)_{24}$+ GARCH(1,1)	SARIMA(1,1,1) $((1,7),1,1)_{24}$	SARIMA(1,1,0) $(1,7),1,1)_{24}$+ GARCH(1,1)
φ_1	0.62	0.14	0.67	0.08
	(0.02)	(0.05)	(0.04)	(0.03)
Φ_1	0.10	0.08	0.11	0.09
	(0.02)	(0.02)	(0.02)	(0.02)
Φ_2	0.24	0.13	0.26	0.13
	(0.01)	(0.01)	(0.02)	(0.01)
θ_1	-0.89		-0.89	
	(0.02)		(0.03)	
Θ_1	-0.85	-0.83	-0.81	-0.85
	(0.02)	(0.01)	(0.02)	(0.01)
Variance equation:				
C		2.41		0.48
		(0.11)		(0.03)
α		0.26		0.69
		(0.03)		(0.05)
β		0.09		0.30
		(0.02)		(0.01)

Notes: Standard errors are given in parenthesis

Appendix C: Distributions of simulated price paths

Figure C.1 indicates number of price values (Y-axis) that hit within the specific price interval (X-axis).

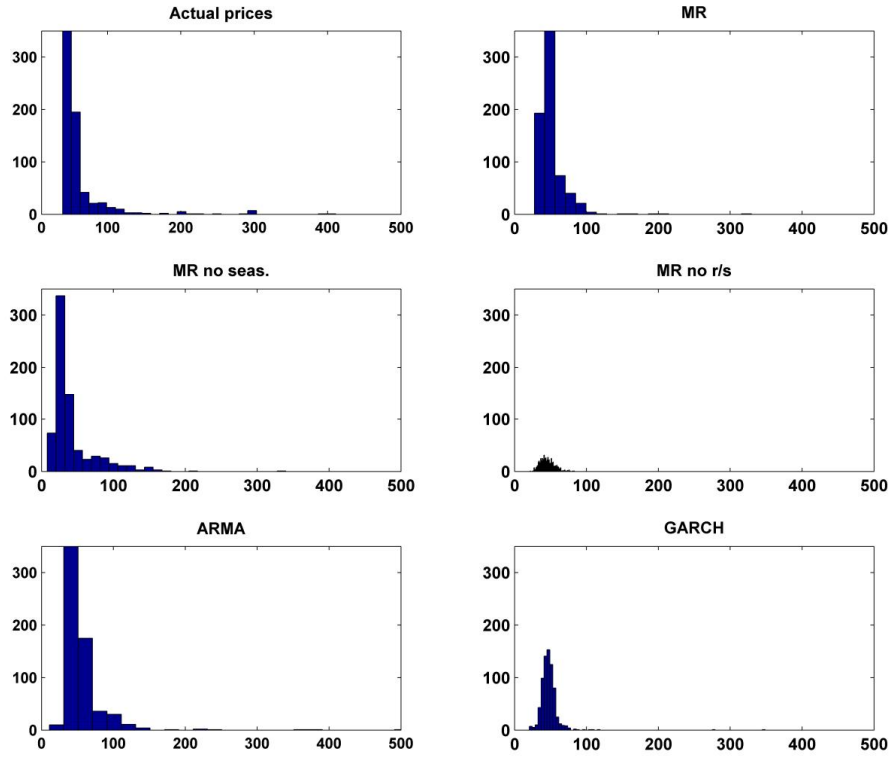

Figure C.1. Distributions of out-of-sample simulated price paths and original prices.

Appendix D: Hybrid electricity price forecasting model

D.1 GMM

When the probability density function (pdf) that describes the data points in a class is not known, it has to be estimated prior to the application of the Bayesian classifier. An arbitrary pdf can be modeled as a linear combination (weighted sum) of several pdfs. Therefore, if a high number of component distributions are used, any distribution can be approximated (Theodoridis and Koutroumbas, 2010). The probability density function for the samples is then given by

$$prob(x \mid \Delta) = \sum_{i=1}^{M} prob(x \mid \mu_i, \Sigma_i) P_{wi} \tag{D.1}$$

where x is a V-dimensional continuous-valued data vector (i.e., measurement of features), $P_{wi}, i = 1,\ldots,M$, are the mixture weights, and $prob(x|\mu_i, \sum_i), i=1,..,M,$ are the component Gaussian densities. Each component density is a V-variate Gaussian function of the form,

$$prob(x \mid \mu_i, \Sigma_i) = \frac{1}{(2\pi)^{V/2} \mid \Sigma_i \mid^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu_i)' \Sigma_i^{-1} (x-\mu_i)\right\} \tag{D.2}$$

with the mean vector μ_i and the covariance matrix $\sum_i$. The mixture weights satisfy the constraint $\sum_{i=1}^{M} P_w(\omega_i) = 1$. The complete Gaussian Mixture model (GMM) is parameterized by the mean vectors, covariance matrices, and mixture weights from all the component densities. These parameters are represented as

$$\Delta = \left\{P_w(\omega_i), \mu_i, \Sigma_i\right\} \quad i = 1,\ldots,M \tag{D.3}$$

Several techniques are available for estimating the parameters of the GMM. By far, the most popular and well-established method is the ML estimation. For a sequence of n training vectors $X=\{x_1,\ldots,x_k\}$, the GMM likelihood, assuming independence between the vectors, can be written as

$$prob(X \mid \Delta) = \prod_{i=1}^{k} prob(x_i \mid \Delta) \tag{D.4}$$

This expression is a nonlinear function of the parameters Δ, and direct maximization is not possible. However, ML parameter estimates can be obtained iteratively using a special case of the expectation-maximization (EM) algorithm. The basic idea of the EM algorithm is, beginning with an initial model Δ, to estimate a new model $\overline{\Delta}$ such that

$prob(X \mid \overline{\Delta}) \geq prob(X \mid \Delta)$.The new model then becomes the initial model for the next iteration, and the process is repeated until some convergence threshold is reached. The EM algorithm for GMM was described in (Reynolds, 2005).

Predicting the occurrence of a spike is a typical binary classification problem. The factors relevant to spikes can be considered the dimensions of the input vector $X=\{x_1,\ldots,x_k\}$ at each time point h where x_i, $i=1,2,\ldots,k$ is the value of a relevant factor. The object is to determine the label y for every input vector, where

$$y = \begin{cases} 1, non-spike \\ -1, spike \end{cases} \tag{D.5}$$

and y denotes whether a spike will occur.

A GMM based on a Bayesian classification algorithm is used to mine the database to find out the internal relationships between the electricity price spikes and the external factors. Basically, for a given input vector $X=\{x_1,\ldots,x_k\}$ and its class label $y \in \{c_1,c_2,\ldots,c_k\}$ the probability classifier calculates the probability that X belongs to class c_i for $i=1,2,\ldots,k$. X is labeled as class c_i, which has the largest probability.

D.2 **KNN**

KNN is a nonparametric technique. It stores all available cases and predicts the numerical target based on a similarity measure (e.g. distance functions). If the KNN is used for feature selection, K points in a given feature set X that are nearest to each point in the numerical target set are used. If the KNN is used for regression, the sum of weighted values of the K closest samples is computed as the unknown sample's value.

The distance metric from an unknown sample or a given target $Z=\{z_1,\ldots,z_k\}$ to the neighboring sample or feature set $X=\{x_1,\ldots,x_k\}$ is determined by the Euclidian distance between two real-valued vectors as given in Eq. (D.6)

$$dist(X,Z) = \sqrt{\sum_{i=1}^{k}(x_i - z_i)^2} \tag{D.6}$$

where k is a vector dimensionality.

D.3 **Parameter estimations of ARMA+GARCH based models**

Structures of proposed models are given below. All data are in daily resolution.

SARIMAX+(GARCH):

$$(1-\varphi_1 B)(1-B)(1-B^7)\,price_h = \eta_1 U_{n\_b.demand,h}(1-B)(1-B^7) + (1+\theta_1 B^1)(1+\Theta_1 B^7)a_h \quad \text{(D.7)}$$

ARMAX+(GARCH):

$$(1-\varphi_1 B-\varphi_7 B^7-\varphi_8 B^8)\,price_h^{decomposed} = \eta_1 U_{n\_b.demand,h}^{decomposed} + (1+\theta_1 B^7)a_h \quad \text{(D.8)}$$

GARCH (for both SARIMAX+GARCH/ARMAX+GARCH models):

$$\sigma_h{}^2 = C + \sum_{i=1}^{1}\alpha_i a_{t-i}{}^2 + \sum_{i=1}^{1}\beta_i \sigma_{h-i}{}^2, a_h \sim N\left(0,\sigma^2\right) \text{ at time } h, h = 1,2,3,...,H \quad \text{(D.9)}$$

The second and third columns of Table D.1 report the results of estimating (D.7-D.9).

Table D.1. ARMAX+GARCH based models obtained from the initial training data set.

	SARIMAX + GARCH(1,1)	ARMAX + GARCH(1,1)
φ_1	0.31	0.85
	(0.07)	(0.03)
φ_7		0.72
		(0.04)
φ_8		-0.58
		(0.05)
η_1	0.01	0.50
	(0.00)	(0.05)
θ_1	-0.67	-0.64
	(0.06)	(0.06)
Θ_1	-0.84	
	(0.02)	
Variance equation:		
C	2.29	0.00
	(0.14)	(0.000)
α	0.41	0.68
	(0.05)	(0.07)
β	0.39	0.16
	(0.04)	(0.04)
Model diagnostics:		
LB_Q	26.41	29.00
	[0.20]	[0.11]
ARCH	0.56	0.89
	[0.46]	[0.35]

Notes: Standard errors are given in parentheses and probability values in square brackets. LB_Q is the Ljung-Box Q-statistic to test for serial correlation in the residuals. ARCH tests for autoregressive conditional heteroscedasticity in the residuals.

D.4 Random walk model

The random walk was implemented as the naïve method since it is the most widely used and simplest naïve benchmark method in forecasting studies. The random walk forecast function is given as

$$X_h = X_{h-1}, \quad \text{at time} \quad h, h = 1, 2, 3, ..., H \tag{D.10}$$

where X_h and X_{h-1} can be considered price values at time h and $h-1$, respectively.

D.5 Performance measurements for the normal range price models

Table D.2. Performance measurements of the models predicting a normal range price up to seven days ahead.

	Horizon, [days]	MSE	MAE	AMAPE, [%]
Naïve benchmark	1	15.27	2.70	7.43
	2	18.67	3.12	8.60
	3	20.84	3.34	9.19
	4	26.14	3.68	10.15
	5	26.30	3.69	10.18
	6	26.20	3.65	10.16
	7	26.44	3.73	10.29
SARIMA+GARCH	1	8.14	2.09	5.76
	2	10.55	2.53	6.96
	3	11.69	2.67	7.36
	4	15.13	2.92	8.05
	5	17.90	3.18	8.76
	6	19.43	3.32	9.15
	7	22.44	3.51	9.66
SARIMAX+GARCH	1	7.35	2.00	5.51
	2	8.57	2.23	6.15
	3	8.96	2.28	6.29
	4	10.97	2.42	6.67
	5	12.51	2.65	7.30
	6	13.95	2.82	7.77
	7	16.56	2.95	8.14
ARMA+GARCH*	1	6.30	1.94	5.34
	2	7.41	2.10	5.80
	3	8.09	2.17	5.97
	4	8.53	2.27	6.25
	5	8.83	2.29	6.38
	6	9.67	2.39	6.63
	7	9.76	2.40	6.75
ARMAX+GARCH*	1	6.02	1.90	5.23
	2	6.87	2.00	5.50
	3	7.00	2.02	5.57

	4	7.18	2.05	5.87
	5	8.36	2.23	6.14
	6	8.75	2.27	6.37
	7	8.79	2.31	6.47
NN with raw data	1	10.14	2.14	5.89
	2	10.55	2.34	6.46
	3	10.80	2.45	6.76
	4	11.62	2.55	7.03
	5	14.33	2.78	7.68
	6	14.98	2.88	7.95
	7	15.73	3.00	8.27
NN with	1	7.09	2.08	5.73
decomposed data	2	7.91	2.20	6.05
	3	8.83	2.27	6.25
	4	9.19	2.38	6.55
	5	9.92	2.48	6.82
	6	11.13	2.56	7.08
	7	11.16	2.64	7.27
Combined with raw	1	7.41	1.96	5.40
data	2	8.34	2.17	5.97
	3	8.56	2.19	6.05
	4	11.20	2.35	6.49
	5	11.48	2.51	6.92
	6	13.23	2.67	7.35
	7	13.70	2.78	7.66
Combined with	1	6.01	1.89	5.20
decomposed data	2	6.58	1.95	5.36
	3	6.63	1.97	5.39
	4	7.18	2.05	5.63
	5	7.70	2.14	5.82
	6	8.16	2.16	5.95
	7	8.24	2.22	6.04

Notes: *Models tested by decomposed data.

Appendix E: Feature selection techniques

E.1 MI

A quantity called entropy is defined for any probability distribution. The entropy of a random variable is a measure of the uncertainty of the random variable; it is a measure of the amount of information required on the average to describe the random variable. Entropy then becomes the self-information of a random variable. The notion of entropy can be extended to define MI, which is a measure of the amount of information one random variable contains about another. MI is a special case of a more general quantity called relative entropy, which is a measure of the distance between two probability distributions (Cover and Thomas, 1991). The entropy *H(X)* of a discrete random variable *X* with values X_1, X_2,…, X_k and probabilities $P(X_1)$, $P(X_2)$,…, $P(X_k)$, respectively, is defined as follows:

$$H(X) = -\sum_{i=1}^{k} P(X_i)\log_2(P(X_i)). \tag{E.1}$$

The joint entropy *H(X,Y)* of a pair of discrete random variables (X,Y) with a joint distribution *P(X,Y)* is defined as

$$H(X,Y) = -\sum_{i=1}^{k}\sum_{j=1}^{m} P(X_i,Y_j)\log_2(P(X_i,Y_j)). \tag{E.2}$$

One can define the conditional entropy *H(Y/X)* of a random variable given by another as the expected value of the entropies of the conditional distributions *P(Y/X)*, averaged over the conditioning random variable. Therefore, the conditional entropy *H(Y/X)* quantifies the remaining uncertainty of *Y*, when *X* is known. The conditional entropy is defined as follows:

$$\begin{aligned} H(Y/X) = &\sum_{i=1}^{k} P(X_i)H(Y/X=X_i) \\ &-\sum_{i=1}^{k} P(X_i)\sum_{j=1}^{m} P(Y_j/X_i)\log_2(P(Y_j/X_i)) \\ &-\sum_{i=1}^{k}\sum_{j=1}^{m} P(X_i,Y_j)\log_2(P(Y_j/X_i)) \end{aligned} \tag{E.3}$$

MI is introduced as a measure of the amount of information that one random variable contains about another random variable. It is the reduction in the uncertainty of one random variable due to the knowledge of the other. Therefore, the mutual information *I(X,Y)* is the information found commonly in two random variables *X* and *Y* with a joint

probability mass function *P(X,Y)* and marginal probability mass functions *P(X)* and *P(Y)*, and can be defined as:

$$I(X,Y) = \sum_{i=1}^{k}\sum_{j=1}^{m} P(X_i,Y_j)\log_2 \frac{P(X_i,Y_j)}{P(X_i)P(Y_j)}. \quad \text{(E.4)}$$

Thus, MI of a random variable with itself is the entropy of the random variable. This is the reason why entropy is sometimes referred to as self-information. The relationship between *H(X)*, *H(Y)*, *H(X,Y)*, *H(X/Y)*, *H(Y/X)*, and *MI(X,Y)* is expressed in a Venn diagram (see Figure E.1). *MI(X,Y)* corresponds to the intersection of the information in *X* with the information in *Y*.

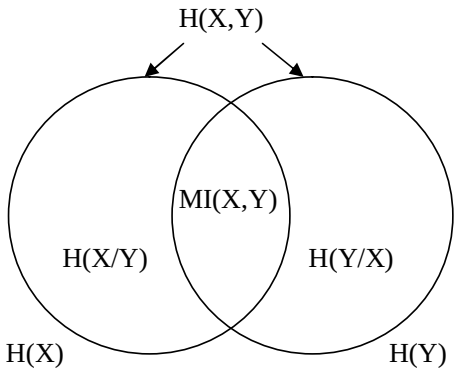

Figure E.1. Relationship between entropy and MI.

E.2 **Relief**

A problem of estimating the quality of attributes (features) is an important issue in the machine learning when a sufficient subset of features to describe the target concept is selected. Relief is a nonlinear technique that is able to detect conditional dependencies between attributes and provide a unified view on the attribute estimation in regression and classification (Robnik-Sikonja and Kononenko, 2003).

The original Relief algorithm considering two-class problems is used to estimate the quality of attributes according to how well their values distinguish between the instances that are near to each other (Kira and Rendell, 1992).

For that purpose, given a randomly selected instance *X*, Relief searches for its two nearest neighbors based on the Euclidian distance measure: one from the same class, called nearest hit $H_{nearest}$, and the other from a different class, called nearest miss

$M_{nearest}$. In the Relief algorithm, the weight of the i^{th} feature W_i is updated according to the following equation:

$$W_i = W_i + |X^{(i)} - M^{(i)}_{nearest}(X)| - |X^{(i)} - H^{(i)}_{nearest}(X)|, \; i = 1, 2, ..., I \tag{E.5}$$

where $X^{(i)}$ is the i^{th} attribute (feature) of the sample X; $M_{nearest}^{(i)}$ and $H_{nearest}^{(i)}$ are the i^{th} features of the nearest miss and the nearest hit of the sample X, respectively, and I is the number of candidate input features.

Therefore, a feature's weight is updated according to how well its values distinguish the sample from its nearest hit and nearest miss. A feature will receive a high weight if it differentiates between samples from opposite classes and has the same value for the samples of the same class (Robnik-Sikonja and Kononenko, 2003). The cycle is repeated for all randomly selected samples, and then, the candidate features are ranked according to the finally obtained weight values.

The original Relief algorithm is limited to two-class problems and becomes inappropriate for a problem of electricity price forecasting where the price value Y is continuous. The difference from the original Relief to adapt it to the regression problem is that, instead of one nearest hit and one nearest miss, Relief uses K nearest hits and misses and averages their contribution to W_i.

Relief's estimate W_i of the quality of the i^{th} attribute is an approximation of the following difference of probabilities (Kononenko, 1994):

$$\begin{aligned} W_i = {} & P(diff.value\ of\ i^{th}\ \ feature / nearest\ \ sample\ \ from\ diff.class) \\ & - P(diff.value\ of\ i^{th}\ \ feature / nearest\ \ sample\ \ from\ same.class) \end{aligned} \tag{E.6}$$

Then, instead of requiring the exact knowledge of whether two samples belong to the same class or not, a kind of probability that the price values Y of two samples are different is introduced. This probability can be modeled with the relative distance between the price value Y (class) of two samples.

Eq. (E.6) can be reformulated, so it can be directly evaluated using the probability that the price values Y of two samples are different:

$$P_{diff\_i} = P(diff.\ value\ of\ i^{th}\ \ feature / nearest\ \ samples) \tag{E.7}$$

$$P_{diff\_Y} = P(diff.\ value\ of\ Y / nearest\ \ samples) \tag{E.8}$$

$$P_{diff\_Y / diff\_i} = P(diff.\ value\ of\ Y / diff.\ value\ of\ i^{th}\ \ feature\ and\ nearest\ samples) \tag{E.9}$$

and from Eq. (E.6) according to Bayes' rule:

$$W_i = \frac{P_{diff\_Y/diff\_i} P_{diff\_i}}{P_{diff\_Y}} - \frac{(1 - P_{diff\_Y/diff\_i}) P_{diff\_i}}{1 - P_{diff\_Y}}. \quad \text{(E.10)}$$

The algorithm to estimate W_i by approximating terms defined by Eqs (E.7)–(E.9) can be found in (Robnik-Sikonja and Kononenko, 2003).

Relief-based algorithms have been recommended for feature selection when classification or regression approaches are adopted (Kononenko, 1994; Robnik-Sikonja, Kononenko, 2003).

Appendix F: Two-step feature selection algorithm

Relevance values are calculated for the target variable (price) and each candidate feature. Further, all candidate features with a relevance value (with respect to the target feature) $\geq V_1$ (0.61) are selected by the relevancy filter (see Figure F.1 and Table F.1). The selected features indicate the short-run trend ($price_{h-1}$, $price_{h-2}$, $price_{h-3}$, $price_{h-4}$), and the daily ($price_{h-24}$) and weekly periodicity ($price_{h-168}$) of the price series.

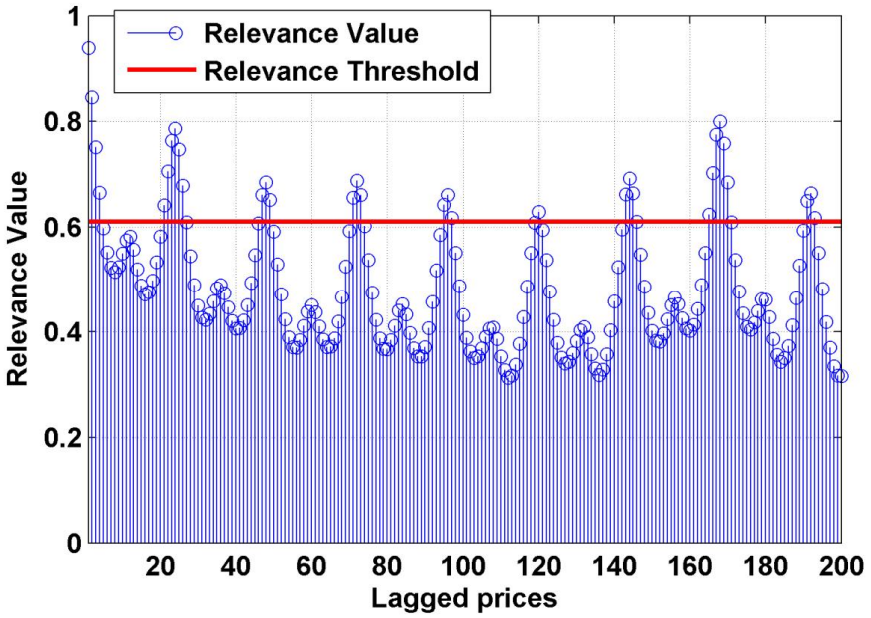

Figure. F.1.Relevance values between the candidate inputs and the target variable.

The results obtained when the redundancy filter is implemented are presented in Table F.2. Competing candidate inputs, their mutual relevance values (correlation coefficient), relevance values between each candidate input and a target variable (correlation coefficient), and the removed inputs are presented in Table F.2.

Table F.1. Results obtained when the relevance filter is implemented.

Selected input	Correlation coefficient
$price_{h-1}$	0.94
$price_{h-2}$	0.85
$price_{h-3}$	0.75
$price_{h-4}$	0.67
$price_{h-21}$	0.64
$price_{h-22}$	0.71
$price_{h-23}$	0.76
$price_{h-24}$	0.79
$price_{h-25}$	0.75
$price_{h-26}$	0.68
$price_{h-47}$	0.66
$price_{h-48}$	0.68
$price_{h-49}$	0.65
$price_{h-71}$	0.66
$price_{h-72}$	0.69
$price_{h-73}$	0.66
$price_{h-95}$	0.64
$price_{h-96}$	0.66
$price_{h-97}$	0.62
$price_{h-120}$	0.63
$price_{h-143}$	0.66
$price_{h-144}$	0.69
$price_{h-145}$	0.66
$price_{h-165}$	0.62
$price_{h-166}$	0.70
$price_{h-167}$	0.77
$price_{h-168}$	0.80
$price_{h-169}$	0.76
$price_{h-170}$	0.69
$price_{h-191}$	0.65
$price_{h-192}$	0.66
$price_{h-193}$	0.62

Table F.2. Results obtained when the redundancy filter is implemented.

Candidate input x_a	Candidate input x_b	RV(x_a, x_b)	RV(x_a,y)	RV(x_a,y)	Removed
$price_{h-1}$	$price_{h-2}$	0.94	0.94	0.85	$price_{h-2}$
$price_{h-3}$	$price_{h-4}$	0.94	0.75	0.67	$price_{h-4}$
$price_{h-21}$	$price_{h-22}$	0.94	0.64	0.71	$price_{h-21}$
$price_{h-22}$	$price_{h-23}$	0.94	0.71	0.76	$price_{h-22}$
$price_{h-23}$	$price_{h-24}$	0.94	0.76	0.79	$price_{h-23}$
$price_{h-24}$	$price_{h-25}$	0.94	0.79	0.75	$price_{h-25}$
$price_{h-47}$	$price_{h-48}$	0.94	0.66	0.68	$price_{h-47}$
$price_{h-48}$	$price_{h-49}$	0.94	0.68	0.65	$price_{h-49}$
$price_{h-71}$	$price_{h-72}$	0.94	0.66	0.69	$price_{h-71}$
$price_{h-72}$	$price_{h-73}$	0.94	0.69	0.66	$price_{h-73}$
$price_{h-95}$	$price_{h-96}$	0.94	0.64	0.66	$price_{h-95}$
$price_{h-96}$	$price_{h-97}$	0.94	0.66	0.62	$price_{h-97}$
$price_{h-143}$	$price_{h-144}$	0.94	0.66	0.69	$price_{h-143}$
$price_{h-144}$	$price_{h-145}$	0.94	0.69	0.66	$price_{h-145}$
$price_{h-165}$	$price_{h-166}$	0.94	0.62	0.70	$price_{h-165}$
$price_{h-166}$	$price_{h-167}$	0.94	0.70	0.77	$price_{h-166}$
$price_{h-167}$	$price_{h-168}$	0.94	0.77	0.80	$price_{h-167}$
$price_{h-168}$	$price_{h-169}$	0.94	0.80	0.76	$price_{h-169}$
$price_{h-191}$	$price_{h-192}$	0.94	0.65	0.66	$price_{h-191}$
$price_{h-192}$	$price_{h-193}$	0.94	0.66	0.62	$price_{h-193}$

Appendix G: RVM and RF forecasting engines

G.1 RVM

Relevance vector machine (RVM), a statistical learning technique based on the Bayesian estimation theory, is developed for regression and classification problems. A detailed mathematical description of the RVM is given in (Tipping, 2001).

Note that in most of real-world problems, the training data are not linearly separable. The method to deal with nonlinear data is to use a map function to map the training data from the input space into some high-dimensional feature space so that they will become linearly separable in the feature space. The related kernel function is used to avoid explicit knowledge of the high-dimensional mapping (Vapnik, 1995). A Gaussian RBF kernel with its specific value of spread σ_{RVM} is selected for the application of RVM (Meng et al., 2009).

G.2 RF

RF produces a forest of a specific number of decision trees (N_{tree}) whose predictions are combined to make an overall prediction for the forest. Bagging is a method to develop improved estimating class probabilities from the decision tree classification algorithm.

Mathematical description of the decision tree and the bagging method can be found in (Breiman, 1984; Provost et al., 2000; Ali et al., 2012).

G.3 RVM and RF with different feature selection techniques

Table G1. AMAPE (%) for the price forecast of the four weeks of the Finnish day-ahead energy market in the year 2009 produced by the RVM employing different feature selection techniques. Input data: historical prices.

	Feature selection technique				
Test week	Linear Corr./ —	Relief/ —	MI/ —	KNN/ —	Sequential selection
Fall	5.38	5.01	5.20	6.01	5.88
Summer	9.97	10.55	10.13	10.83	11.25
Spring	6.21	6.54	6.01	5.99	6.89
Winter	6.02	7.01	6.13	6.88	7.20
Average	6.90	7.28	6.87	7.43	7.81
	Linear Corr. / Linear Corr.	MI/ MI	Relief/ Linear Corr.	Relief / MI	Heuristic
Fall	4.69	4.78	4.86	4.81	5.81
Summer	9.38	9.27	10.37	10.05	10.57
Spring	5.87	5.94	5.31	4.99	8.52
Winter	5.21	5.03	5.47	5.25	7.48
Average	6.29	6.26	6.50	6.28	8.10

Table G.2. AMAPE (%) for the price forecast of the four weeks of the Finnish day-ahead energy market in the year 2010 produced by the RVM employing different feature selection techniques. Input data: historical prices.

	Feature selection technique				
Test week	Linear Corr. / —	Relief / —	MI/ —	KNN/ —	Sequential selection
Fall	3.22	3.50	2.66	3.15	4.99
Summer	5.64	6.02	5.23	5.79	5.66
Spring	8.79	9.03	8.90	8.81	8.13
Winter	23.77	19.88	20.66	19.05	18.90
Average	10.36	9.61	9.36	9.20	9.42
	Linear Corr. / Linear Corr.	MI/ MI	Relief / Linear Corr.	Relief/ MI	Heuristic
Fall	2.64	2.54	3.03	3.17	3.05
Summer	4.79	4.40	5.64	6.96	7.05
Spring	8.58	7.23	8.83	8.71	9.13
Winter	19.00	17.86	19.19	18.32	26.86
Average	8.75	8.03	9.17	9.29	11.52

Table G3. AMAPE (%) for the price forecast of the four weeks of the Finnish day-ahead energy market in the year 2009 produced by the RF regression employing different feature selection techniques. Input data: historical prices.

	Feature selection technique				
Test week	Linear Corr. / —	Relief/ —	MI —	KNN/ —	Sequential selection
Fall	5.56	5.98	5.47	5.80	5.68
Summer	9.26	9.90	9.29	9.93	10.15
Spring	5.99	6.10	5.68	5.99	6.10
Winter	6.09	6.29	6.00	6.10	6.89
Average	6.73	7.07	6.61	6.96	7.13
	Linear Corr. / Linear Corr.	MI/ MI	Relief / Linear Corr.	Relief / MI	Heuristic
Fall	4.50	4.63	5.06	5.31	5.93
Summer	8.95	9.19	9.31	10.28	10.22
Spring	5.63	5.63	5.59	5.76	7.12
Winter	5.77	5.52	5.30	5.67	6.25
Average	6.21	6.24	6.32	6.76	7.38

Table G4. AMAPE (%) for the price forecast of the four weeks of the Finnish day-ahead energy market in the year 2010 produced by the RF regression employing different feature selection techniques. Input data: historical prices.

	Feature selection technique				
Test week	Linear Corr. / —	Relief/ —	MI/ —	KNN/ —	Sequential selection
Fall	3.08	3.12	2.68	3.44	3.73
Summer	6.90	5.17	4.68	4.76	6.06
Spring	10.03	9.63	9.23	9.05	9.33
Winter	17.24	18.01	16.89	17.25	17.96
Average	9.31	8.98	8.37	8.63	9.27
	Linear Corr. / Linear Corr.	MI/ MI	Relief / Linear Corr.	Relief / MI	Heuristic
Fall	2.66	2.53	3.13	3.17	3.03
Summer	5.64	4.45	5.79	4.96	6.76
Spring	9.14	8.47	8.83	8.71	9.74
Winter	16.60	14.59	16.45	16.32	23.15
Average	8.51	7.51	8.55	8.29	10.67

Appendix H: Simultaneous price and demand forecasting

H.1 Inputs selected by the two-step feature selection

The selected inputs and their ranked correlation coefficients (with respect to the target variable) are given for price and demand prediction (see Table H.1).

Table H.1. Feature selection results for the approximation wavelet subseries of price and demand for 15 Nov 2009 in the Finnish day-ahead energy market.

Selected features for the approximation subseries of price at hour h		Selected features for the approximation subseries of demand at hour h	
Selected inputs	Correlation coefficient	Selected inputs	Correlation coefficient
$A3_{price,h-1}$	0.99	$A3_{demand,\ h-1}$	0.99
$A3_{price,h-2}$	0.97	$A3_{SARIMA\_demand,h}$	0.97
$A3_{SARIMA\_price,h}$	0.82	$A3_{demand,\ h-167}$	0.92
$A3_{SARIMA\_demand,h}$	0.79	$A3_{demand,\ h-3}$	0.87
$A3_{demand,h-167}$	0.77	$demand_{h-167}$	0.83
$A3_{price,h-170}$	0.74	$A3_{demand,\ h-4}$	0.79
$A3_{price,h-167}$	0.73	$demand_{h-167}$	0.83
$A3_{price,h-169}$	0.73	$demand_{h-146}$	0.69
$A3_{price,h-25}$	0.71	$A3_{price,h-145}$	0.66
$A3_{demand,h-170}$	0.69	$A3_{price,h-144}$	0.66
$A3_{demand,h-172}$	0.68	$A3_{SARIMA\_price,h}$	0.62
$A3_{price,h-21}$	0.67	$A3_{price,h-1}$	0.62
$A3_{demand,h-146}$	0.66	$A3_{price,h-142}$	0.59
$A3_{demand,h-194}$	0.66	$A3_{demand,\ h-48}$	0.59
$A3_{demand,h-142}$	0.66	$A3_{demand,\ h-20}$	0.59
$price_{h-169}$	0.66	$A3_{demand,\ h-140}$	0.55
$A3_{demand,h-22}$	0.65	$A3_{price,\ h-167}$	0.55
$A3_{demand,h-4}$	0.64	$A3_{price,h-3}$	0.55
$price_{h-170}$	0.64	$A3_{demand,\ h-94}$	0.54
$A3_{price,h-6}$	0.63	$demand_{h-163}$	0.54
$A3_{price,h-165}$	0.63		

H.2 **Model performance for a period of one year**

Table H.2. Results of the proposed simultaneous price and demand forecasts for all 52 weeks of the year 2009 in the Finnish day-ahead energy market.

Week	1	2	3	4	5	6
AMAPE	5.08/4.02	4.45/2.72	5.76/2.09	3.20/1.84	2.88/0.99	3.78/1.65
Week	7	8	9	10	11	12
AMAPE	4.17/1.25	3.66/1.17	2.48/1.18	5.10/2.17	3.59/1.56	2.93/1.60
Week	13	14	15	16	17	18
AMAPE	5.97/1.74	3.38/1.31	5.68/3.91	6.43/2.89	5.79/1.90	7.07/2.17
Week	19	20	21	22	23	24
AMAPE	6.20/1.68	4.17/1.98	6.35/3.71	7.10/2.94	4.46/1.40	5.40/2.33
Week	25	26	27	28	29	30
AMAPE	6.20/4.04	3.97/2.69	5.93/0.87	4.45/1.39	5.58/1.60	8.06/1.61
Week	31	32	33	34	35	36
AMAPE	7.25/2.13	6.58/1.85	6.34/1.91	6.86/1.05	5.49/0.98	6.14/0.95
Week	37	38	39	40	41	42
AMAPE	5.00/1.16	4.49/1.12	4.96/1.14	5.87/1.81	4.89/1.73	3.27/1.73
Week	43	44	45	46	47	48
AMAPE	2.55/1.96	3.85/2.79	1.92/2.05	3.18/1.85	2.61/2.16	3.82/1.39
Week	49	50	51	52		
AMAPE	7.22/1.55	6.38/2.22	8.73/2.86	7.09/2.77		

Average of price AMAPEs = 5.07%; Average of demand AMAPEs = 1.95%

Table H.2 shows that the AMAPE values for all 52 weeks of the year 2009 are close to the AMAPE values of the four considered test weeks. The averages of the AMAPEs for the 52 weeks are slightly higher than or similar to the averages of the AMAPEs of the four test weeks for both the price (5.07% versus 4.48%) and the demand (1.95% versus 1.79%).

Appendix I: Iterative forecasting methodology with separate normal price and price spike frameworks

I.1 PNN

Probabilistic neural networks (PNN) are a kind of a radial basis network suitable for classification problems. A PNN is closely related to the Parzen window probability density function estimator (Duda et al., 2001). A PNN is organized into a multilayered feed-forward network with four layers: an input layer (set of measurements), a pattern layer (the Gaussian functions), a summation layer (average operation of the outputs from the second layer for each class), and an output layer (a vote, selecting the largest value). Mathematical details of PNN can be found in (Specht, 1988). The spread of a Gaussin RBF σ_{PNN} is an adjustable parameter of the PNN. If the spread is close to zero, the network acts as a nearest neighbor classifier. As the spread becomes larger, the designed network takes into account several nearby design vectors.

I.2 Forecasting performance of competing approaches

Forecasted price curves obtained from SARIMA, WT+SARIMA, WT+NN, and WT+SARIMA+NN models for the four spiky weeks of the year 2010 in the Finnish day-ahead energy market are presented in Figures I.1–I.4, respectively.

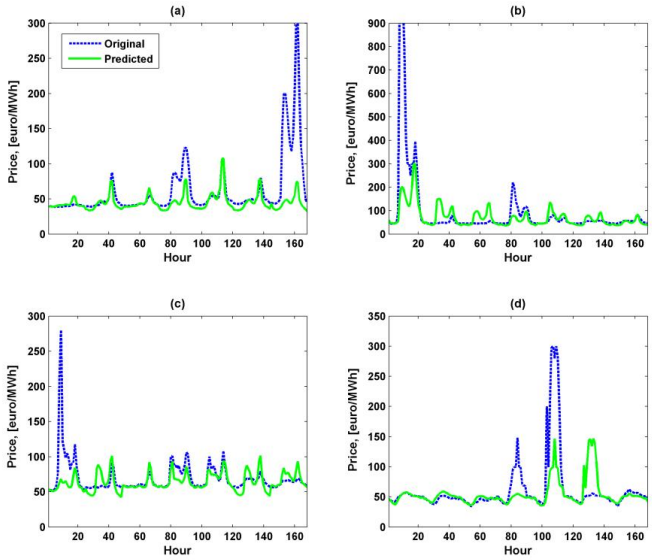

Figure I.1. SARIMA: (a) Week 1; (b) Week 2; (c) Week 5; (d) Week 28.

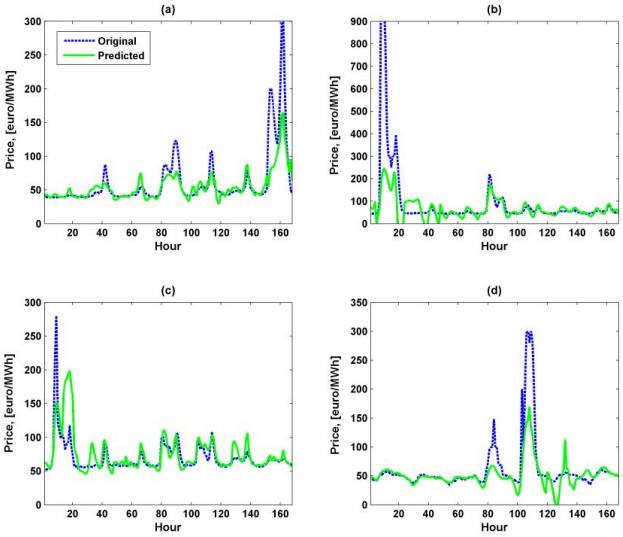

Figure I.2. WT+SARIMA: (a) Week 1; (b) Week 2; (c) Week 5; (d) Week 28.

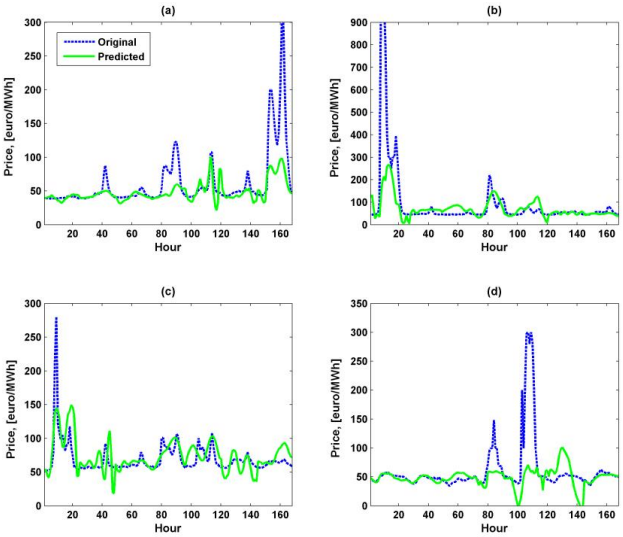

Figure I.3. WT+NN: (a) Week 1; (b) Week 2; (c) Week 5; (d) Week 28.

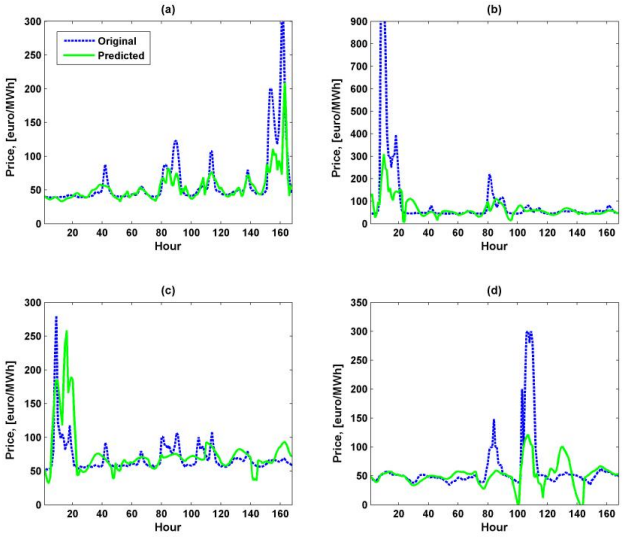

Figure I.4. WT+NN+SARIMA: (a) Week 1; (b) Week 2; (c) Week 5; (d) Week 28.

Appendix J: Short-term operation planning

The support decision-making tool generating an optimal production schedule and bidding strategy for a demand-side market customer based on the electricity price forecast is discussed.

Scheduling of the demand-side market participants' operation is formulated as an optimization problem, which is solved to minimize the expected energy costs of the market participant. The problem of minimizing electricity costs over the next 24 hours for a demand-side market participant can be given as:

$$\min_{P_h} Cost = \sum_{h=1}^{24} price_h \cdot P_h \tag{J.1}$$

where P_h is the net power purchased from the market at hour h and $price_h$ is the market price at hour h. It should be noted that Eq. J.1 is subject to technical constraints (e.g. generation constraints, transmission constraints).

In a real case, when the optimization problem to schedule day-ahead operation has to be solved, realized electricity market prices are not known. Therefore, price forecasts generated from the corresponding forecasting model are given as the expected day-ahead prices and considered in Eq. J.1 as the realized market prices.

A demand-side market customer considered in this study is presented as a typical CHP industry process having own on-site generation and both thermal and electrical energy demand. The thermal and electrical energy demand profiles are assumed to be known for an operational day (see Figure J.1).

The real case of CHP power plant operation within an electricity market environment can be presented as in Figure J.2.

Therefore, the optimization objective function can be given as:

$$\min_{h} Cost = \sum_{h=1}^{24} Cost_h^{Production} + \sum_{h=1}^{24} price_h^{forec} \cdot P_{Elspot,h}^{Import} - \sum_{i=h}^{24} price_h^{forec} \cdot P_{Elspot,h}^{Export} \tag{J.2}$$

where $Cost_h^{Production} = 300+40\cdot(P_h^{Heat} + P_h^{Electr.})+0.002\cdot(P_h^{Heat} + P_h^{Electr.})^2$ (euro/hour) is assumed to be an approximation of the total production costs based on the amount of generated heat and electrical energy at hour h; $price_h^{forec}$ is the forecasted market price at hour h; P_h^{Heat} is the heat energy generated by CHP; $P_h^{Electr.}$ is the electrical energy generated by CHP; $P_{Elspot,h}^{Import}$ is the electrical energy imported from the market at hour h, and $P_{Elspot,h}^{Export}$ is the electrical energy exported to the market at hour h.

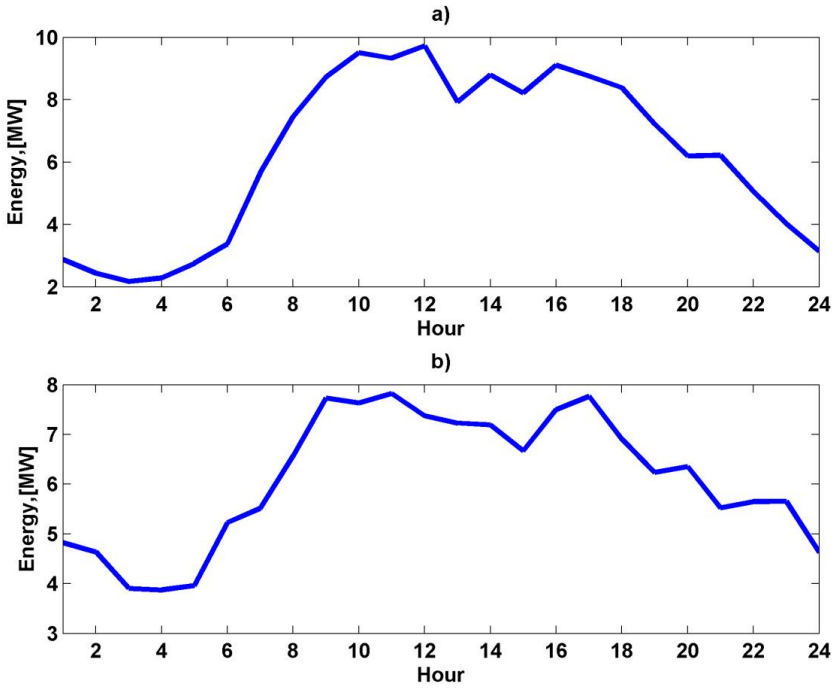

Figure J.1. a) Thermal and b) electrical energy demand profiles.

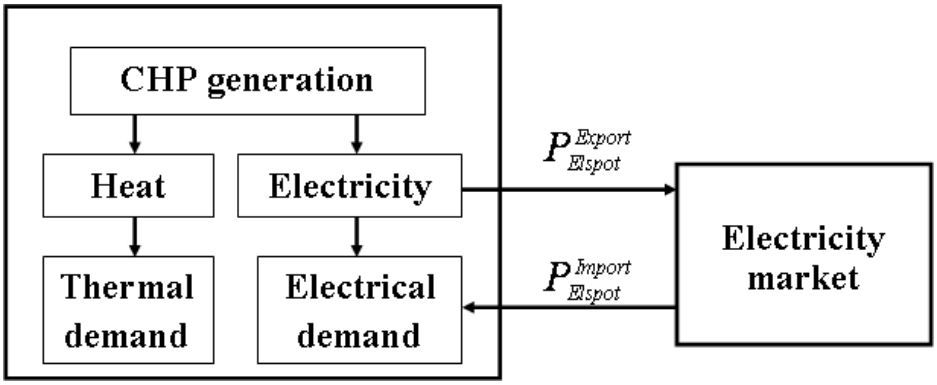

Figure J.2. Scheme of the CHP power plant operation within an electricity market.

The thermal demand must be met at all hours by the thermal energy produced at the power plant. The electrical demand must be met either by the energy produced by the power plant or energy purchased from the market. The energy balance constraints with added transmission losses are written as:

$$P_h^{Heat} = Heat_h \tag{J.3}$$

$$P_h^{Electr.local} + P_{Elspot,h}^{Import} = Electricity_h + \alpha^{loss} \cdot (P_{Elspot,h}^{Import})^2 \tag{J.4}$$

$$P_h^{Electr.} = P_h^{Electr.local} + P_{Elspot,h}^{Export} + \alpha^{loss} \cdot (P_{Elspot,h}^{Export})^2 \tag{J.5}$$

$$P_{Elspot,h}^{Export} \le transmission\_limit \tag{J.6}$$

$$P_{Elspot,h}^{Import} \le transmission\_limit \tag{J.7}$$

$$P_{min}^{Heat} \le P_h^{Heat} \le P_{max}^{Heat} \tag{J.8}$$

$$P_{min}^{Electr.} \le P_h^{Heat} \le P_{max}^{Electr.} \tag{J.9}$$

where $Heat_i$ and $Electricity_i$ are the hourly thermal electrical demand, respectively; $P_i^{Electr.local}$ is the electric power from the power plant supplying local electricity demand at hour i; P_{min}^{Heat}, P_{max}^{Heat}, $P_{min}^{Electr.}$, $P_{max}^{Electr.}$ are the heat/electricity generation limits of the CHP power plant, and α^{loss} is the transmission loss coefficient.

To maintain the CO_2 emissions produced by the CHP power plant, a certain constraint on the volume of the produced CO_2 is given as:

$$\alpha^{CO_2} \cdot \sum_{h=1}^{24} (P_h^{Heat} + P_h^{Electr.}) \le CO_{2\_Limit} \tag{J.10}$$

where α^{CO2} is the coefficient indicating the volume of CO_2 (ton) produced per MWh of energy generated by the CHP power plant; CO_{2_Limit} is the specified limit of CO_2 produced (ton/day).

With the thermal and electrical demand profiles of the CHP power plant, the optimization problem has been solved for a CHP power plant operating within the Finnish day-ahead energy market on a single day, 15 Feb 2010. The values of α^{loss}, α^{CO2}, CO_{2_Limit} are considered to be 0.0012, 0.43 (ton/MWh), 100 (ton/day), respectively. The energy import/export/generation schedules of the CHP power plant (presented as in Figure J.2) for a single test day when using actual prices and two different price forecasts of low and high accuracy are shown in Figures J.3–J.4. Here, the forecasts of high and low accuracy correspond to price forecasts produced by the proposed separate forecasting methodology and simple SARIMA, respectively. These two forecasting models are considered in Chapter 6 of the doctoral thesis.

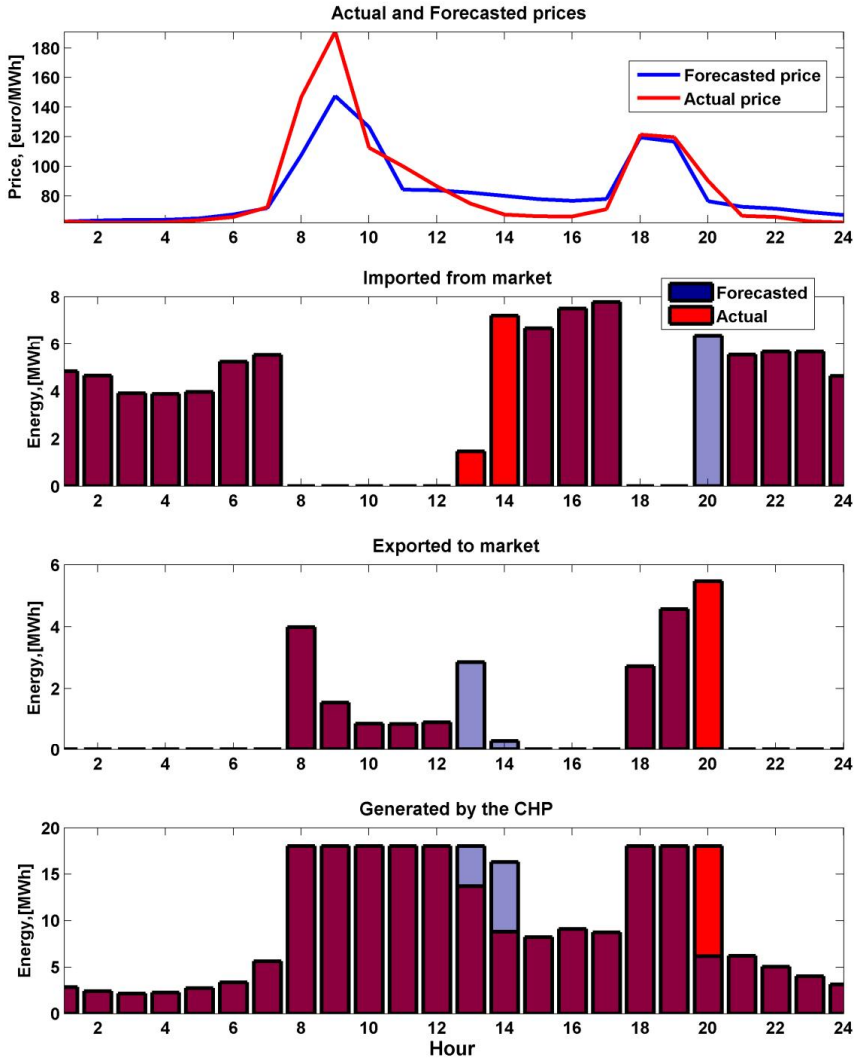

Figure J.3. Energy scheduling of the CHP power plant during a single day, 15 Feb 2010, based on the price forecast obtained from the separate forecasting methodology proposed.

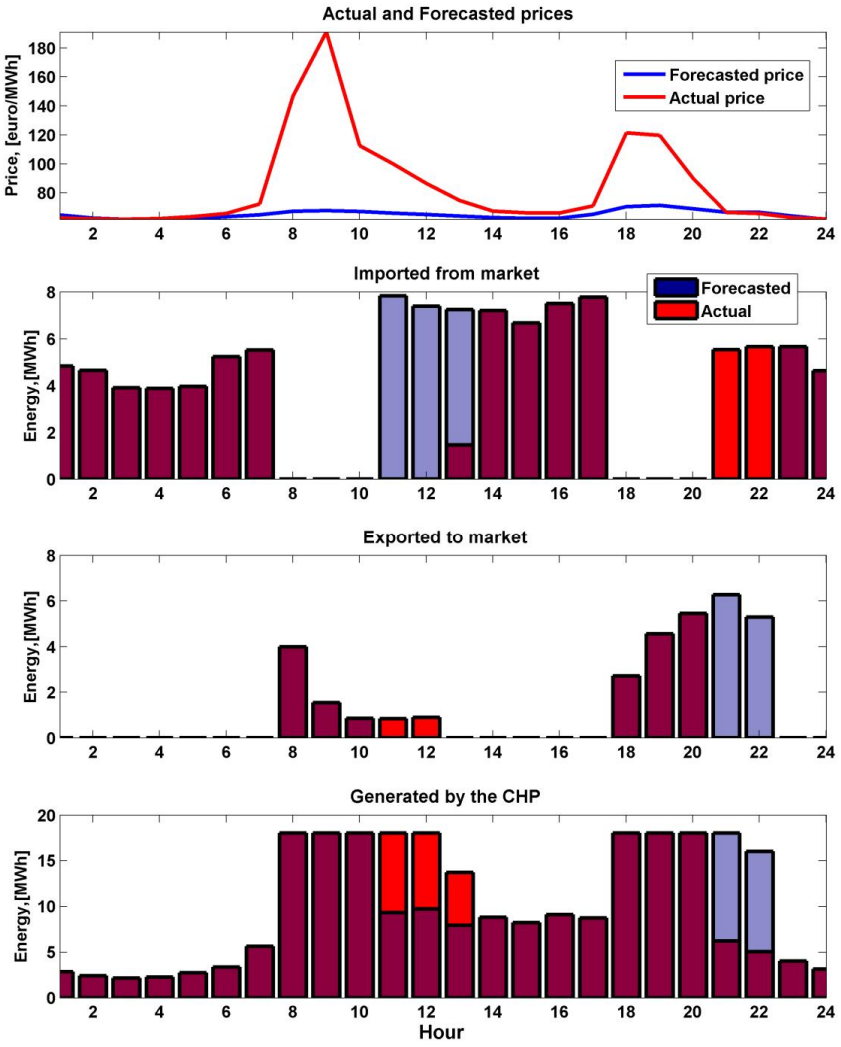

Figure J.4. Energy scheduling of the CHP power plant during a single day, 15 Feb 2010, based on the price forecast obtained from the SARIMA model.

The total CHP costs based on three different market price paths are presented in Table J.1.

Table J.1. The total CHP power plant costs when using actual market prices and two different price forecasts for a single day, 15 Feb 2010.

Actual costs, [euro]	Estimated costs when using the proposed separate methodology, [euro]	Estimated costs when using the SARIMA model, [euro]
19542	19700	20153

The cost deviation information aims at evaluating the overall economic impact of using the specific market price forecast in the operation scheduling of the specific market participant. The cost deviation is based upon the following relation

$$Cost\ Deviation = \frac{Estimated\ Costs - Actual\ costs}{Estimated\ Costs} \cdot 100\% \qquad \text{(J.11)}$$

Therefore, the cost deviation values can only be calculated after the realized market prices are available. Cost deviations (%) (related to the actual power plant costs corresponding to the ideal schedules) and AMAPE (%) values when two different price forecasts used are illustrated in Figure J.5.

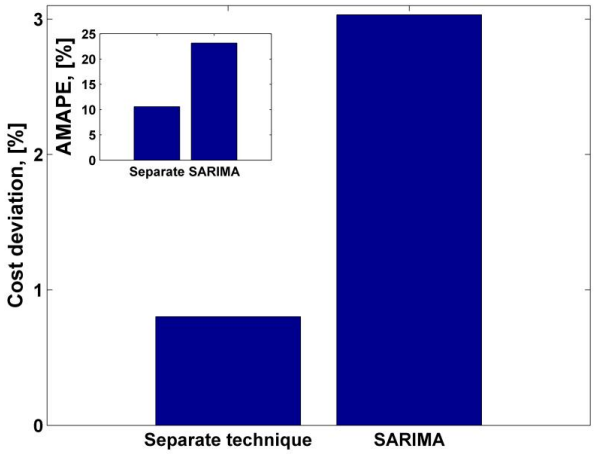

Figure J.5. Cost deviations of the CHP power plant and the AMAPE values when two different price forecasts are used for a single day, 15 Feb 2010.

In this study, scheduling of the next-day operation of the CHP power plant based on the 24 hours ahead electricity price forecasts of low and high accuracy is described. As demonstrated, the electricity market price forecast can be effectively employed to schedule the operation 24 hours ahead. Linear correlation between the forecast error measures and the corresponding cost deviations exists.

ACTA UNIVERSITATIS LAPPEENRANTAENSIS

489. PAANANEN, MIKKO. On innovative search: the use of internal and external sources of innovation among Finnish innovators. 2012. Diss.

490. BELOVA, POLINA. Quasiclassical approach to the vortex state in iron-based superconductors. 2012. Diss.

491. HIETANEN, IIRO. Design and characterization of large area position sensitive radiation detectors. 2012. Diss.

492. PÄSSILÄ, ANNE. A reflexive model of research-based theatre Processing innovation of the cross-road of theatre, reflection and practice-based innovation activities. 2012. Diss.

493. RIIPINEN, TOMI. Modeling and control of the power conversion unit in a solid oxide fuel cell environment. 2012. Diss.

494. RANTALAINEN, TUOMAS. Simulation of structural stress history based on dynamic analysis. 2012. Diss.

495. SALMIMIES, RIINA. Acidic dissolution of iron oxides and regeneration of a ceramic filter medium. 2012. Diss.

496. VAUTERIN, JOHANNA JULIA. The demand for global student talent: Capitalizing on the value of university-industry collaboration. 2012. Diss.

497. RILLA, MARKO. Design of salient pole PM synchronous machines for a vehicle traction application. 2012. Diss.

498. FEDOROVA, ELENA. Interdependence of emerging Eastern European stock markets. 2012. Diss.

499. SHAH, SRUJAL. Analysis and validation of space averaged drag model for numerical simulations of gas-solid flows in fluidized beds. 2012. Diss.

500. WANG, YONGBO. Novel methods for error modeling and parameter identification of redundant hybrid serial-parallel robot. 2012. Diss.

501. MAXIMOV, ALEXANDER. Theoretical analysis and numerical simulation of spectral radiative properties of combustion gases in oxy/air-fired combustion systems. 2012. Diss.

502. KUTVONEN, ANTERO. Strategic external deployment of intellectual assets. 2012. Diss.

503. VÄISÄNEN, VESA. Performance and scalability of isolated DC-DC converter topologies in low voltage, high current applications. 2012. Diss.

504. IKONEN, MIKA. Power cycling lifetime estimation of IGBT power modules based on chip temperature modeling. 2012. Diss.

505. LEIVO, TIMO. Pricing anomalies in the Finnish stock market. 2012. Diss.

506. NISKANEN, ANTTI. Landfill gas management as engineered landfills – Estimation and mitigation of environmental aspects. 2012. Diss.

507. QIU, FENG. Surface transformation hardening of carbon steel with high power fiber laser. 2012. Diss.

508. SMIRNOV, ALEXANDER. AMB system for high-speed motors using automatic commissioning. 2012. Diss.

509. ESKELINEN, HARRI, ed. Advanced approaches to analytical and systematic DFMA analysis. 2013.

510. RYYNÄNEN, HARRI. From network pictures to network insight in solution business – the role of internal communication. 2013. Diss.

511. JÄRVI, KATI. Ecosystem architecture design: endogenous and exogenous structural properties. 2013. Diss.

512. PIILI, HEIDI. Characterisation of laser beam and paper material interaction. 2013. Diss.

513. MONTO, SARI. Towards inter-organizational working capital management. 2013. Diss.

514. PIRINEN, MARKKU. The effects of welding heat input usability of high strength steels in welded structures. 2013. Diss.

515. SARKKINEN, MINNA. Strategic innovation management based on three dimensions diagnosing innovation development needs in a peripheral region. 2013. Diss.

516. MAGLYAS, ANDREY. Overcoming the complexity of software product management. 2013. Diss.

517. MOISIO, SAMI. A soft contact collision method for real-time simulation of triangularized geometries in multibody dynamics. 2013. Diss.

518. IMMONEN, PAULA. Energy efficiency of a diesel-electric mobile working machine. 2013. Diss.

519. ELORANTA, LEENA. Innovation in a non-formal adult education organisation – multi-case study in four education centres. 2013. Diss.

520. ZAKHARCHUK, IVAN. Manifestation of the pairing symmetry in the vortex core structure in iron-based superconductors. 2013. Diss.

521. KÄÄRIÄINEN, MARJA-LEENA. Atomic layer deposited titanium and zinc oxides; structure and doping effects on their photoactivity, photocatalytic activity and bioactivity. 2013. Diss.

522. KURONEN, JUHANI. Jatkuvan äänitehojakautuman algoritmi pitkien käytävien äänikenttien mallintamiseen. 2013. Diss.

523. HÄMÄLÄINEN, HENRY. Identification of some additional loss components in high-power low-voltage permanent magnet generators. 2013. Diss.

524. SÄRKKÄ, HEIKKI. Electro-oxidation treatment of pulp and paper mill circulating waters and wastewaters. 2013. Diss.

525. HEIKKINEN, JANI. Virtual technology and haptic interface solutions for design and control of mobile working machines. 2013. Diss.

526. SOININEN, JUHA. Entrepreneurial orientation in small and medium-sized enterprises during economic crisis. 2013. Diss.

527. JÄPPINEN, EERO. The effects of location, feedstock availability, and supply-chain logistics on the greenhouse gas emissions of forest-biomass energy utilization in Finland. 2013. Diss.

528. SÖDERHOLM, KRISTIINA. Licensing model development for small modular reactors (SMRs) – focusing on the Finnish regulatory framework. 2013. Diss.

529. LAISI, MILLA. Deregulation's impact on the railway freight transport sector's future in the Baltic Sea region. 2013. Diss.