Aalto University
School of Science
Degree Programme of Engineering Physics and Mathematics

Jussi Hirvonen

Stochastic approach to mid- and long-term forecasting of ERCOT real-time electricity price

Master's Thesis
Espoo, April 13, 2016

Supervisor: Prof. Pauliina Ilmonen
Instructor: M.Sc. Jyrki Leino

Aalto University
School of Science
Degree Programme of Engineering Physics and Mathematics

ABSTRACT OF
MASTER'S THESIS

Author: Jussi Hirvonen

Title:
Stochastic approach to mid- and long-term forecasting of ERCOT real-time electricity price

Date: April 13, 2016 **Pages:** 111

Professorship: Statistics **Code:** Mat-2

Supervisor: Prof. Pauliina Ilmonen

Instructor: M.Sc. Jyrki Leino

The purpose of this work is to build understanding of real-time (RT) price creation in Electric Reliability Council of Texas (ERCOT) and construct a stochastic simulation methodology to create RT price time series several years to future. Simulation methodology takes into account impact of identified drivers of RT price.

Forecast electricity prices are needed as inputs, for example, for power plant investment decisions. In ERCOT, revenues of power plants consist of selling electricity in three markets (day-ahead (DA), RT, and ancillary services (AS) markets). Traditionally, mainly DA and partly AS market are considered in investment profitability calculations. Flexible power plants can get revenue also from RT market. They can obtain additional profit by benefiting price difference between DA and RT markets. Long term price forecasting is needed to support investment decisions due to long lifetime (15+ years) of power plants.

Fundamental and stochastic models are two main classes of electricity price forecasting models. There are established methods for DA electricity price forecasting and commercial software can do the task. Short term (up to a month), typically deterministic, real-time price forecasts are used by market participants, too. However, such solutions do not exist for long-term (3+ years) RT price forecasting.

In this study, statistical analysis is conducted to identify RT price drivers. Simulation methodology is constructed using bootstrap method with several adjustments. It is seen that RT price spikes can be largely explained by available generation capacity exceeding demand and DA price. On most common RT price level forecast error of surplus capacity, change speed of net-load, DA price and previous behaviour of RT price explain price fluctuations. Future values of the chosen explanatory variables can be simulated by conducting a fundamental DA market simulation using a dedicated commercial software and calibrated model.

Functioning of simulation method is validated by comparing simulated RT price series to historical RT price. Moreover, two case studies are conducted. It is seen that increasing share of wind capacity in ERCOT market increases RT price volatility.

Keywords: Forecasting, bootstrap, electricity price, ERCOT

Language: English

Aalto-yliopisto
Perustieteiden korkeakoulu
Teknillisen fysiikan ja matematiikan tutkinto-ohjelma

DIPLOMITYÖN TIIVISTELMÄ

Tekijä: Jussi Hirvonen

Työn nimi:
Sähkön hinnan ennustaminen stokastisin menetelmin

Päiväys:	13. huhtikuuta 2016	**Sivumäärä:**	111
Professuuri:	Tilastotiede	**Koodi:**	Mat-2
Valvoja:	Prof. Pauliina Ilmonen		
Ohjaaja:	M.Sc. Jyrki Leino		

Tämän työn tavoitteena on mallintaa päivän sisäistä sähkönhintaa Teksasissa sijaitsevalla ERCOT-markkinalla. Lisäksi kehitetään simulointimenetelmä, jolla voidaan luoda hinta-aikasarjoja useaksi vuodeksi tulevaisuuteen. Selittäjämuuttujien vaikutus huomioidaan mallinnuksessa.

Sähkönhinnan ennustaminen on tärkeää esimerkiksi voimalaitosinvestointien kannattavuutta laskettaessa. ERCOT:ssa voimalaitokset myyvät sähköä kolmella markkinalla: seuraavan päivän (DA), päivän sisäisellä (RT) sekä reservituotemarkkinalla (AS). Tavallisesti vain DA-markkinan hinta huomioidaan investointilaskelmissa. Nopeat voimalaitokset voivat kuitenkin hyötyä merkittävästi RT-markkinahinnan liikkeistä ja erityisesti sen ja DA-markkinahinnan erosta.

Sähkönhinnan ennustamiseksi on olemassa kahdenlaisia menetelmiä - fundamentaalisia ja stokastisia. DA-markkinahinnan ennustamiseen käytetään yleisesti tietokoneohjelmia, joiden antamat tulokset ovat varsin tarkkoja niin lyhyellä kuin pitkälläkin ennustehorisontilla. RT-markkinahinnalle ei kuitenkaan ole olemassa vastaavia pitkän aikavälin ratkaisuja.

Tässä työssä tutkitaan tilastollisesti RT-markkinahinnan muodostumista. Löydettyihin selittäjämuuttujiin perustuen rakennetaan bootstrap-menetelmää käyttäen stokastinen ennustemalli. Korkeiden hintapiikkien nähdään useimmiten tapahtuvan korkean DA-markkinahinnan ja alhaisen vapaan sähköntuotantokapasiteetin aikoina. Tavallisimmalla hintatasolla selittäjämuuttujiksi valitaan DA-markkinahinta, vapaan sähköntuotantokapasiteetin ennustevirhe, tuulituotannon ylittävän kulutuksen muutosnopeus sekä edellinen RT-markkinahinta. Selittäjämuuttujien arvot simuloidaan sähkömarkkinamallinnukseen tarkoitetulla tietokoneohjelmalla ja Markovin ketjuihin perustuvalla stokastisella menetelmällä.

Simulointimenetelmän toimivuus varmistetaan ennustamalla menneisyyden hintoja ja havaitsemalla ennusteet tarpeeksi samanlaisiksi toteutuneiden kanssa. RT-markkinahintaa simuloidaan kahdessa tulevaisuuden skenaariossa. Tuulituotannon määrän kasvun nähdään kasvattavan hintavaihteluita RT-markkinalla.

Asiasanat:	Ennustaminen, bootstrap, sähkön hinta
Kieli:	Englanti

Acknowledgements

First and foremost, I would like to thank my instructor Jyrki Leino for his inspiring leadership style, endless expertise, and insightful comments. It has been simply a pleasure to work with him. I will always remember the trip to the U.S. and all other fun moments we had with Jyrki.

Second, I thank my supervisor professor Pauliina Ilmonen for helping me navigate through this project. Her advice was priceless in both qualitative and quantitative challenges I faced during this project.

I thank Emilia, Elina, Sofia, Mikko and all other wonderful people in Sales & Marketing team. This thesis would not exist without Kristian Mäkelä and Matti Rautkivi. I thank Wärtsilä for opportunity to work on this extremely interesting research topic.

I am grateful to my brother Roope and all my friends who have occasionally interrupted my study evenings for other, more important, aspects of life. I thank Tuuli a thousand times for her patience and being always happy to get me home even after my late nights at work.

Finally, I thank my parents Mia and Pekka for supporting me in everything I have decided to do in my life.

Espoo, April 13, 2016

Jussi Hirvonen

Contents

Abbreviations and Acronyms

Ancillary services (AS)	Services other than electricity production provided to market operator by electricity generators for grid stability maintaining purpose
Annual peak load	The highest load of a year
Base load	A constant demand of electricity over a long time period
Base load plant	A power plant that usually does not start and stop within day
Black-out	A situation in which electricity demand of part of or whole market cannot be served
Capacity factor	Mean generation in a period as percentage of generation capacity
Clearing	Action performed by market operator to match demand and supply based on bids and offers provided by electricity buyers and sellers
Combined cycle gas turbine (CCGT)	A gas turbine power plant with a steam turbine that drives an additional generator for improved fuel efficiency
Congested line	A transmission line transmission capacity of which is limiting electric current in it
Congested node	A node that is connected to grid by transmission lines that are often congested
Congestion	A situation in which a current in a transmission line is limited by its transmission capacity
Current operating plan (COP)	A plan that generating company provides to inform ERCOT about its intended power plant availability
Day-ahead (DA) market	Electricity market where sold energy is delivered one day after market clearing
Day-ahead (DA) price	Price of electricity in day-ahead market

Dispatching	Process which optimises generation needed to serve load and provides running schedule to power plants. Conducted by market operator
Energy-only market	Electricity market type, in which generators are not paid for their on-line generation capacity, but only for electricity they sell
ERCOT	Electric Reliability Council of Texas. ISO of Texas.
Fundamental model	An electricity price forecasting model that is based on modelling generation of different power plant types
Gas turbine (GT)	A gas turbine power plant
Gate closure	Dead-line in minutes for bids and offers before delivery of electricity, 5 minutes in ERCOT RT market
Grid	All electric components (generation, transmission, consumers) of a market
Grid stability	Capability of power system to deliver electricity to consumers without large frequency deviations and black-outs
Hub	A region of grid that consists of several nodes
Hub average price	Electricity price that is calculated as simple average of all ERCOT hub prices
Hub price	Electricity price that is calculated as average of LMPs within the hub
ICE	Internal combustion engine
Independent power producer (IPP)	A private company that owns generation capacity and sells electricity to consumers
Independent system operator (ISO)	An organization that acts as market operator and coordinates power system operation, e.g. ERCOT
Intermediate load	Electricity demand that is present for 10 to 18 hours per day due to regular daily seasonal pattern in demand
Intermittent generation	Electricity generation of wind and solar power plants
Load	Total electricity demand of market
Load reduction	AS product in which a voluntary electricity consumer gets compensation for committing to reduce its consumption when needed
Locational marginal price (LMP)	Electricity price of a node

Market operator	An organization that matches load and generation in electricity market
Net-load	The load exceeding intermittent generation
Nodal market	An electricity market type in which a typically large number of nodes have separate prices, e.g. ERCOT
Node	A place in grid where consumption or generation is located
Non-intermittent generation	Electricity generation of all other than wind and solar power plants
Non-spinning reserve	Ancillary service product that requires power plant to be ready to start generating electricity in 30 minutes when needed
Operational flexibility	Ability of a power plant to change its output in short time
Peak load	The load that exceeds intermediate load
Peaking power plant (peaker)	A power plant that only runs during peak load hours
Price-adder	A mechanism that increases electricity price, when surplus capacity in the system is low. Introduced in ERCOT to incentivize building new generation capacity.
Price-setter (marginal power plant)	The power plant with the highest offer price that is dispatched
Price-taker	Power plant that is dispatched, but is not price-setter
Ramp	To change output power of a power plant.
Ramping capability	Ability of a power plant to change its output quickly
Ramping constraint	Inability to ramp up as fast as would be needed. Can be a problem either on the level of a single power plant or entire system.
Real-time (RT) market	Intra-day electricity market where sold energy is delivered typically very shortly (e.g. 5 minutes in ERCOT) after market clearing
Real-time (RT) price	Price of electricity in real-time market
Regulation-down	Ancillary service product to decrease generation when needed
Regulation-up	Ancillary service product to increase generation when needed

Responsive reserve	Ancillary service product that requires power plant to be ready to start generating electricity in a few seconds when needed
Risk premium	Amount by which the expected value of risky return must exceed risk-free return to make a market player indifferent between risky and risk-free return
Surplus capacity	Available generation capacity exceeding load
Three-part supply offer	An offer to sell electricity that specifies (i) start-up offer, (ii) minimum-energy offer, and (iii) energy offer curve
Transmission capacity	Maximum electric current of a transmission line
Utility	Typically very large company that generates and distributes electricity
Zonal market	Electricity market type in which prices are regional, e.g. Nordpool

Notations

RT	Real-time electricity price
DA	Day-ahead electricity price
$\Delta = RT - DA$	Spread between day-ahead and real-time price
$\Delta_{hist.}$	Historical Δ value
IGA	Intermittent generation, actual (MW)
IGF	Day-ahead forecast of intermittent generation (MW)
LA	Demand, actual (MW)
LA	Day-ahead forecast of demand (MW)
$NICA$	Available non-intermittent generation capacity, actual (MW)
$NICF$	Day-ahead forecast of available non-intermittent generation capacity (MW)
FCA	Available generation capacity exceeding demand, actual (MW)
FCF	Day-ahead forecast of available generation capacity exceeding demand (MW)
$FCFE = FCA - FCF$	Forecast error of available generation capacity exceeding demand (MW)
$NLA = LA - IGA$	Net-load, actual(MW)
NLC	Change of net-load from the previous 5-minute period (MW)
C	Dynamic time step category
C_{spike}	Dynamic time step category corresponding to spike
n_C	Number of different dynamic time step categories used in RT price simulation
X_{spike}	Binary variable that indicates spike period, 1: spike, 0: non-spike
$\Delta_C \in \{0, 1, 2, ...\}$	A variable that indicates which group Δ of certain period belongs to

$RT_C \in \{0, 1, 2, ...\}$	A variable that indicates which group RT of certain period belongs to
n_{RT_C}	Number of price classes used in RT price sampling
n_{Δ_C}	Number of delta classes used in RT price sampling
l_{RT_C}	A vector that contains limits of price classes
l_{Δ_C}	A vector that contains limits of delta classes
$DA_{limits,C}$	A vector that contains limits of DA price for determining dynamic time step category C
$FCFE_{limit}$	A limit of $FCFE$ for determining dynamic time step category C
NLC_{limit}	A limit of NLC for determining dynamic time step category C
$l_{spike,RT}$	A limit above which all RT prices are spike prices
$l_{spike,DA}$	A limit above which all DA prices are spike prices
$DA_{limit,prob}$	A limit of DA price for conditional probability of spike period starting
$DA_{limit,prob}$	A limit of DA price for conditional probability of spike period starting
$FCA_{limit,prob}$	A limit of FCA for conditional probability of spike period starting
G_{RT_c}	Group of historical RT prices that have occurred when RT class was $RT_C, C = 0, 1, 2, ...n$
G_{Δ_c}	Group of historical Δ values that have occurred Δ class was $\Delta_C, C = 0, 1, 2, ...n$
G_L	A set that includes lengths of historical spike periods
$G_{RT_{spike}}$	A vector that contains RT prices of historical spike periods in chronological order
S_{spike}	Minimum share of spike prices within a spike period
X_{start}	Binary variable, 1: spike period starts, 0: spike period does not start
L_B	Block length used in simulation of future spike prices with block bootstrap method
XF	Day-ahead forecast value of explanatory variable X
XA	Actual value of explanatory variable X
$XF_{fut.}$	Future day-ahead forecast value of explanatory variable X
$XA_{fut.}$	Future actual value of explanatory variable X
$XF_{hist.}$	Historical day-ahead forecast value of explanatory variable X
$XA_{hist.}$	Historical actual value of explanatory variable X
$IGF_{fut.}$	Future day-ahead forecast of intermittent generation

$IGA_{fut.}$	Future actual value of intermittent generation
$IGF_{hist.}$	Historical day-ahead forecast value of intermittent generation
$IGA_{hist.}$	Historical actual value of intermittent generation
$LF_{fut.}$	Future day-ahead load forecast
$LA_{fut.}$	Future actual load
$LF_{hist.}$	Historical day-ahead load forecast
$LA_{hist.}$	Historical actual load
$NICF_{fut.}$	Future day-ahead forecast of available non-intermittent generation capacity
$NICA_{fut.}$	Future actual value of available non-intermittent generation capacity
$NICF_{hist.}$	Historical day-ahead forecast of available non-intermittent generation capacity
$NICA_{hist.}$	Historical actual value of available non-intermittent generation capacity
XF_C	Class of day-ahead forecast of explanatory variable X
n_{XF_C}	Number of classes XF_C of day-ahead forecast used in simulation of actual value of explanatory variable X
l_{XF_C}	Limits of classes XF_C of day-ahead forecast used in simulation of actual value of explanatory variable X
$l_{XA_{XF_C}}$	Limits of conditional quantiles Q_{XA} used in simulation of actual values of explanatory variable per each class of day-ahead forecast XF_C
Q_{XA}	Conditional quantile of actual value of explanatory variable X
q	Number of conditional quantiles Q_{XA} used in simulation of actual values of explanatory variable
$I_a(x)$	Indicator function

$$I_a(x) = \begin{cases} 1, & \text{if } x \geq a \\ 0, & \text{otherwise} \end{cases}$$

Chapter 1

Introduction

Forecasting electricity price in competitive markets is a task studied a lot by statisticians. However, most scientific work on forecasting electricity prices until today has been about day-ahead (DA) electricity price, i.e. the price of electricity delivered one day after the market clearing. Flexible power plant technology allows benefiting movements of Real-Time (RT) price, i.e. the price of electricity to be delivered in only a few minutes after market clearing. Thanks to their ability to start and stop flexibly for just a few minutes without starting cost or impact on maintenance, those power plants can take advantage of all price movements. In this work we study fundamentals of RT price creation in Electric Reliability Council of Texas (ERCOT). We also build a method to create simulated RT price forecasts for several years. The method is based on a statistical method known as bootstrapping with several modifications to capture the impact of several explanatory variables.

1.1 Background

Power plant investments are typically long-term investments. A major driver of profitability is price for which generated electricity can be sold. Future price movements, of course, are not known with certainty at the time that investment decision is made. Traditionally only DA market and ancillary services (AS) market, in which different reserve products are traded, have been considered in investment decisions (in personal communication with M.Sc. Jyrki Leino, Senior Power System Analyst, Wärtsilä, 15 March 2016, Turku). However flexible power plants can achieve profits by selling some or all of their generation in RT market. A flexible power plant can also get additional profit by benefiting price differences between markets. For example, it can sell electricity in DA market and, instead of generating electricity to fulfil

its commitment, buy the commitment back from RT market and not start at all in RT. This strategy is profitable if RT price is lower than marginal generation cost of the plant. Wärtsilä is one manufacturer of flexible power plants and has a need to understand the value of power plant flexibility in ERCOT. Goal of this study is to find the drivers of electricity price in ERCOT RT market and create simulated RT price time series for several years. Forecast price series will be used to determine profitability of power plant investments in ERCOT.

ERCOT market was chosen for the study since electricity prices there have been volatile in recent years. Price volatility has given opportunities for flexible power plants to make profits by making transactions in different markets as described above. Secondly, there is a need to understand what impact some changes, such as growing wind generation capacity, will have on price volatility in ERCOT. RT price volatility is mainly a result of large share of residential consumers, the demand of which is often inflexible with respect to the price due to low consumption and limited chances to observe electricity price. Also, the generation in ERCOT market is more difficult to predict than in many other markets, because of the large share of renewable generation capacity. Renewable generation is exposed to weather changes, which leads occasionally to unexpected events with impact on RT price. Third reason behind volatile RT price is low transmission capacity of the interconnectors that enable transferring electricity from or to adjacent markets. Therefore, ERCOT is almost an electric island. Unexpected electricity grid or generation outages cannot be compensated by co-operating with other markets and may lead to high prices.

1.2 Research problem

The purpose of this thesis is to understand RT electricity price creation in ERCOT and build a methodology to create forecast price time series for several years. Due to significant uncertainty related to future events we do not try to create one precise RT price prediction. Instead, we simulate hundreds of time series and study the distribution of price among all forecast series. We put a special emphasis on studying the impact of increasing wind generation capacity on price volatility. The needed steps are (i) building hypotheses and conducting statistical analysis to identify explanatory variables of RT price movements, (ii) determining quantitative measures of correlation between RT price and each explanatory variable, (iii) constructing a stochastic price creation method based on correlations and forecast time series of explanatory variables, (iv) creating simulated price scenarios using predictions

of explanatory variables provided by ERCOT, and (v) performing "what-if" analyses to see what impact e.g. demand and wind generation capacity growing slightly faster than expected could be expected to have on future prices.

1.3 Methods

We conduct different statistical analyses to identify drivers of electricity price in ERCOT RT market. The methods include univariate (e.g. duration curves and empirical autocorrelation functions) and bivariate (e.g. scatter plots) analyses. We use price data from years 2011-2015 and a lot of other market data published by ERCOT and collected and delivered to us by Genscape.. Data selection is based on topic understanding achieved through reviewing literature and interviewing electricity market experts.

To create future price scenarios, we use bootstrap method that was introduced by Bradley Efron [7] and has since been used in many contexts, such as machine learning [2], volatility prediction in stock market modelling [28], and sex determination [26]. Bootstrap is a simple method to apply when reliable history data is available. The principle is easy to understand intuitively even though the method has rigorous mathematical foundation. It also does not need too many assumptions about the unknown underlying distribution that generated the observations. For more information about bootstrapping see e.g. [8] or [5].

We use bootstrap method to generate a large number of scenarios for electricity price in ERCOT RT market for several years. The historical observations of RT price and spread between DA and RT price are used as bootstrapping populations. The impact of different explanatory variables determined by the statistical analyses is taken into account when constructing price forecasts for each future time step. Our approach differs in many ways from fundamental electricity price market modelling methods often used to create short-term price forecasts. Instead of building one prediction of future price we create several possible scenarios. This approach enables us to more robustly determine price distribution than in single point estimation by fundamental modelling and explicitly take into account the uncertainty related to future. A large number of price series can also be used in price risk evaluation of power plant investment. Statistical approach was chosen because of future uncertainty and complicated nature of RT price creation involving many drivers. Some of the drivers may be unknown to us and some of them are virtually impossible to observe. Our forecast series are not accurate predictions based on explanatory variable values. But neither would be

any prediction based on any explanatory variables, as there is always uncertainty about future in electricity markets. However, the long forecast horizon tends to cancel out forecast errors made in estimating prices of individual future time-steps. Therefore, the estimate of overall price volatility can be expected to be far more accurate than estimates of RT prices of individual future time steps.

1.4 Structure of the study

We start the study by going briefly through in Chapter 2 the functioning of modern electricity markets, where price of electricity is determined through auctions arranged every few minutes. We define many technical terms that are needed to understand the remaining parts of the work and specifically take a look at the structure and most important market processes of ERCOT. Next, we briefly review literature on electricity price forecasting. In Chapter 3 we conduct statistical analyses to identify drivers of ERCOT RT price. In Chapter 4 we build a stochastic model to create simulated RT price series. We apply the method to two future scenarios and compare simulated RT price series in the scenarios. We see how increasing share of wind generation capacity in ERCOT market impacts RT price volatility. Finally, in Chapter 5 we present our conclusions from the study and propose some ideas for future research.

Chapter 2

Electricity markets

Many electricity markets have liberalised from the 1990s' regulated fixed-price contract markets to modern competitive markets where electricity price is determined in auctions organised for example every five minutes [22]. Electricity sellers and buyers send their offers and bids to a computer system held by *market operator*. Then market operator *clears* the market by matching demand and supply and *dispatches* power plants, i.e. schedules power plants to run so that generation and demand equal every second.

To incentivize building new capacity, some markets have deployed so called capacity market. In such markets, generators receive payments for each megawatt of their on-line generation capacity. Energy-only markets do not include such capacity payments and all revenues for generators come from selling energy and AS products.

A *zonal market* is divided to several zones, each of which has its own price for electricity. It is possible that price in one zone is consistently higher than in an other zone. This may happen as a result of *congestion*, i.e. limited inter-zonal transmission capacity between the two zones. Price difference may encourage investors seeking greatest returns on their investments to build new generation capacity in the high-price zone. Added generation capacity may lead to prices in zones to converge. One example of zonal market is Nordpool [1]. Sometimes significant congestion can occur also within a zone. In zonal market setting intra-zone congestion does not encourage building more capacity in the region with greatest scarcity. Therefore, *nodal market* structure has been introduced in many markets. In such setting there is individual price for each *node*, i.e. connection point for load or generation, of the grid. In this work we study one such market, ERCOT.

2.1 Characteristics of electricity

With today's technology, electricity can not be stored for long times in large quantities in a cost-efficient way [31], [33]. Therefore, consumption and generation must be matched every second. This work is done by market operator by dispatching power plants as needed to serve the demand and, if adjusting is needed, by deploying regulating reserves. Reserves are generation capacities that are able to respond to required quick output changes. Generators receive monetary compensation for providing this adjustment service to the system. For a comprehensive presentation of matching electricity demand and generation, see [22] and [25].

2.1.1 Demand

Electricity is consumed by households, firms and public sector in different ways. Aggregate electricity demand exhibits several seasonal patterns that differ by country [22]. In cold countries, such as Finland, demand in winter time is generally higher than in summer time because of the greater need for electricity for lighting and warming houses. In warm regions the demand is higher in summer due to the greater need for air conditioning. Intra-day and weekly patterns are highly similar in many countries. The demand is low during people's low-activity time, i.e. at night and at weekend. The demand is often highest in the working days' afternoons when public traffic consumes a lot of energy and appliances such as washing machines and computers are switched on in many homes. However, there are differences in electricity consumption patterns between consumer groups. Industrial customers have less clear seasonal patterns than residential customers. Therefore, the difference between the highest and lowest demand of a day is often greater in countries with lower share of industrial electricity consumption. Demand is often called *load*, when discussed from generation and transmission system point of view.

2.1.2 Generation

In the traditional electricity markets, there were only a few generating companies that sold electricity to all consumers [22]. The number of companies selling electricity has increased rapidly with market liberalisation. For example ERCOT has 198 certified competitive retail electric providers [12]. Large public companies, *utilities*, may have a fleet of dozens of power plants. Some firms, *independent power producers* (IPP), may operate only a single plant.

Table 2.1: Some characteristics of most common power plant types in ERCOT.

Type	Investment (USD/kW)	Base-load, intermediate-load, peaker, intermittent	Variable operational cost (USD/MWh)	Typical plant size (MW)
Coal	1000	base	25	2000
Nuclear	4000	base	7	3000
CCGT	700	base/intermediate	50	1000
OCGT	500	peaker	70	100
Engine	600	intermediate/peaker	60	100
Wind	1000	intermittent	5	100
Solar	3000	intermittent	6	10

All types of power plants have their own role in the well-functioning power system. So called *base load plants* can generate electricity at very low cost and their output is kept stable regardless of load. Nuclear and coal power plants, as well as combined cycle gas turbine plants (CCGT) are often base load plants. Base load plants have often very high generation capacity and require high construction investment. Intermediate plants operate only when load exceeds the total capacity of the base load plants. Their cost per start and required initial investment are often high. Intermediate plants can not generate electricity at as low a cost as base load plants. Therefore, the electricity price is generally higher, when intermediate plants are needed to serve the load. For example, CCGTs may be operated as intermediate plants. Peaking power plants run only during high-demand periods. They have lower cost per start than base load plants and intermediate plants and they are capable of adjusting their output power very fast. Initial investment is often low compared to base load and intermediate plants, but they have higher marginal generation cost resulting from poor fuel-efficiency. Suitable peaker technologies include gas turbines (GT) and internal combustion engines (ICE).

Power plants that use renewable energy sources can generate electricity at very low cost since their fuel is free. In many markets their generation is always used when it is available [24]. However, generation of wind and solar power plants is intermittent by nature as their electricity production stops immediately, when the sun stops shining or there is no more wind. Generation of wind and solar power plants is therefore called *intermittent generation*. In many markets renewable technologies are subsidised by political decision, because they would not be economic choice otherwise. Table 2.1 shows some characteristics of different power plant types discussed above [22], [21], [19]. Numbers are only directional.

The share of load that can not be served by intermittent plants is called *net-load*. It can change in two ways, (i) when load changes or (ii) generation of intermittent plants changes. Net-load changes require ramping from non-

intermittent plants, which requires some degree of operational flexibility from them. Current trend of adding intermittent generation in many markets will make net-load changes faster and larger. Therefore, need for flexible generation capacity will increase in future. Figure 2.1 illustrates load, net-load and intermittent generation of one example day in ERCOT. Graph shows load increasing from 25 GW to almost 45 GW during the day. Simultaneous decrease in wind generation from 12 GW to less than 4 GW makes net-load growth even faster than load growth. This kind of rapid net-load growth requires non-intermittent power plants to ramp up fast.

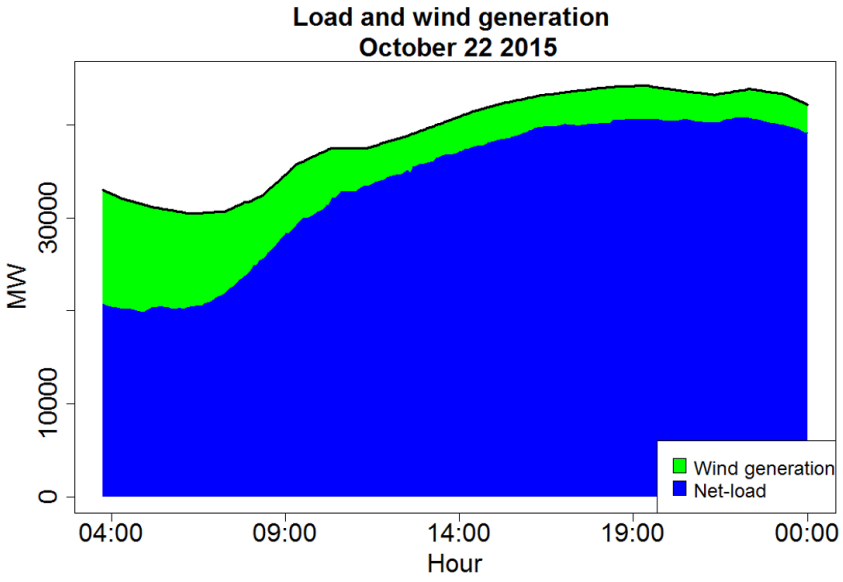

Figure 2.1: Load, wind generation, and net-load of October 22 2015 in ERCOT. Load is the black line on top of the graph. Wind generation is the green region on top of the graph. Net-load is the blue region below wind generation. Load is the sum of net-load and wind generation. Load increases until it reaches its maximum of approximately 44000 MW. Wind generation is strong in the night, but decreases at the same time that load increases making the net-load change even more rapidly than load. Steep increase of net-load is seen as rapid widening of the blue region at night and in the morning.

2.1.3 Transmission

Electricity is transmitted to the place of consumption as electric current via transmission lines. However, there is upper limit for current called *trans-*

mission capacity of transmission line. All transmission lines and *nodes* of a market together are called *grid*. A node is a place where load or generation is electrically connected to a transmission line. A transmission line which is transmitting at its full transmission capacity, is called *congested line*, and the situation when one or more transmission line of a grid is congested is called *congestion*. A *congested node* is a node that is connected to a transmission line that is often congested. It is a task of market operator to ensure that the price of electricity for the users in each node is the cheapest possible. If a node is connected to a low-cost generation unit by transmission lines with abundant capacity, the price in that node will probably be often cheaper than price in an other node that is very congested and has only high-cost generating units near it. The prices of two nodes can differ, if and only if, all transmission lines connecting these nodes are congested or not functioning at all due to line outages. Therefore, one way to reduce market price of electricity is reducing congestion by increasing transmission capacity of transmission lines that are most often congested. It would be technically possible to remove all congestion by adding transmission lines with higher transmission capacity, but it might require high investments. As a result the total cost of electricity to the users might be even higher than in situation where occasional differences in prices of congested nodes occur.

2.2 Different markets

In competitive markets, such as ERCOT, electricity price is determined by demand and supply, and is free to fluctuate as the two change. Electricity is sold in two main markets: DA market and RT market. In addition there is AS market, where market operator buys several reserve products from generators. Reserves are needed to ensure that unexpected events do not risk grid stability, leading at worst case to black-outs.

2.2.1 Day-ahead market

In DA market, a daily auction is arranged, where electricity buyers and sellers provide their bids. The clearing of market is executed by market operator one day before actual delivery of electricity. Therefore, by definition, DA market commitments are future contracts, where the underlying is certain amount of electricity and delivery time is specified to a certain time of next day. In many markets, such as ERCOT, DA market is hourly. Hourly DA market means that market participants can give separate bids for each hour of the next day [20].

2.2.2 Real-time market

Even the best predictions of electricity demand made for next day are just estimates and many things can change before the delivery time of DA contracts. Matching demand and generation exactly for next day would be virtually impossible and often one or both of the two would differ. Difference between generation and demand might lead occasionally to a situation where a part of the regions load would have to be shed to avoid black-out in the whole grid. For that reason there is RT electricity market, which is sometimes called intra-day market. In RT market electricity auctions are arranged very often, even every five minutes. Delivery of electricity happens e.g. five minutes after *gate closure*, which is a dead-line for sending bids and offers. Dispatching is executed by market operator in a similar way to DA market. All plants can participate in the market, but slow-ramping plants can only be dispatched respecting their ramping limits.

A plant that has sold electricity in DA market can sometimes use RT price movements to its advantage. Flexibility is required to make operative decisions only a few minutes before action is needed. Often power plants simply fulfil their DA commitments by generating the electricity that they have sold. Alternatively they can buy electricity in RT market and deliver it to the grid (not physically), so fulfilling their commitment without starting at all. This is more profitable strategy, if buying the electricity from RT market is cheaper for the power plant than generating it themselves. Such situation occurs, when RT price is lower than the marginal generation cost of the power plant.

2.2.3 Ancillary services market

Unexpected generator or transmission outages can occur after RT-market clearing or load may be slightly different than was expected. If there would not be any way to compensate for these changes in supply and demand, there would be a good chance that the generation and load would not equal. It would be seen in the frequency of alternating current, and might lead to *black-outs* in extreme cases. AS products are used if system stability is in danger. Market operator buys these products from generators as commitments to increase or decrease their output when needed. *Regulation up* capacity sold by a power plant means that it must increase its output by certain amount, whenever the frequency of alternating current falls below a certain threshold. *Regulation down* is a commitment of a power plant to decrease its output by certain amount whenever the frequency exceeds a certain threshold determined by market operator. Regulation products can be

provided only by power plants that can respond to changes in frequency very quickly, only in a few seconds. There are also reserves that can be deployed by market operator in a situation where a power plant trips, or for another reason generation suddenly drops significantly. One product that can be used to avoid load shedding is *load reduction*. It is a commitment of a voluntary electricity consumer, e.g. a large industrial site, to reduce its consumption when needed [18]. Some power plants, such as ICEs, can get a large share of their revenue from selling AS products on top of their energy sales, by e.g. selling their entire output in the DA market and regulation-down in AS market. Some power plant types, e.g. nuclear, hardly ever participate in AS market. For a thorough discussion of frequency response, see [21]. The purpose and functioning of different reserves are discussed in [25]. In this work we do not study AS market or prices.

2.3 Forecasting electricity prices

Ability to forecast electricity prices is important for many parties. Power plant owners do investment decisions based on long-term forecasts and operative decisions based on short-term forecasts of future prices. For example, building new flexible generation capacity requires enough price volatility to be profitable. Regulators and politicians try to create market rules so that they encourage building sufficient generation capacities of different types so that load can be served reliably and cost-effectively.

Different methods to create price forecasts exist. Selection of methodology depends on intended purpose of forecast future price time series. Two main approaches to price forecasting are *deterministic* and *stochastic models*. Deterministic models are used to create one prediction of time series, some kind of maximum likelihood estimate. Deterministic models are often used for short-term (up to a month) forecasting. One example of use of deterministic price modelling is creation of power plant bidding strategy for DA market by a trader working for a utility. Stochastic models are based on simulating many forecast price time series, and studying the distribution of simulated time series. For example, 95 % confidence intervals of electricity price of next years Christmas could be created using stochastic modelling.

There are established methods for DA price forecasting, and commercial software can do the task. Short-term RT price forecasts are used by many market participants, too. However, we do not know any such solutions for long-term RT price forecasting. In this work we build a RT price modelling method that uses simulated DA price time series and several other input variables. A study that introduces a forecasting method for AS price is done

in [16].

2.3.1 Day-ahead price forecasting

DA price creation is well-known on fundamental level and there are several methods to create forecasts. Commercial software products exist for short-term and long-term forecasting [16], [17], [3]. electricity market participants can use them supporting their decision making [32].

For more information on DA price forecasting models see e.g. [34], [32], [4], or [23].

In this work we do not focus on DA price forecasting, but we use simulated DA price series as input for our RT price forecasting method. It is possible to forecast DA price accurately enough to be used as an input for stochastic RT price forecasting model.

2.3.2 Real-time price forecasting

Fundamental models for short-term forecasting of RT price exist and are used by market participants. However, methodologies to forecast RT price reliably many years to future do not exist (Rhodri Williams, Regional director - ERCOT, Genscape, Boston, 10 February 2016). Generally RT price forecasting is more difficult than forecasting DA price, as RT price is more volatile (short but very high spikes) and impacted by many more variables. In this work we build a stochastic method to create simulated RT price forecasts for several future years.

2.4 ERCOT

2.4.1 Overwiev

ERCOT market covers about 90 % of Texas load and serves 24 million consumers [12]. Installed generation capacity of ERCOT is more than 74 GW and current record demand, 69.6 GW, occurred on August 10 2015. ERCOT is a nodal energy-only market, which means that generating companies receive all their revenues from selling energy and AS products, not from their on-line capacity. There is a separate RT price, *locational marginal price* (LMP), for each node. LMPs differ from each other when congestion occurs [11]. So called *hub prices* are average prices of node LMPs. We will forecast *hub average price* in this study, which is the simple average of all ERCOT hub prices.

As mentioned earlier, ERCOT is almost an electric island due to limited capacity of interconnections with adjacent markets. This transmission capacity is only 1.3 GW [9].

Figure 2.2 shows shares of installed generation capacity in ERCOT in 2014 by fuel type. 55 % of generation capacity is natural gas -fired. Coal power plants account for 24 % and nuclear plants for 6 %. Wind accounts for 14 % of installed capacity and wind generation capacity continues to grow in future years [9]. Solar capacity is less than 1 % of installed capacity.

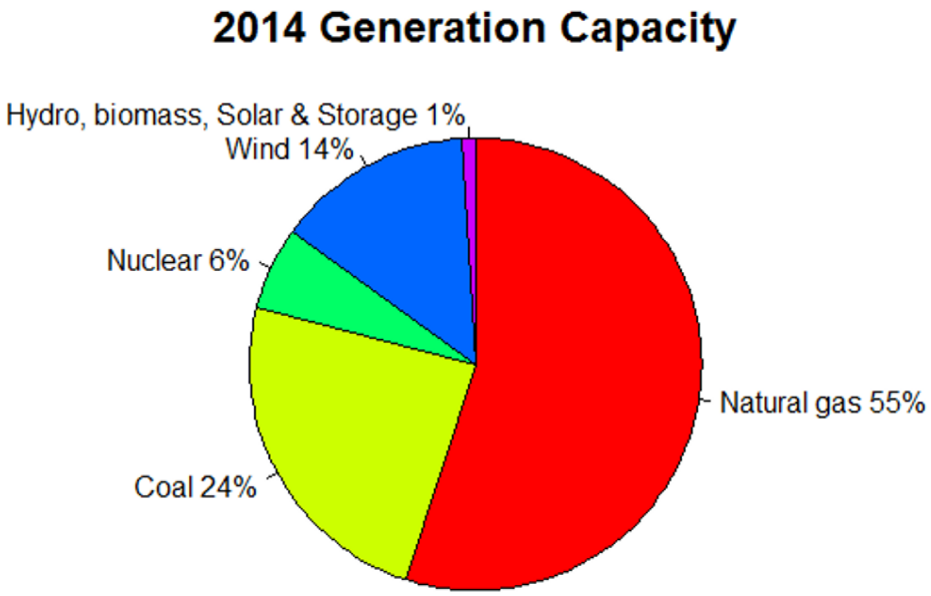

Figure 2.2: Shares of generation capacities by fuel in 2014 in ERCOT [12].

Therefore, we do not treat it separately in this study. However, the trend is increasing in ERCOT [9], which means that impact of solar generation may need to be added to price forecasting models in future. Since solar generation has very low marginal production cost, but is intermittent by nature, it can be supposed that addition of solar generation will lower prices on sunny days. On the other hand, solar generation can be supposed to increase price volatility by increasing magnitude of net-load changes and need to dispatch fast-ramping power plants.

2.4.2 Different markets of ERCOT

In ERCOT, there are two energy-only markets: DA energy market and RT energy market. For ancillary services (AS) there is only DA market. An overview to market procedures of each market is provided in [11]. A thorough description of all market rules and processes can be found in ERCOT market rules and protocols, e.g. [14] and [10].

2.4.2.1 ERCOT Day-ahead energy market

DA market clearing is performed every morning at 10 a.m. for electricity delivery of the next day. DA market is hourly. Participating in DA market is voluntary for generators. The generators provide for each hour their *three-part supply offers*, specifying their start-up offer (USD they want for starting to generate), minimum-energy offer (minimum stable load and respective price), and energy offer curve (price-quantity pairs above minimum stable load). Market is cleared by ERCOT as least-cost solution. This means that ERCOT optimises the system based on all offers from generators so that the generation equal to load will be carried out with least cost possible to the electricity consumers [11].

2.4.2.2 ERCOT Real-time energy market

RT market clearing is made every five minutes. Gate closure is five minutes, which means that actual electricity delivery happens five minutes after market clearing. Generators provide their three-part supply offers, as they do in DA market. Market clearing is done as least-cost solution [11]. Volume of RT market in MW is lower than that of DA market. However, prices in DA market and longer-term forward markets that are bigger in volume closely follow RT price, which makes financial impact of RT market bigger than its MW-volume suggests [29].

2.4.2.3 ERCOT Day-ahead ancillary services market

In ERCOT, there are four AS products: regulation up, regulation down, responsive reserve, and non-spinning reserve. Like DA energy market, AS market is cleared at 10 a.m. for the next day. AS market is hourly [15].

Chapter 3

Statistical analysis of RT price

3.1 Analysis of RT electricity price and its potential drivers

We will next study historical market data published by ERCOT and collected and delivered to us by Genscape. The goal is (i) to build quantitative understanding of drivers of RT price and (ii) to choose an appropriate simulation methodology to create future time series of RT price.

3.1.1 DA price

As explained in chapter 2, DA price DA is determined by market clearing executed by ERCOT on basis of bids and offers sent by market participants. Bidding behaviour of market participants is largely impacted by expectations of probabilities of specific conditions occurring on the next day (in personal communication with Kevin Hanson, Supervisor, Market Operations Support, ERCOT, 8 February 2016, Taylor TX).

Figure 3.1 shows DA price from the beginning of year 2011 to the end of year 2015. Price is most of the time less than 100 USD/MWh. Median price is 27.44 USD/MWh. The highest prices are in summer of 2011, around 2500 USD/MWh. Periods when price is above 500 USD/MWh are rare. Such periods have occurred less than ten times in years 2011-2015. Even though there are only a few high spikes in DA price, they have a significant financial impact on electricity sellers and buyers. Duration curve is a graph commonly used to illustrate empirical distribution of electricity price. Vertical axis represents prices from the smallest to highest in the period of interest. Horizontal axis represents the hours of year (8760 in total). It can be scaled if the inspection period length differs from one year. Then it

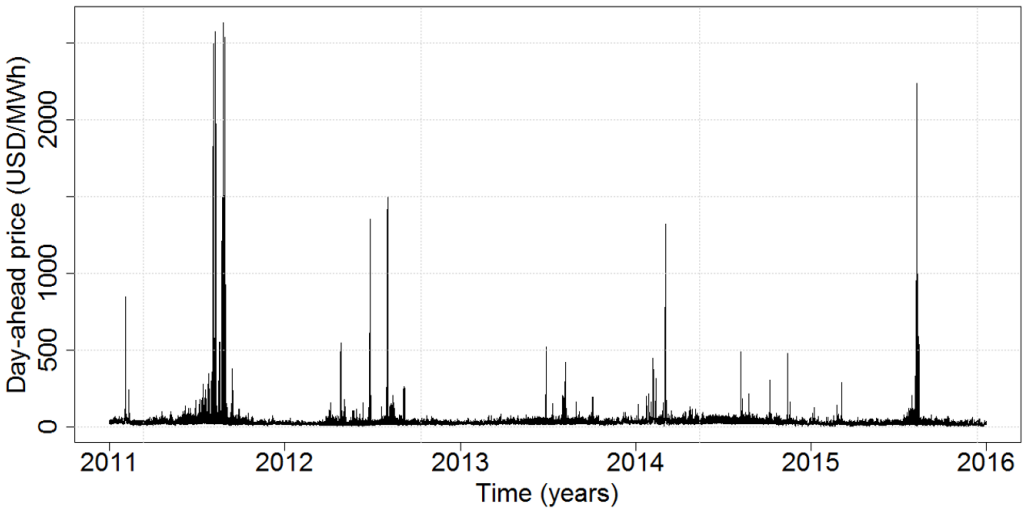

Figure 3.1: Day-ahead price from year 2011 to 2015. Most of the time DA price is below 100 USD/MWh, but occasional spikes of more than 100 USD/MWh occur.

represents average hours of year in the inspection period. Duration curve shows how many hours in a year (value on horizontal axis) electricity price was above any price level (value on vertical axis). Duration curve is always decreasing. Exactly same information could be shown in empirical cumulative distribution function or histogram (though infinite number of intervals would be needed in histogram in general case). Duration curve of DA price in figure 3.2 shows that highest DA price in 2011-2015 has been around 2600 USD/MWh and lowest has been approximately zero. DA price has exceeded 100 USD/MWh only a few hundred hours per year. Figure 3.3 shows DA price and demand of an example week in June 2014. Both variables show clear seasonal variation at one day season length. Values of both variables are often low at night and high in the day-time. Demand of last two days of the inspection period is higher than in the other days. DA price has reached its maximum in these days, too. DA price is largely impacted by demand since price is determined by the dispatched power plant with the most expensive offer price. In low-demand times it is enough to dispatch renewable and base load plants with cheap offer prices. As demand increases intermediate and peaking plants with higher offer prices need to be dispatched to serve the demand. Consequently price increases.

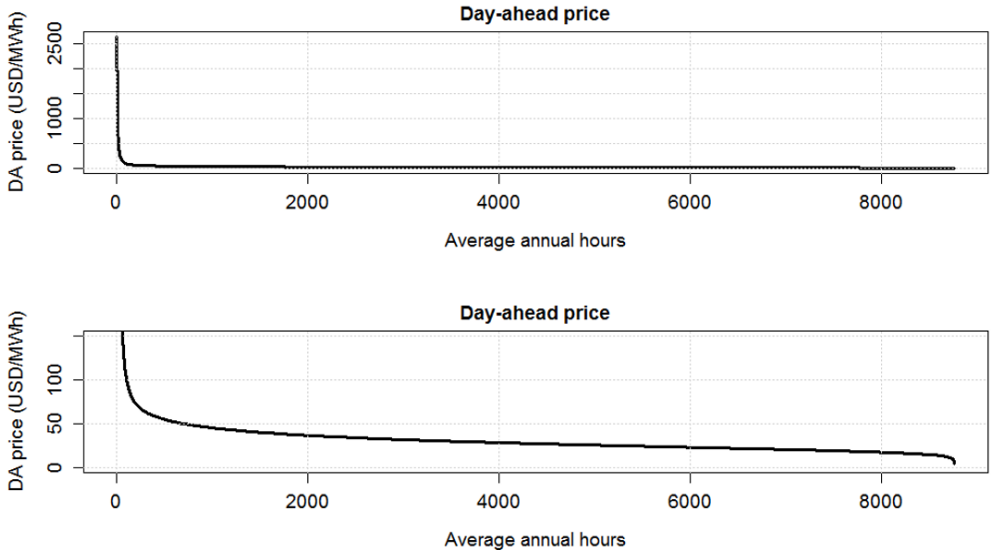

Figure 3.2: Duration curve of DA price in top figure shows that only a tiny share of DA prices was above 500 USD/MWh. Bottom figure shows part of duration curve where $DA < 150$. Price is below 50 USD/MWh approximately 8000 hours per year. Almost linearly decreasing duration curve shows that price distribution is rather uniform between 0 and 50 USD/MWh.

3.1.2 RT price

ERCOT has specified a price cap, which limits how high the RT price RT can be. Table 3.1 shows historical levels of price cap as specified in [6]. Figure 3.4 shows RT price from the beginning of year 2011 to the end of year 2015. Price is most of the time less than 100 USD/MWh. Median RT price is 25.2 USD/MWh. Cap prices have occurred unregularly in the five-year period and comparison with figure 3.1 shows that high RT prices are more frequent than high DA prices. Because of greater height and frequency, the financial impact of price spikes on market participants is even larger in RT than DA market. However, as the DA price is hourly and RT price is 5-minute price, the duration of RT spikes can be, and often is, a lot shorter than in DA. A histogram of historical durations of price spikes is shown in figure 3.5. Majority of spikes lasts only 5 minutes and only a tiny share of spikes is longer than 3 hours (36 5-minute periods). A power plant that can capture the value of short price spikes by being able to ramp up and down very fast, can make significant profit in RT market in addition to its DA market profit [30]. Majority of the highest RT prices are in the year 2011, on cap level 3000 USD/MWh. The reason for high prices in the summer

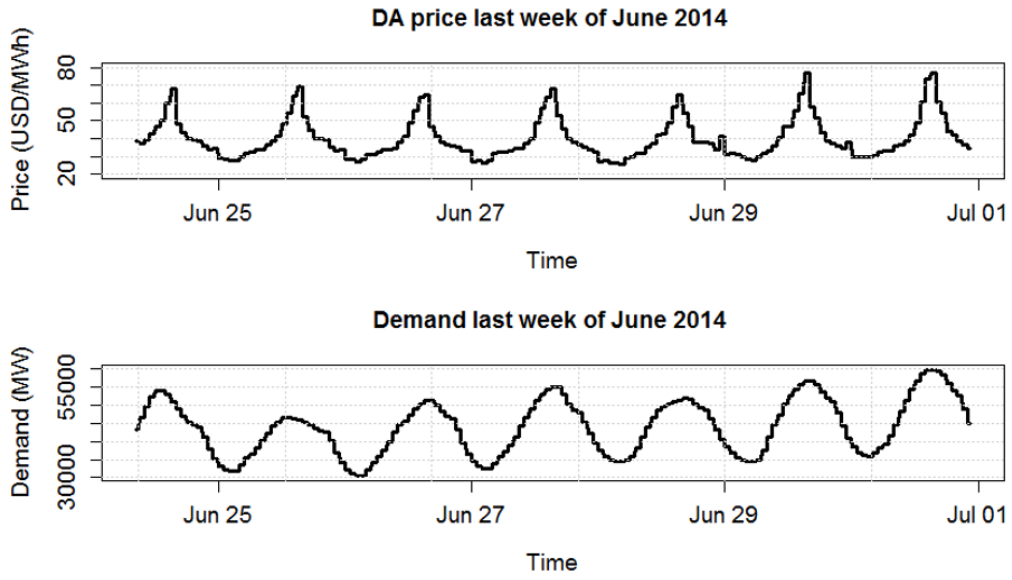

Figure 3.3: Top graph shows DA price of an example week in June 2014 and bottom graph shows demand in the same period. Both time series have a clear seasonal variation at one day season length. Price is often high, when demand is high, and opposite.

of 2011 was very hot weather in Texas leading to very high demand and to generation outages [27]. The price spikes in February 2011 are a result of generation outages and problems with gas equipment caused by very cold weather [27]. In RT price not only prices above 1000 USD/MWh, but also prices between 500 and 1000 USD/MWh are more frequent than in case of DA price. These prices are mainly due to quick net-load changes that require GTs to be started, since there is not enough cheaper fast-ramping power plant generation capacity available (in personal communication with Kevin Hanson, Supervisor, Market Operations Support, ERCOT, 8 February 2016, Taylor TX). GTs have high cost per start and the need to dispatch them increases the price to a high level for at least the first 5-minute period that they are running.

Duration curve of *RT* in figure 3.6 shows that highest price in 2011-2015 has been approximately 5000 USD/MWh and lowest has been negative, approximately -250 USD/MWh, where also price floor is located. *RT* has exceeded 100 USD/MWh only a few hours per year. Excluding both ends of duration curve, the rest is very stable. This means that more than 8500 hours a year, price has been inside a small interval. As spikes are so rare, we can model them separately in the forecasting method.

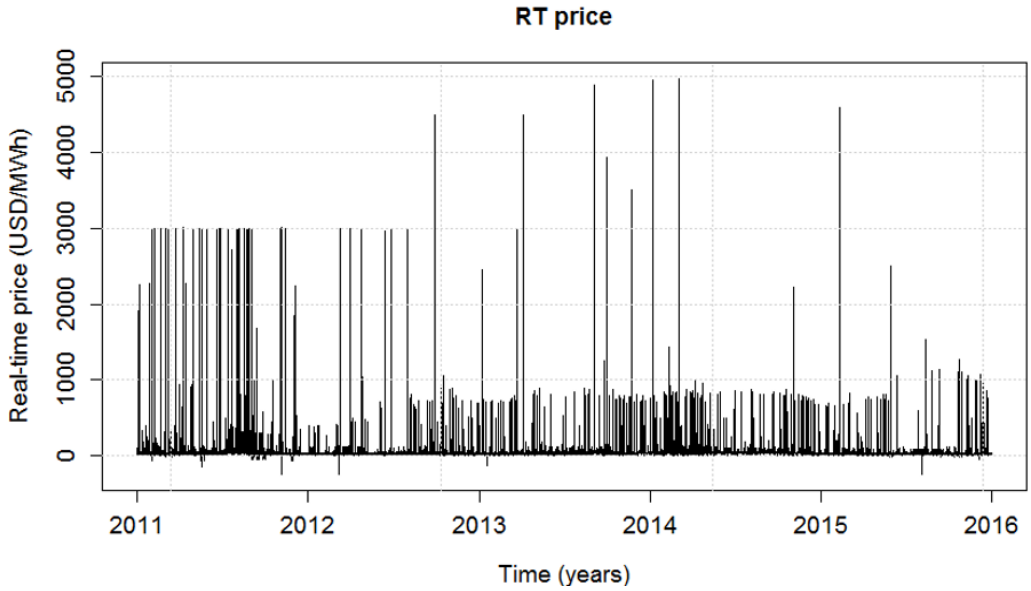

Figure 3.4: Real-time price from year 2011 to 2015. Most of the time price is low, but there are sometimes spikes. Year 2011 shows most huge price spikes of 3000 USD/MWh. Prices between 500 and 1000 USD/MWh are most frequent in the period starting at second half of 2012.

RT is strongly positively autocorrelated on short lags. This can be seen in figure 3.7, showing the empirical autocorrelation function of *RT*. At lag 100, i.e. 8 hours, autocorrelation has almost vanished already, but there is some positive 24-hour seasonal autocorrelation, around lag 288. This is due to daily variation that was seen even more clearly in *DA* and is also present in *RT*. To account for the strong autocorrelation of RT price, we can use the price of previous time step to explain the next price.

Figure 3.8 shows *RT* and *DA* of an example week in June 2014. *DA* moves seasonally up and down by time of day and *RT* follows it to some

Table 3.1: ERCOT price cap 2011-2015

Start date	Price cap
N/A	3000
1 Aug 2012	4500
1 Jun 2013	5000
1 Jun 2014	7000
1 Jun 2015	9000

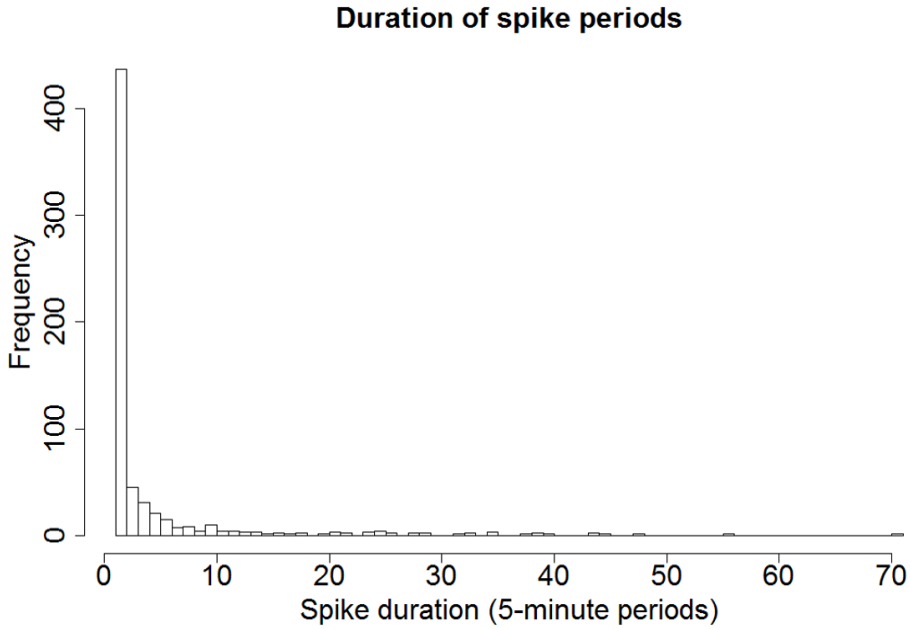

Figure 3.5: Histogram of durations of historical price spikes. More than 400 of total 600 spikes lasts only 5 minutes (one time step). The longest price spike in 2011-2015 has been almost 6 hours (70 5-minute periods) long.

extent. However, RT movements are not as regular as those of DA. highest RT price is often lower than highest DA price of the day. This systematic difference between RT and DA can be seen as a risk premium. It represents the expected profit loss that risk-averse electricity buyers are willing to accept in DA market to avoid having to buy their electricity in RT market where price volatility is higher and consequently risk of price spike is higher. Load is low at night and unexpected events occur less often than in the daytime. Consequently risk of price spike is lower at night. Therefore, risk-premium is lower at night (in personal communication with Kevin Hanson, Supervisor, Market Operations Support, ERCOT, 8 February 2016, Taylor TX). An estimate of risk premium can be obtained from mean difference of RT and DA, which is 2.97 USD/MWh. Risk premium estimated in the same way using only prices that have occurred between 3 p.m. and 4 p.m. is a lot higher, 19.65 USD/MWh. This difference results from the higher spike probability of RT price spike in the afternoon. In addition, there are two short spike periods in RT in the example week. The risk of extremely high price realized for electricity buyers that have not bought their electricity

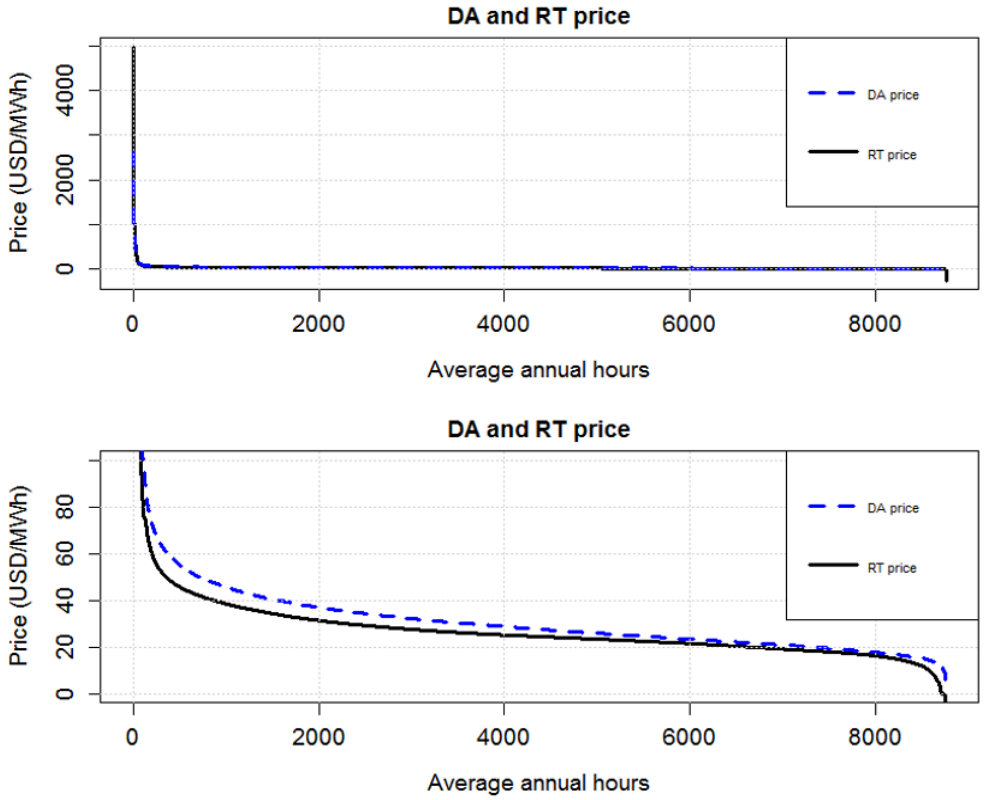

Figure 3.6: Duration curves of DA and RT price in top graph shows that huge majority of both prices are low. Huge prices of above 1000 USD/MWh and negative prices are rare. Bottom graph shows part of duration curve where price is in range 0-100 USD/MWh. It can be seen that DA prices in range 40-100 USD/MWh are slightly more frequent than RT prices in the same range. On the contrary, RT prices below 40 USD/MWh are more frequent, and especially very low prices are relatively more frequent in RT than in DA market.

from DA market. Figure 3.9 shows empirical joint distribution of DA and RT. There is no clear linear or non-linear correlation visible. However, vast majority of points in the graph is close to origin, which makes it virtually impossible to draw conclusions from the graph. It is possible that there is a dependency between DA and RT, when both are at their most common level. A zooming to prices below 100 USD/MWh is shown in figure 3.10. There is clear linear correlation between DA and RT when DA is between 0 and around 40 USD/MWh and then the two prices are often almost equal. When DA is higher than 40 USD/MWh, RT price distribution is more uniform and

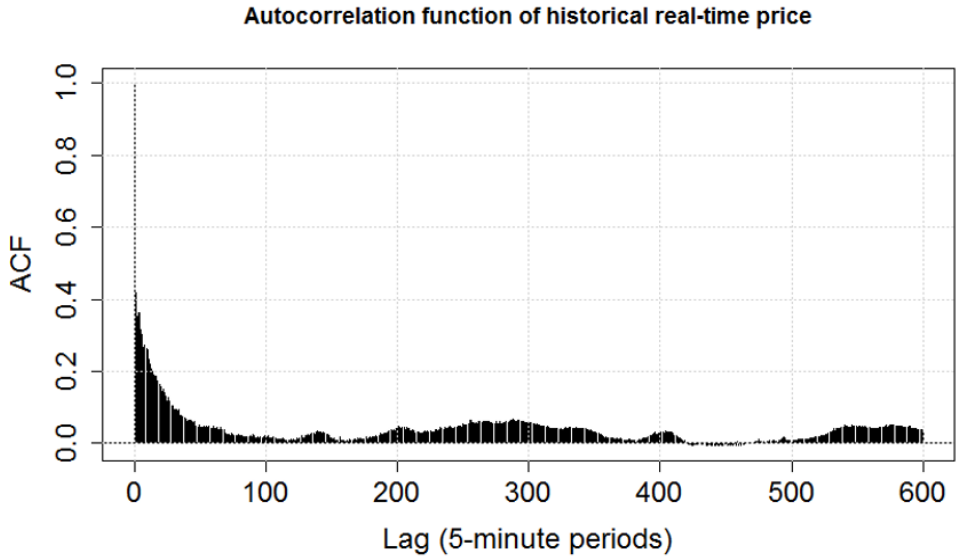

Figure 3.7: Empirical autocorrelation function of RT price shows strong positive autocorrelation on lags shorter than 50 time-steps (5-minute periods). There is seasonal autocorrelation at lags near 288, i.e. 24 hours.

RT is more often below DA. This results from the risk premium relating to DA and RT markets described above. DA is often almost equal to RT, when there is little uncertainty related to RT. In such situations prices are often low. When uncertainty of RT is greater, DA is more often above 40 USD/MWh and RT does not correlate so strongly with DA.

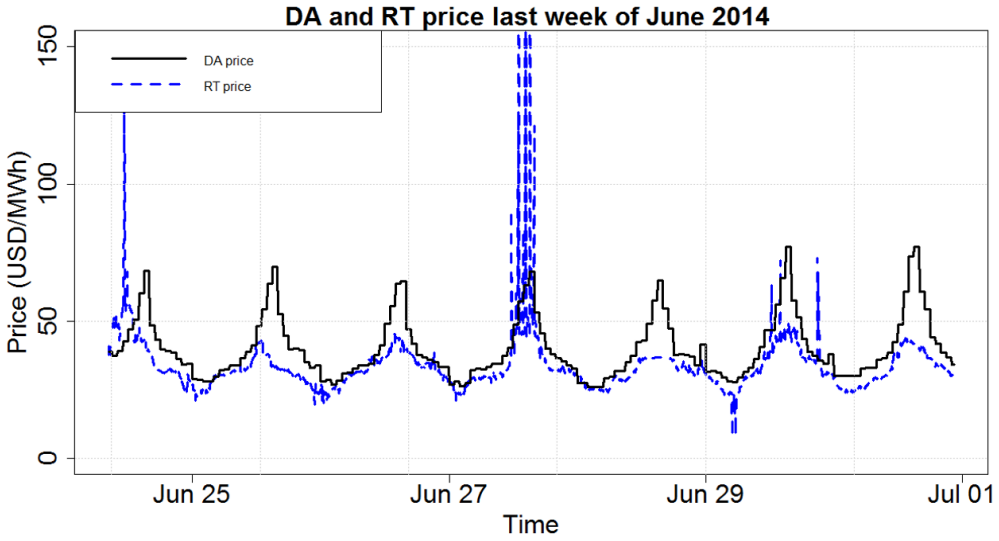

Figure 3.8: DA and RT prices of last week of June 2014. RT price does not show as clear daily pattern as DA price. There are two RT price spikes in the period. Highest DA price of day is often above the highest RT price of day.

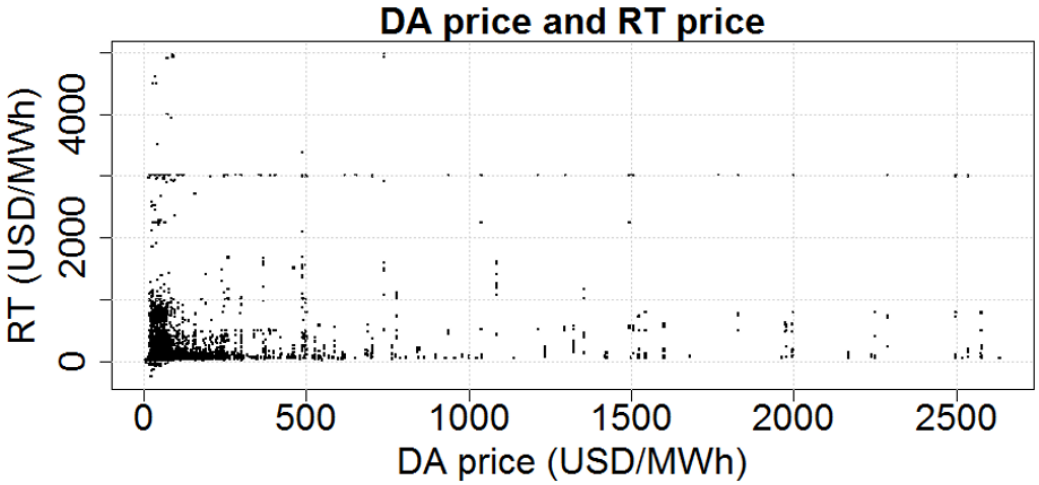

Figure 3.9: Scatter plot of DA price and RT price. Vast majority of points are close to origin. RT prices above 1000 USD/MWh have occurred at all levels of DA price, but their relative share is greater when DA price exceeds 500 USD/MWh. Negative RT prices have only occurred when DA price has been at low level.

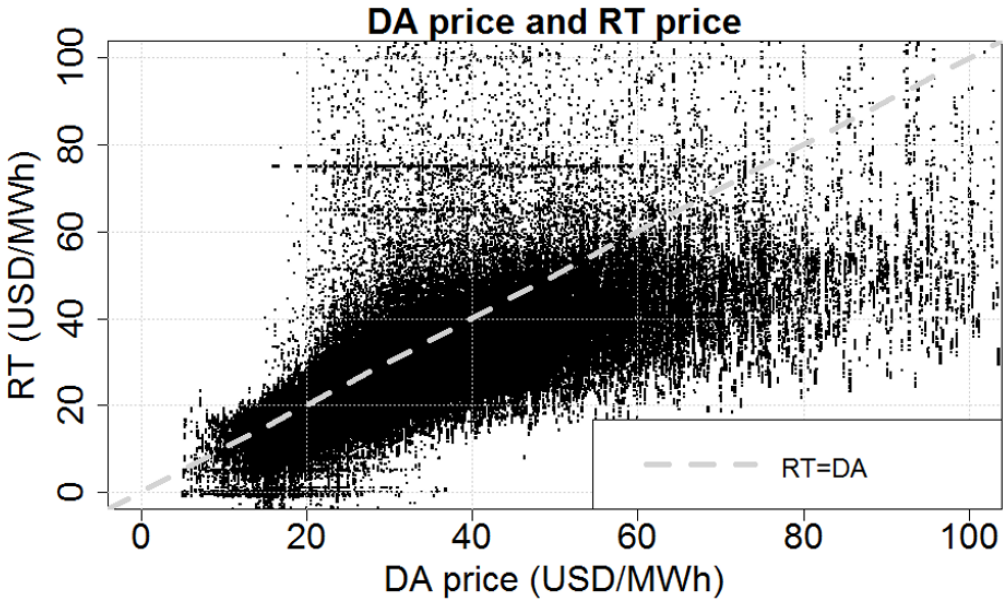

Figure 3.10: Scatter plot of DA price and RT price when both are below 100 USD/MWh. Grey dashed line represents situation when RT and DA prices are equal. RT price clearly increases as DA price increases between 10 and 40 USD/MWh. Systematic increase of RT price with DA price stops when DA price exceeds 40 USD/MWh.

3.1.3 Spread between RT and DA price

We denote the spread between DA and RT price by $\Delta = RT - DA$. Figure 3.11 shows Δ of years 2011-2015 as a function of time. There are large positive Δ values in the same periods that RT has spiked. Due to price floor RT can be significantly below DA only when DA is high. Therefore, large negative Δ values can occur only at periods that DA is high. As high DA prices have been rare, there are only a few large negative Δ values in the period.

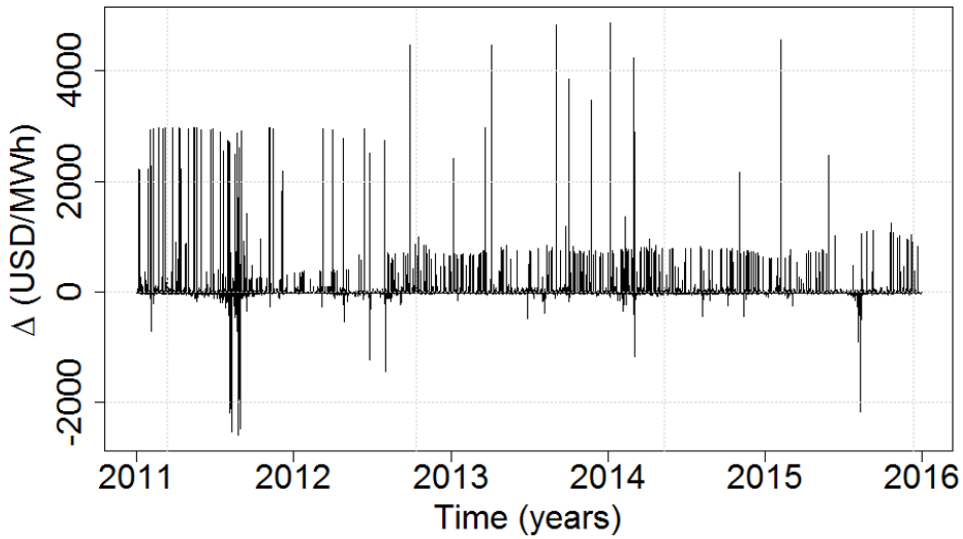

Figure 3.11: Δ, i.e. spread between RT and DA price from year 2011 to 2015. Most of the time Δ is close to zero. Occasional large negative and positive deltas occur. Majority of them are positive.

Duration curve of Δ in figure 3.12 shows that the highest Δ value in 2011-2015 was around 5000 USD/MWh and lowest Δ value was around -3000 USD/MWh. Absolute value of Δ exceeded 100 USD/MWh only a few hours per year. Duration curve is flat when both ends are excluded. This means that more than 8500 hours a year, spread between DA and RT has been very small. Mean Δ is -2.97 USD/MWh, meaning that on average RT is that much lower than DA.

Empirical autocorrelation function of Δ is shown in figure 3.13. Δ exhibits strong positive autocorrelation on lags shorter than 50 time-steps (5-minute periods). There is a positive seasonal autocorrelation at lags near 288, i.e. 24 hours, much stronger than in case of RT. This is possibly a result of risk

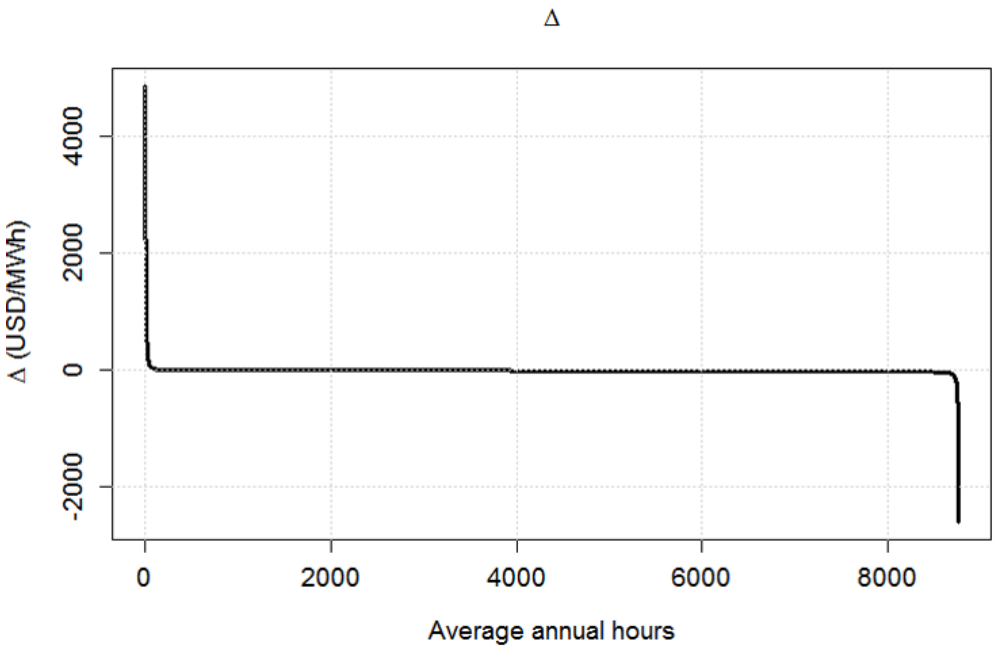

Figure 3.12: Duration curve of Δ shows that vast majority of Δ values are close to zero, but tails of Δ distribution are very long.

premium of DA regularly widening in high-load day-time and getting smaller in the night-time.

Scatter plot of DA price and Δ in figure 3.14 shows clear negative correlation between the two variables. This is obvious, resulting from the way we defined Δ as $RT - DA$. Vast majority of points are close to origin. Large positive Δ values are possible only when DA price is low enough due to price cap. Similarly, large negative Δ values are possible only when DA price is high enough due to price floor. As we saw in figure 3.10, RT follows DA closely when DA is in its most common level of less than 40 USD/MWh. Then Δ is often close to zero and small unexpected events in market can make it slightly negative or positive. The uniform distribution of Δ means that DA being below 40 USD/MWh, RT price simulation can be carried out by first simulating DA and then sampling Δ and adding it to RT. However, simulation of RT can not be conducted in the same way when DA is above 40 USD/MWh since distribution of Δ is strongly dependent on DA as we saw in figure 3.14. In that region, however, distribution of RT is rather uniform as seen in figure 3.9. Therefore, we can simulate RT price by sampling RT

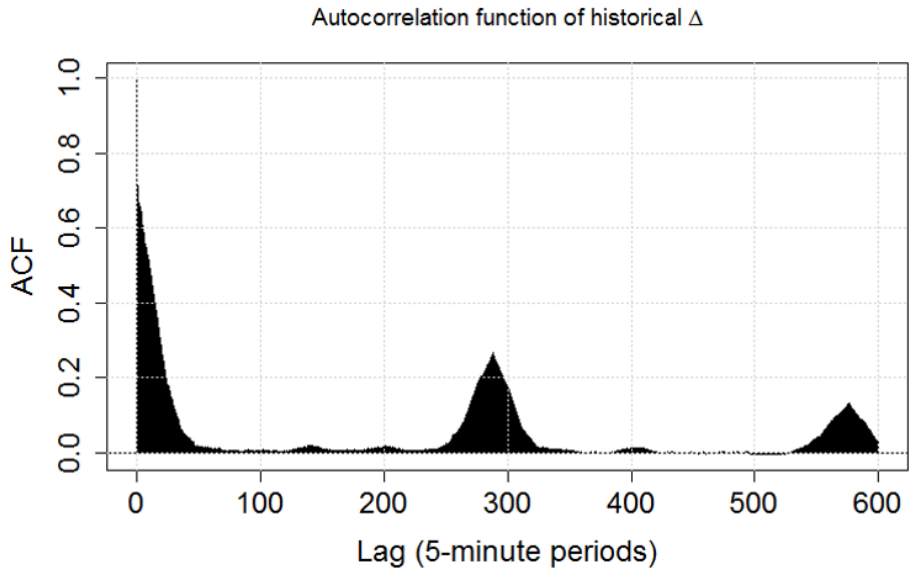

Figure 3.13: Empirical autocorrelation function of Δ shows strong positive autocorrelation on lags shorter than 50 time-steps (5-minute periods). There is clear positive seasonal autocorrelation at lags near 288, i.e. 24 hours.

directly, when $DA \geq 40$.

3.1.4 Demand

Figure 3.15 shows demand of electricity in ERCOT from the beginning of year 2011 to the end of year 2015. The demand varies between 20000 and 70000 MW. There is clear yearly seasonal variation in demand because of the greater need for air-conditioning in summer and for warming houses in winter. There might also be an increasing trend as the demand seems to increase slightly towards the end of the period. Furthermore, there is a clear seasonal variation in demand on daily basis as we saw in figure 3.3. Load increases in the morning. Night-time load is often significantly lower than in the day-time. Figure 3.16 shows a duration curve of demand. It can be seen that demand being over 60000 MW is rare, but maximum demand is around 70000 MW. This means that upper tail of demand distribution is long. To be able to serve the demand in the highest demand times even if some outages occur, system needs more than 10000 MW of capacity, that is unused most of the year. Seasonal qualities at different lengths described

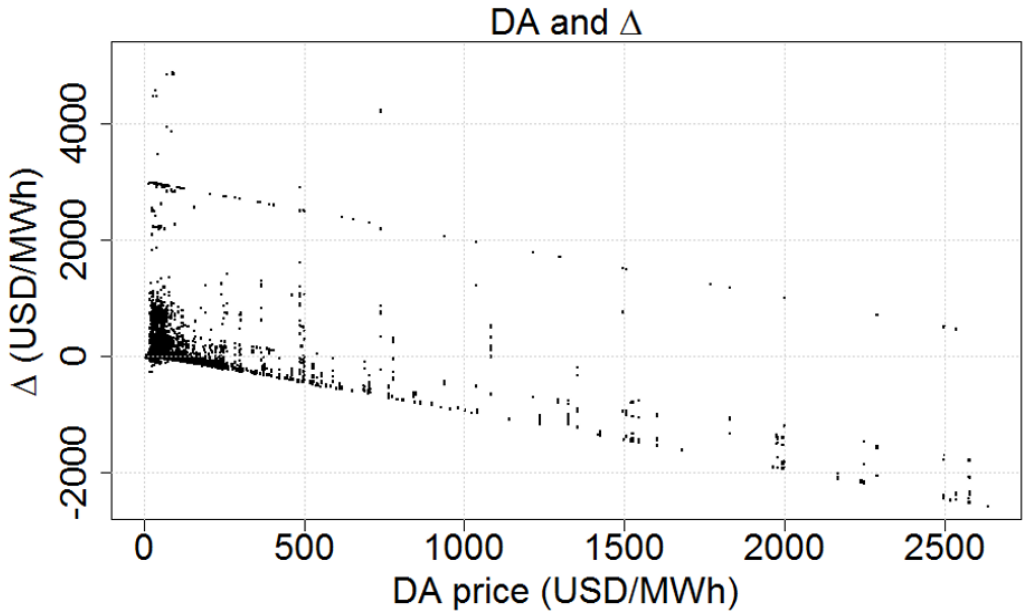

Figure 3.14: Scatter plot of DA price and Δ shows clear negative correlation between the two variables. Vast majority of points are close to origin. Large positive Δ values have occurred only when DA price is low enough due to price cap. Similarly, large negative Δ values have occurred only when DA price is high due to price floor.

above can be seen in the empirical autocorrelation function of demand, too, shown in figure 3.17, where clear peak is seen around lags 288 (=24 hours) and its multiples. This results from the fact that demand is usually close to what it was 24 hours earlier. Autocorrelation on short lags is close to one, meaning that demand changes between consecutive 5-minute periods are often rather small.

Top graph of figure 3.18 shows scatter plot of demand and RT price. It seems that demand does not explain RT price movements. High prices are rather evenly distributed at all demand levels excluding lowest demand values of the period. Slightly more cap prices occur when demand is close to its maximum than otherwise. Bottom graph shows scatter plot of demand and Δ. Like high RT values, high Δ values have occurred evenly at all demand levels. However, large absolute values of Δ are a lot more frequent when demand is high, over 60000 MW. These are caused by demand being almost equal to total available generation capacity. Price increases naturally when more expensive generation has to be dispatched, but also through price-

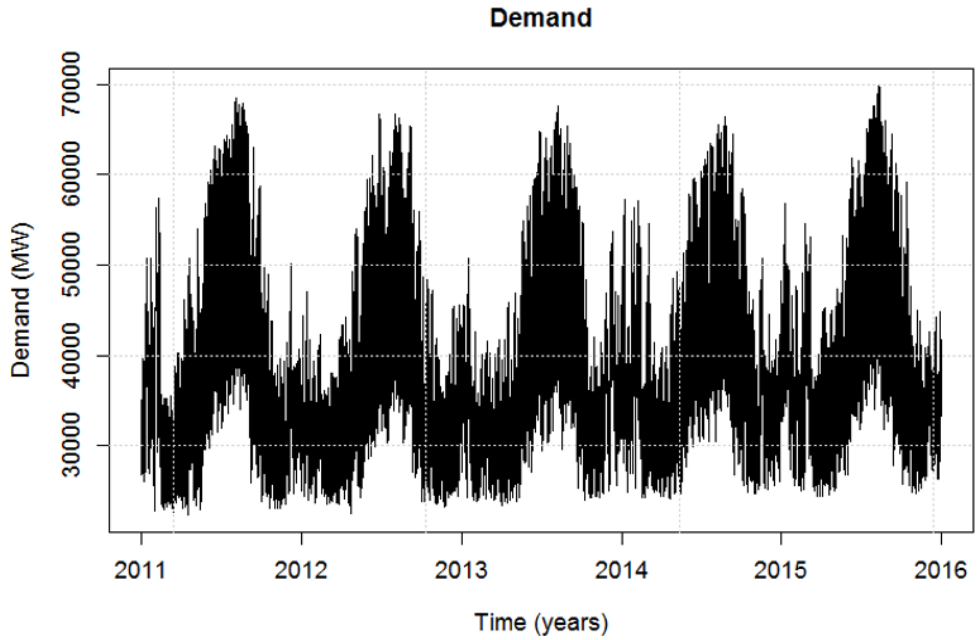

Figure 3.15: ERCOT demand 2011-2015 varies between 20000 and 70000 MW. There is clear yearly seasonal variation, summer-time demand being higher than rest of the year. There are also some high-demand periods in winter-time, especially in years 2011, 2014, and 2015.

adder, that was deployed in ERCOT in 2014. Price-adder increases price when surplus capacity, i.e. available generation capacity exceeding demand is less than 5 GW [13]. A zooming to most common Δ level in figure 3.19 shows that distribution of Δ is rather uniform at its normal levels conditional on demand. Therefore, we do not need demand as an explanatory variable of Δ. Most high RT prices that have occurred at high-demand times are not a result of high demand itself, but rather of low surplus capacity.

Scatter plot of day-ahead forecast and actual demand in figure 3.20 shows very strong positive correlation between the two variables. This means that demand forecast errors are small most of the time. Sometimes, though, actual value differs slightly from forecast. Both ends of point cloud are thinner than middle part. It suggests that forecasts are often more accurate when demand is close to its annual maximum or minimum. However, this may as well be a result of natural variation and a lot more points being in the middle of cloud.

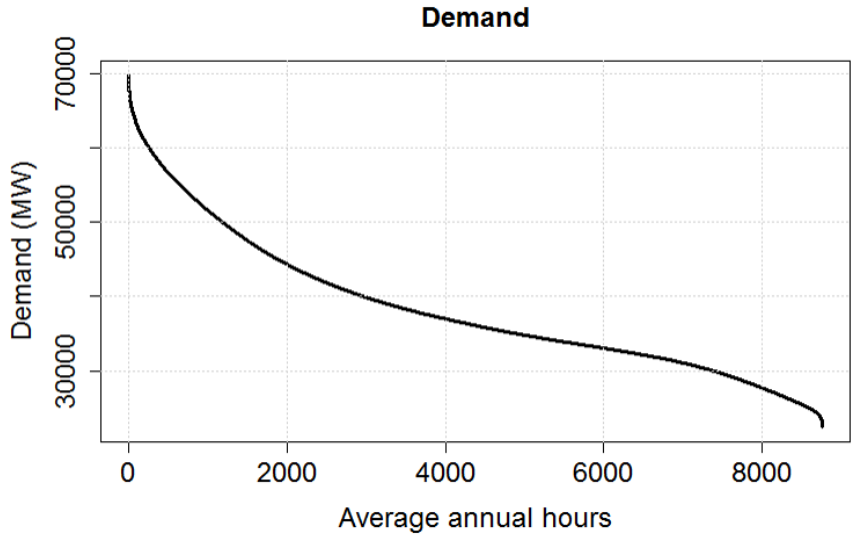

Figure 3.16: Duration curve of demand shows that there are less than 500 hours per year in which demand exceeds 60000 MW. Majority of hours, demand is between 30000 and 50000 MW.

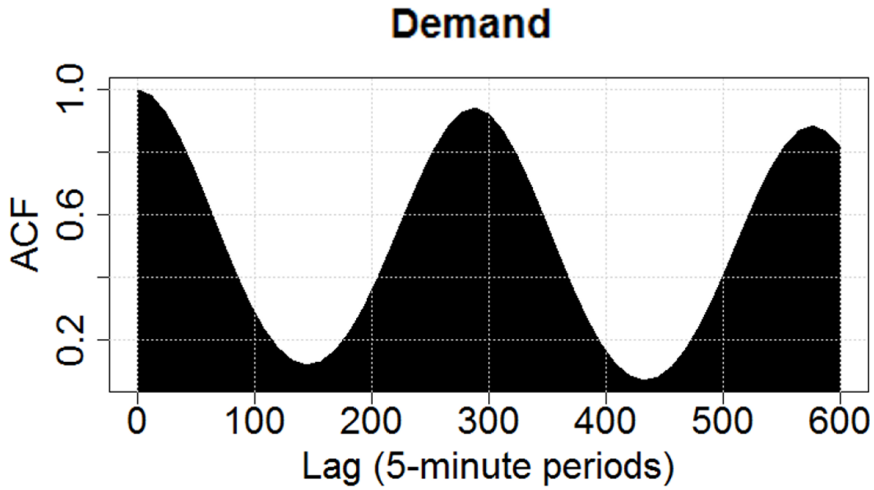

Figure 3.17: Autocorrelation function of demand shows very strong autocorrelation at lags shorter than 20 time-steps, i.e. around two hours. There is very strong seasonal 24-hour autocorrelation.

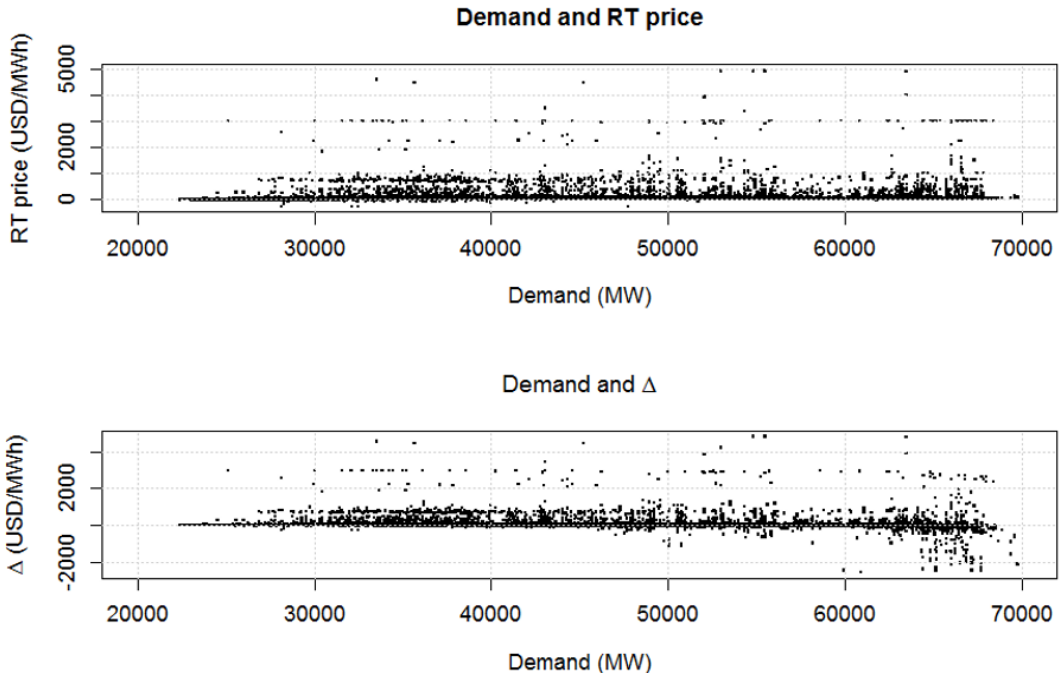

Figure 3.18: Top graph shows scatter plot of demand and RT price. No clear correlation is visible. However, there are slightly more high RT prices when demand is close to its maximum due to cap prices of the year 2011. Bottom graph shows scatter plot of demand and Δ. Variance of Δ is higher when demand is high.

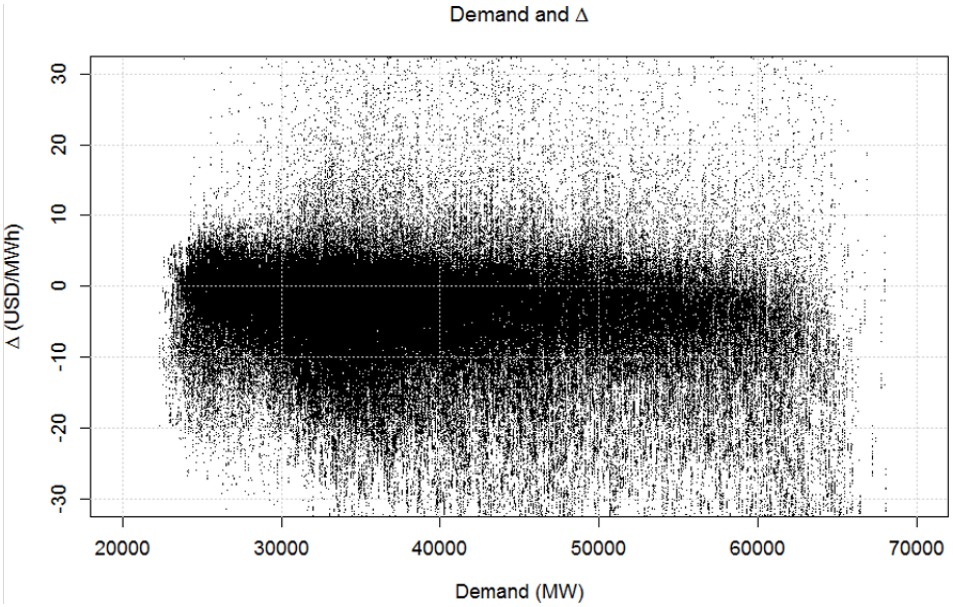

Figure 3.19: Scatter plot of demand and Δ limited to $-30 - 30$ USD/MWh, its most common level. There is no clear correlation between the two variables.

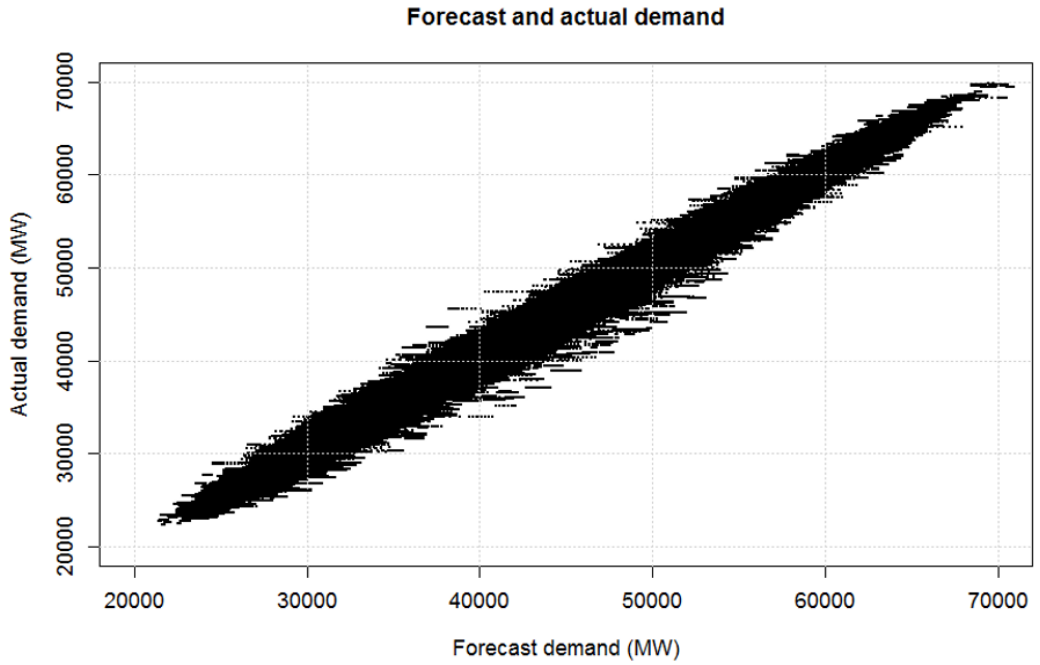

Figure 3.20: Scatter plot of forecast and actual demand shows very strong positive correlation between the two variables. Both ends of the cloud are slim, which tells us that very strong and very weak wind conditions are easiest to forecast.

3.1.5 Wind generation

Rapid increasing trend in installed wind generation capacity makes studying impact of wind generation on RT price very important for the purpose of our study. If we could not capture the impact accurately enough, there would be a good chance that our forecast series would not exhibit true future characteristics of future ERCOT RT price.

Figure 3.21 illustrates ERCOT wind generation from year 2011 to 2015. Top graph shows that wind generation varies between close to zero and more than 10000 MW. There is an increasing trend as a result of growing installed wind generation capacity. Bottom graph shows wind generation as a percentage of installed capacity. For example, 100% would mean that all wind generating units are generating at their full capacity. There is no clear trend and each year the values have varied from close to zero to a little over 80 %. A mean of all these values, *capacity factor*, is 32 %, which tells that on average each wind generating unit has run on 32 % of its maximum output in 2011-2015. This number sounds reasonable, as it is a lot lower than 100 %, which would require idealistic perfect wind conditions and no maintenance breaks at generators. Top graph of figure 3.22 shows duration curve 2011-2015 of wind generation in MW. It shows that generation of more than 8000 MW has occurred only 1000 hours per year on average. However, this graph is not very informative due to rapid growth of installed wind generation capacity. Bottom graph shows duration curve of wind generation as a percentage of installed capacity. We saw in figure 3.21 that there is no trend in percentage values. Therefore, percentage values give more constistent information about wind conditions. Output exceeding 80 % of installed capacity is rare (100 hours per year) and zero generation situations occur hardly ever. Distribution between these extremes is rather even. Empirical autocorrelation function of wind generation on top graph of figure 3.23 shows very strong autocorrelation at lags shorter than 50 time-steps, i.e. around four hours, because of slowly moving weather fronts. There is also a weak seasonal 24-hour component in autocorrelation possibly due to regular daily variation of wind speed in Texas. Bottom graph shows empirical autocorrelation function of error of day-ahead forecast of wind generation. It exhibits strong short-term autocorrelation, too, but not as strong as wind generation. There is no clear 24-hour peak in the graph. As we may expect, accuracy of forecasts made for same time of two consecutive days do not correlate. Top graph of figure 3.24 shows scatter plot of wind generation and RT price. There seems to be no clear correlation between the two variables. High RT prices, though, occur a little more frequently, when wind generation is low than otherwise. Bottom graph shows scatter plot of wind generation and

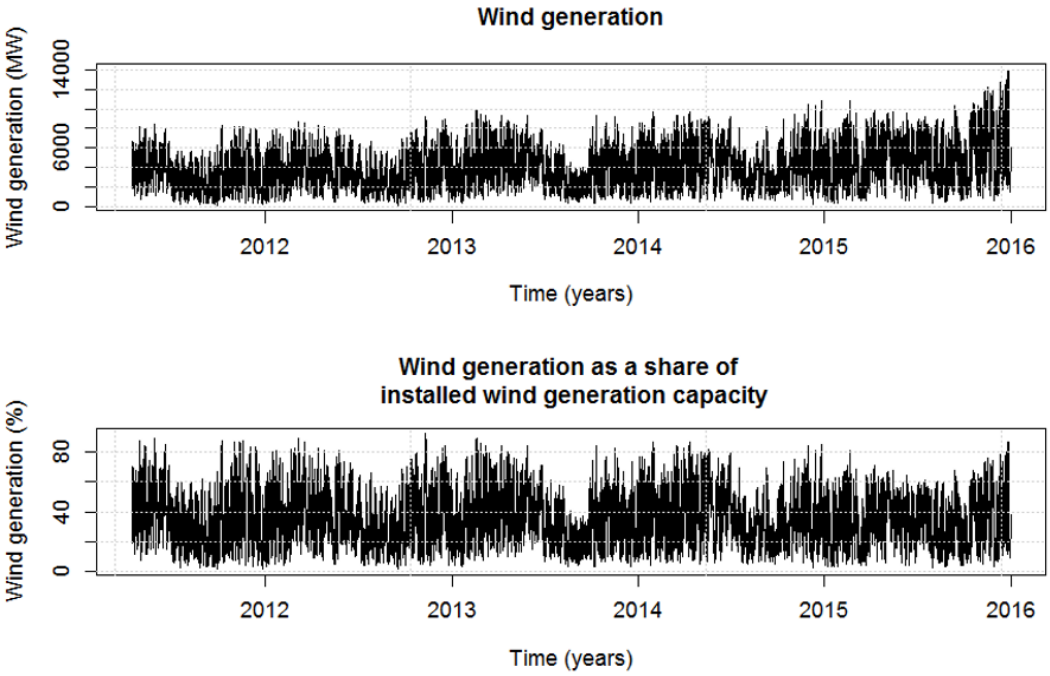

Figure 3.21: Top graph shows wind generation from 2011 to 2015. It varies a lot between 0 and more than 10000 MW. There is an increasing trend visible. The highest values are in the end of year 2015 due to windy conditions and largest installed wind generation capacity. Bottom graph shows wind generation as a percentage of installed capacity. There is no clear trend and each year the values have varied from close to zero to a little over 80 %.

Δ. There is no clear correlation visible. Absolute value of Δ is almost always very small, when wind generation exceeds 10000 MW. Scatter plot of historical day-ahead forecast made by ERCOT and actual wind generation in figure 3.25 shows that the values of the two variables often differ only a little. It means that wind generation forecasts are relatively accurate most of the time. A few points are far from the cloud, which means that actual wind generation sometimes differs significantly from forecast. The ends of the cloud are slim. Forecasts are generally more accurate, when there is very strong or hardly any wind, than when wind speed is normal. This is possibly a result of the usual shape of generation curve of a wind generator as function of wind speed. Generation increases rapidly at intermediate wind speeds causing more variance in generation output, when it is at its middle level [22].

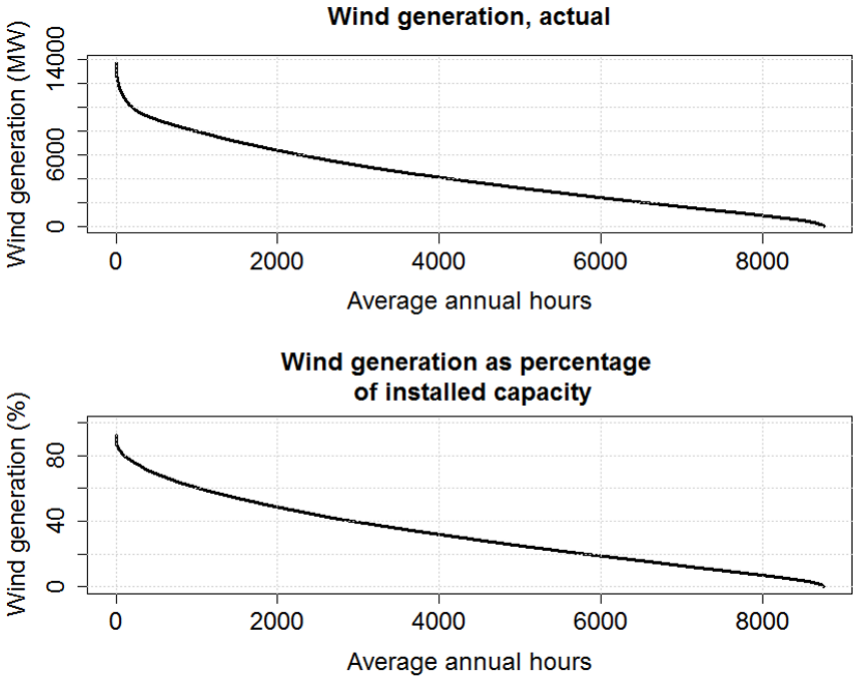

Figure 3.22: Top graph shows duration curve of wind generation in MW. It shows that generation of more than 8000 MW has occurred only 1000 hours per year on average. The highest values are even more rare. Bottom graph shows wind generation as a percentage of installed wind generation capacity. Wind generation exceeds 30 % of installed capacity on average 4000 hours per year. Zero generation situations occur hardly ever.

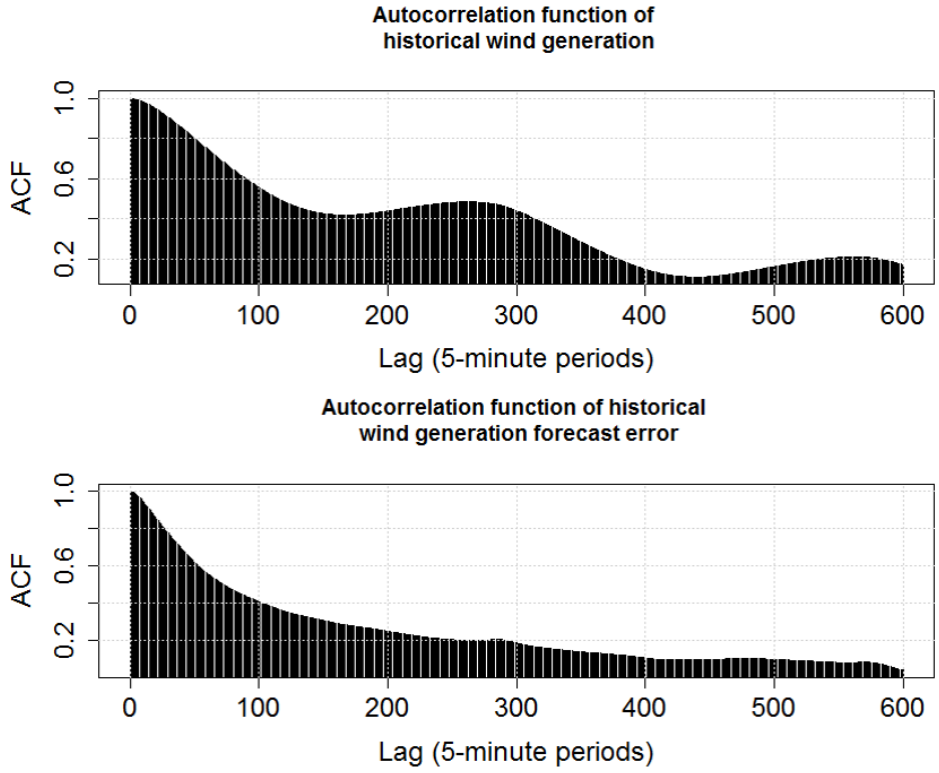

Figure 3.23: Empirical autocorrelation function of wind generation on top graph shows very strong autocorrelation at lags shorter than 50 time-steps, i.e. around four hours. There is also seasonal 24-hour component in autocorrelation. Bottom graph shows empirical autocorrelation function of wind generation forecast error. It exhibits strong short-term autocorrelation, too, but not as strong as wind generation. There is no clear 24-hour peak in the autocorrelation function.

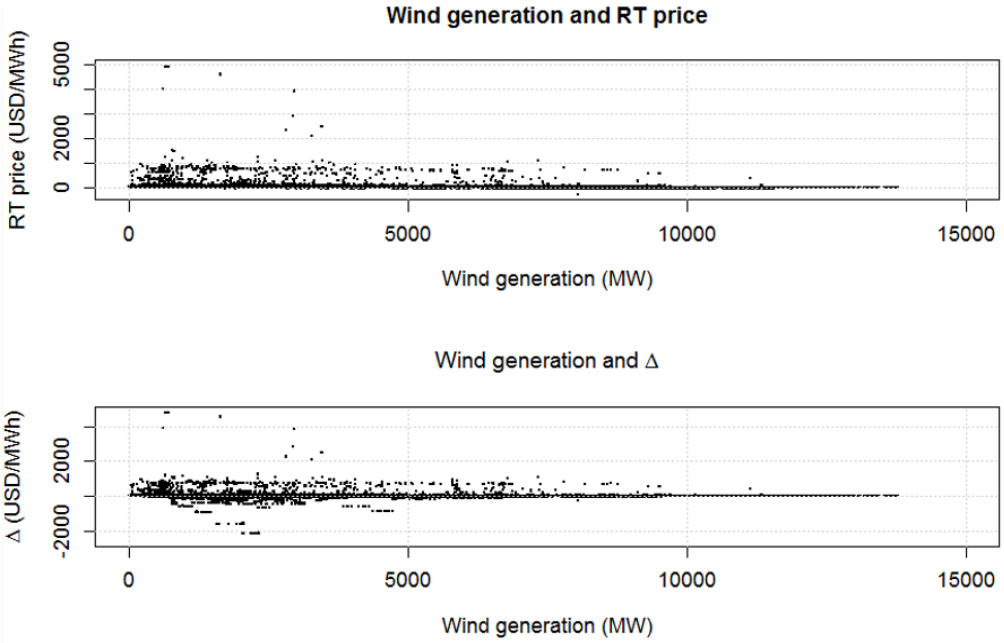

Figure 3.24: Top graph shows scatter plot of wind generation and RT price. No clear correlation is visible. However, there are slightly more high RT prices when wind generation is close to zero. Bottom graph shows scatter plot of wind generation and Δ. There are no clear correlations either. Δ is almost always very small in absolute value, when wind generation exceeds 10000 MW.

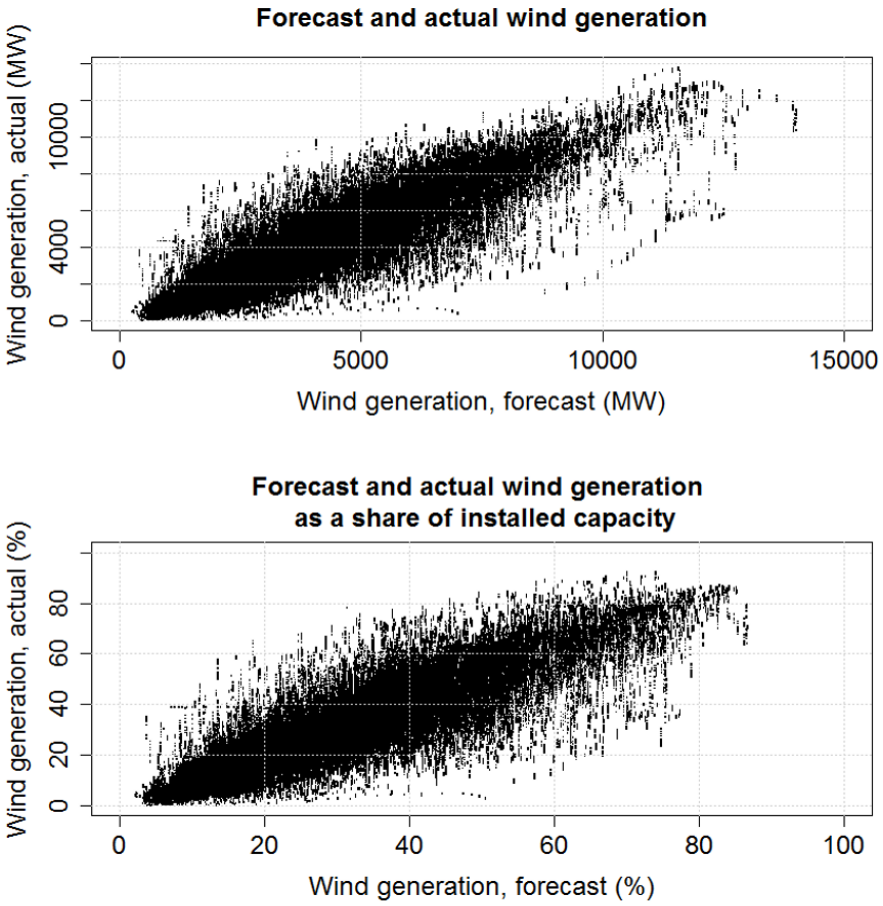

Figure 3.25: Scatter plot of day-ahead forecast and actual wind generation shows positive correlation between the two variables, indicating that wind generation forecasts are relatively accurate most of the time. There are some points far from the middle of the cloud, which means that actual wind generation sometimes differs significantly from forecast. The ends of the cloud are slim, which shows that forecasts are generally more accurate, when there is very strong or hardly any wind, than when wind speed is normal.

3.1.6 Net-load

Figure 3.26 shows a scatter plot of net-load and RT on top graph. There is no clear correlation between the two variables. High prices occur at all levels of net-load except the lowest net-loads of less than 20000 MW. Bottom graph shows scatter plot of net-load and Δ. It seems that net-load does not explain occurrence of high Δ values very well. Largest negative Δ values, however, have occurred at very high net-load times. Possibly DA is high because of high probability of high RT due to forecast surplus capacity being low. However, sometimes the risk does not realize and RT is consequently low.

When net-load increases rapidly, only fast-ramping power plants can respond to the change. Some of these power plants, especially GTs, have high marginal generation cost. Need to dispatch them leads to high electricity prices. Therefore, we can expect to see positive correlation in net-load change speed and RT. Top graph of figure 3.27 shows scatter plot of 5-minute net-load change, NLC, and RT. NLC is positive when net-load is increasing. Clearly, there are more high RT values at positive than negative net-load change values. Bottom graph shows scatter plot of NLC and Δ. Large negative Δ values are more frequent when NLC is negative. Similarly, large positive Δ values are more frequent when NLC is positive. The clearest limit for Δ is not $NLC = 0$, but $NLC = 200$ MW. We can use NLC being above or below 200 MW to explain Δ distribution.

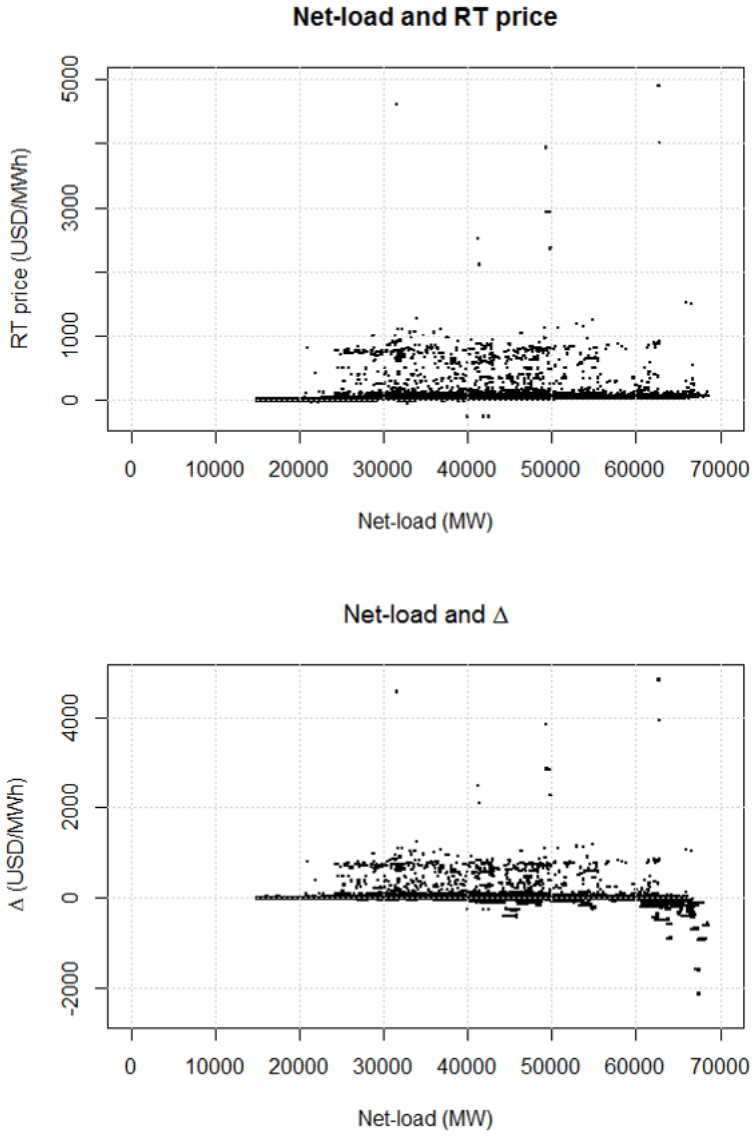

Figure 3.26: Top graph shows scatter plot of net-load and RT price. No clear correlation is visible, but it seems that RT price has always been close to zero when net-load has been below 20000 MW. Bottom graph shows scatter plot of demand and Δ. Large negative Δ values are more frequent, when net-load is high.

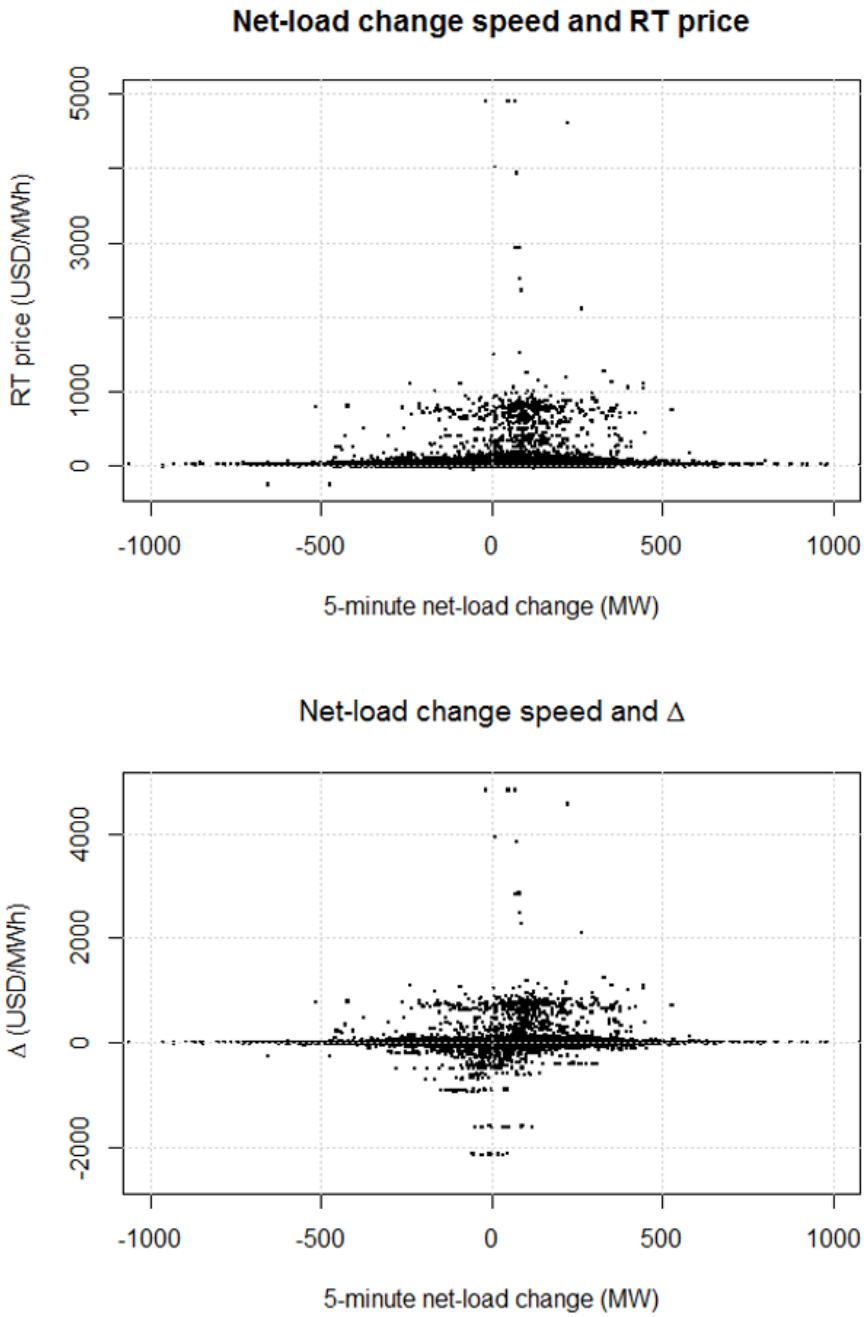

Figure 3.27: Top graph shows scatter plot of 5-minute net-load change and RT price. Positive value means growing net-load. High prices are more frequent when net-load change is positive. Bottom graph shows scatter plot of net-load change and Δ. Large negative Δ values are more frequent when net-load change is negative and large positive Δ values are more frequent when net-load change is positive.

3.1.7 Non-intermittent generation

Actual available non-intermittent generation capacity, $NICA$, represents the maximum net-load that could be served with available generation resources. Generators submit their *current operating plans* (COP) to ERCOT. COPs submitted for next day can be used as day-ahead forecast of available capacity, $NICF$. Figure 3.28 shows scatter plot of $NICF$ and $NICA$. The graph shows very strong positive correlation between the two variables, indicating that last-day changes to COPs are small most of the time. There are several points far from the rest of points, which means that sometimes moderate deviations from forecast have occurred. These deviations are possibly a result of unexpected power plant outages.

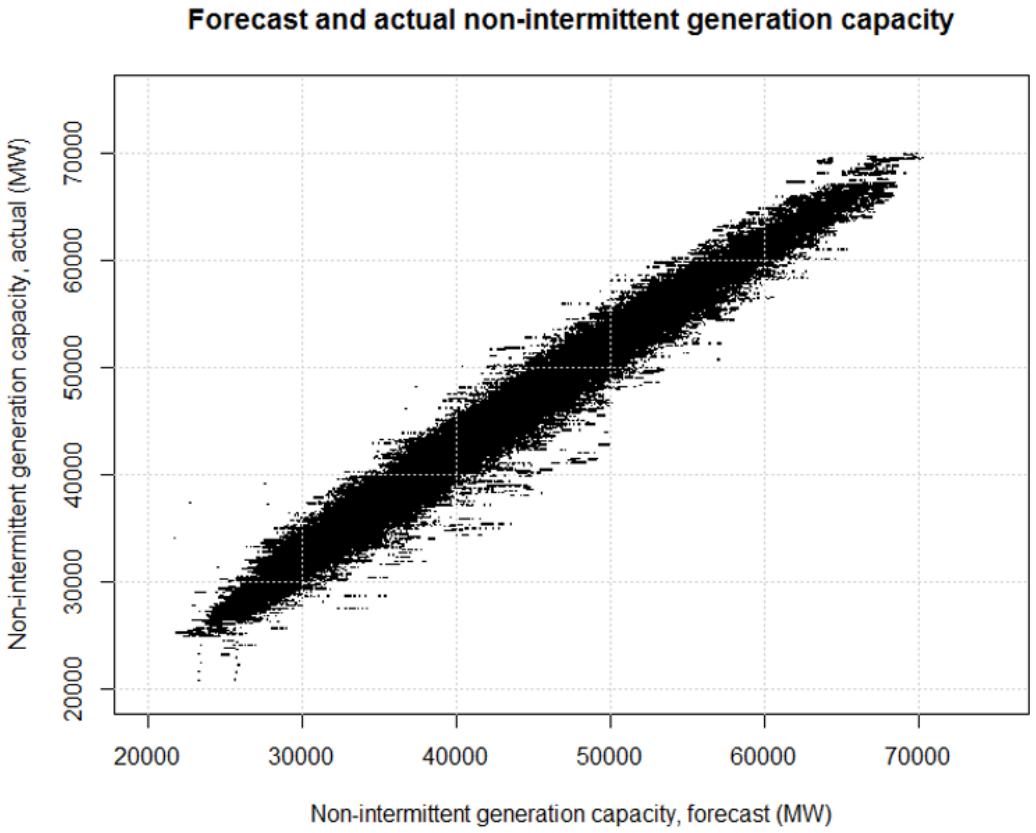

Figure 3.28: Scatter plot of day-ahead forecast and actual values of available on-line non-intermittent generation capacity shows very strong positive correlation, indicating that forecast errors are very small most of the time. There are several points far from the rest of the cloud, which means that sometimes moderate changes from forecast occur.

3.1.8 Surplus capacity

ERCOT implemented an Operating Reserve Demand Curve (ORDC) on June 1, 2014 to incentivize building new generation capacity. ERCOT Investment Incentives and Resource Adequacy in 2012 had showed that system reliability might be threatened otherwise [27], [13]. As a result, scarcity pricing mechanism increases RT price when available reserves are below 5000 MW. Scatter plot of actual surplus capacity $FCA = NICA + IGA - LA$ and RT price in figure 3.29 shows that spike prices mostly occur when surplus capacity is low. Surplus capacity represents the amount by which available generation capacity exceeds load. This is partly a result of scarcity pricing mechanism and partly of power plants with higher offer prices being dispatched.

We can use amount of surplus capacity as an explanatory variable for price spike probability in our simulation methodology.

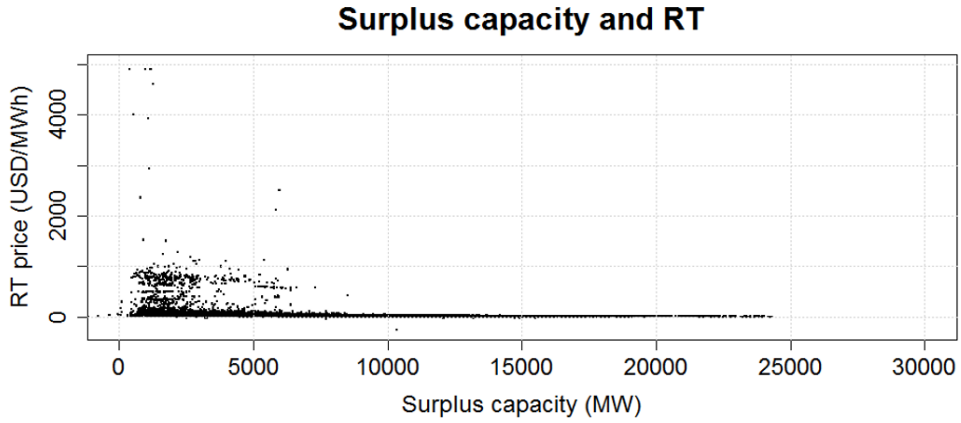

Figure 3.29: Scatter plot of surplus capacity and RT price shows that high prices mostly occur when surplus capacity is low. Share of high RT prices is a lot higher, when surplus capacity is below 7000 MW.

3.1.9 Forecast error of surplus capacity

We define forecast error of surplus capacity, $FCFE$, as difference between actual surplus capacity FCA and day-ahead forecast FCF, $FCFE = FCA - FCF$, where $FCA = NICA + IGA - LA$ and $FCF = NICF + IGF - LF$. Therefore, $FCFE$ represents the difference between actual surplus capacity and day-ahead forecast of surplus capacity, which was available when DA market was cleared. Negative $FCFE$ means that actual surplus capacity is lower than was expected day-ahead. Figure 3.30 shows histogram of $FCFE$. Most of the time $FCFE$ is between -5000 and 2000 MW.

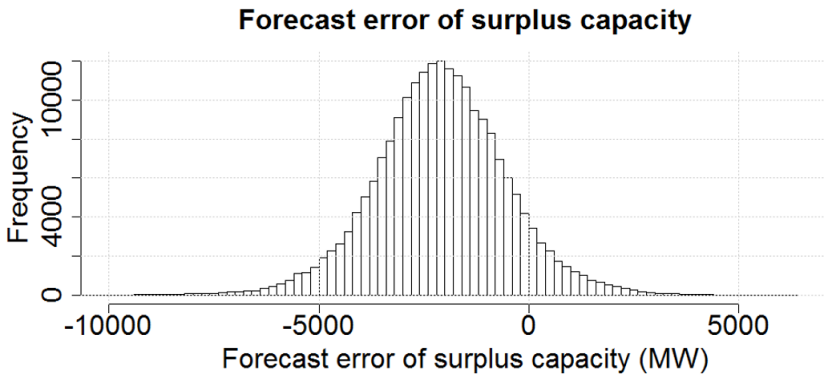

Figure 3.30: Histogram of forecast error of surplus capacity. Positive value means that actual surplus capacity is larger than day-ahead forecast. Histogram shows that forecast slightly underestimates actual surplus capacity most of the time.

We expect that RT is often higher than DA, when $FCFE$ is negative, since there is less extra resources available than was expected. Figure 3.31 shows one example night from June 2014, when a price spike occurred because of largely negative $FCFE$. Top graph shows actual and forecast wind generation. It can be seen that after 22:00 actual generation drops rapidly even though it was forecast to increase slightly. Middle graph shows forecast error of surplus capacity in the same period. It falls sharply from close to zero to -8000 MW. At the same time there is a huge spike in Δ, shown in the bottom graph. This occurred six days after ORDC was implemented and price-adder was one factor that increased RT price.

Figure 3.32 shows scatter plot of forecast error of surplus capacity and RT price on top graph. There is no clear correlation between the two variables. Bottom graph shows scatter plot of forecast error of surplus capacity and Δ. Large negative values of Δ occur only, when forecast error of surplus

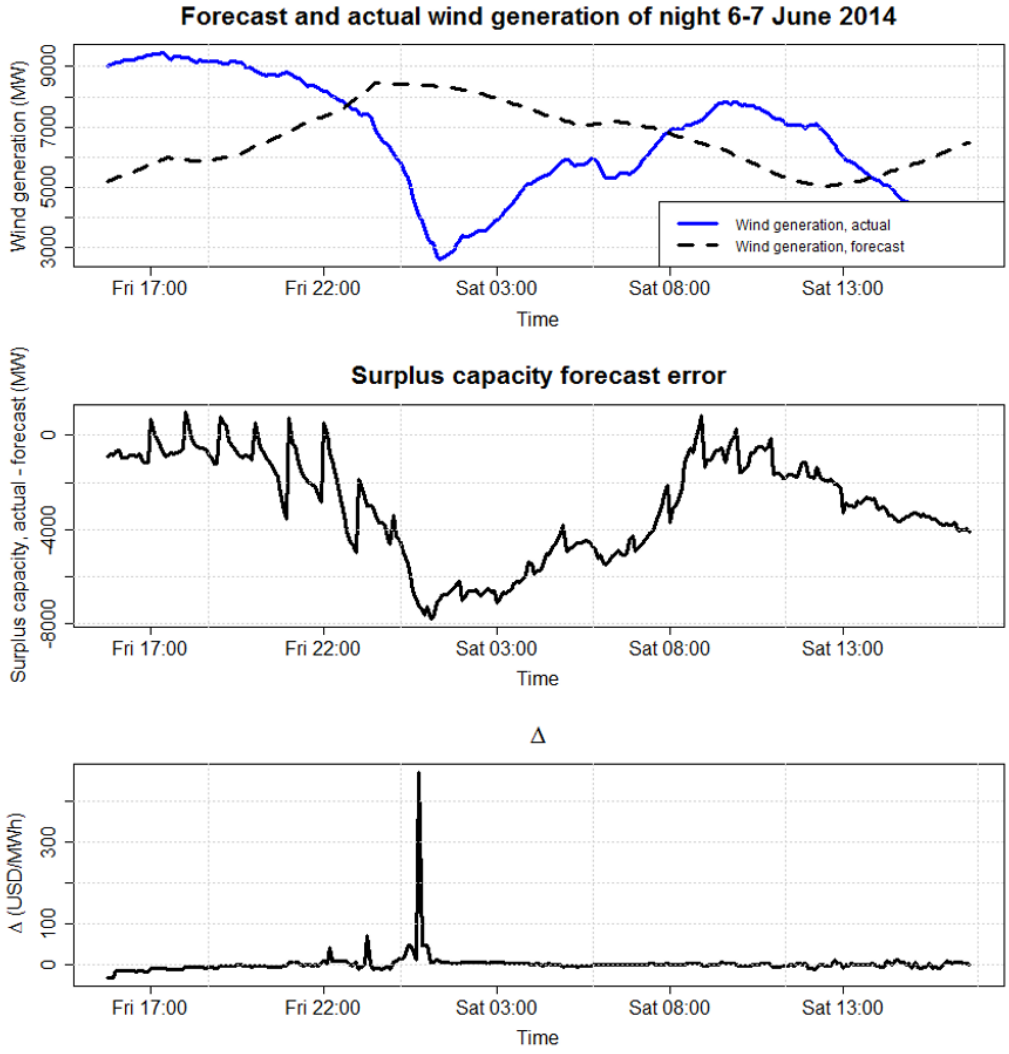

Figure 3.31: A graph of wind generation of one example night in June 2014 shows how the generation unexpectedly drops rapidly from 8000 MW to less than 3000 MW. Middle graph shows that forecast error of surplus capacity drops consequently to -8000 MW. Bottom graph shows that a large negative Δ has occurred at the same time.

capacity is positive and large positive values of Δ are more frequent, when forecast error of surplus capacity is negative. Therefore, we can use forecast error of surplus capacity to explain Δ movements. Clearest difference in Δ distribution is seen when FCFE is divided into two classes with limit $FCFE = -1800$ MW.

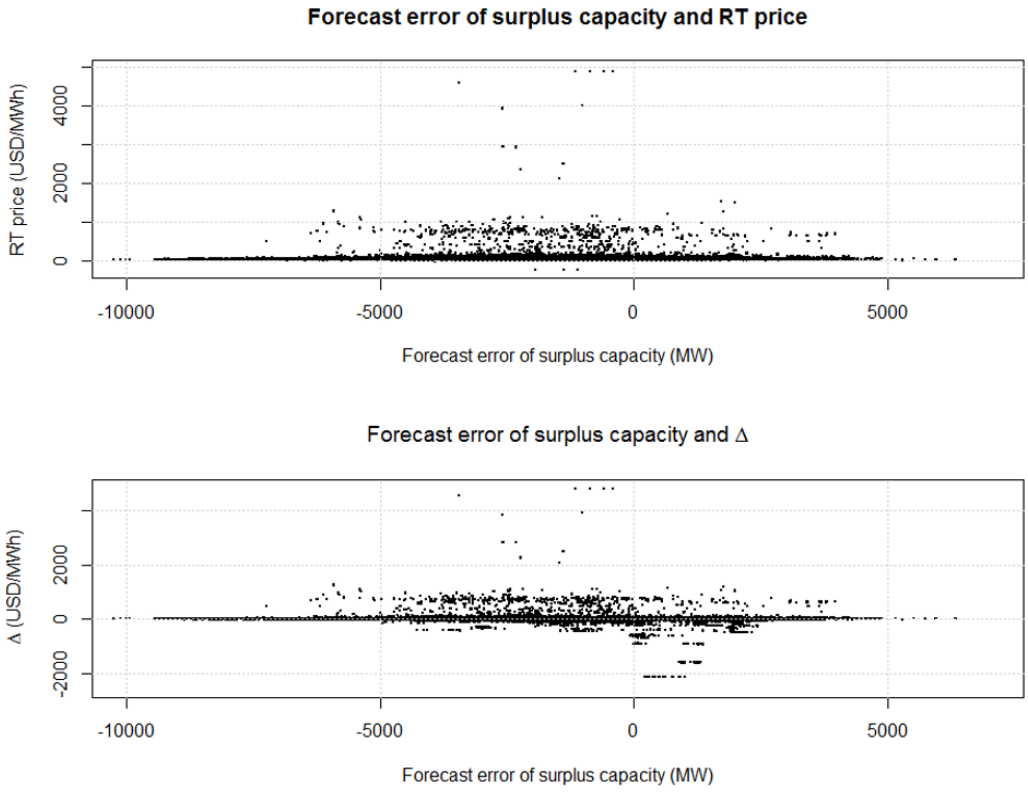

Figure 3.32: Top graph shows scatter plot of forecast error of surplus capacity and RT price. Positive value on horizontal axis means that actual surplus capacity is larger than day-ahead forecast. There is no clear correlation between the two variables. Bottom graph shows scatter plot of forecast error of surplus capacity and Δ. Large negative Δ values are more frequent, when forecast error of surplus capacity is positive. Large positive Δ values are a little more frequent, when forecast error of surplus capacity is negative.

3.1.10 Node prices and hub prices

We have used as RT price in analyses the hub average RT price, which is simple average of hub prices in ERCOT. Next we will briefly study differences in prices of different nodes and hubs. Each power plant can choose whether it sells its electricity for node price or hub price (in personal communication with M.Sc. Matti Rautkivi, Origination, Wärtsilä North America, 8 February 2016, Taylor TX). Therefore, it is important to know what impact on profitability the selection of reference price has.

Figure 3.33 shows that congestion causes larger differences in RT price between hubs (bottom graph) than within a hub (top graph). This is just one example analysis and a lot more node prices should be analysed to build comprehensive understanding of general differences in hub and node prices. We will do simulations in this study using hub average price. If future prices of a specific node are needed, it is safest choice to do simulations for the price of that node. It is simple as only one input time series needs to be changed. However, there is a risk that adding new generation or transmission capacity to that node has an impact on future prices. Therefore, hub price is more conservative choice, as it is more robust with respect to local changes in the grid.

3.2 Selection of simulation methodology

As expected, our analysis did not show perfect correlation between RT price and any other variable. However, we found some patterns that explain a part of RT price variation. Because of no perfect correlations and significant future uncertainty that is present in our simulations we can use a stochastic method in RT price simulation. Due to large variation in RT price we divide it to several price classes. We build a modified bootstrap model where price class is sampled first and then price from the right class. Explanatory variables have impact on sampling probabilities of price classes. Since DA market modelling is relatively straightforward and accurate, we can use DA market modelling software to create simulated DA price and DA-forecasts of explanatory variable values as inputs for our RT price modelling method. We will sample Δ, when $DA < 40$ and RT otherwise. Stochasticity and impact of explanatory variables are applied to selection of price class for each time-step.

We simulate spike prices separately and define spike period loosely as a period in which RT price is mostly in it highest percentile. A precise definition will be given later in the study. We will use the surplus capacity

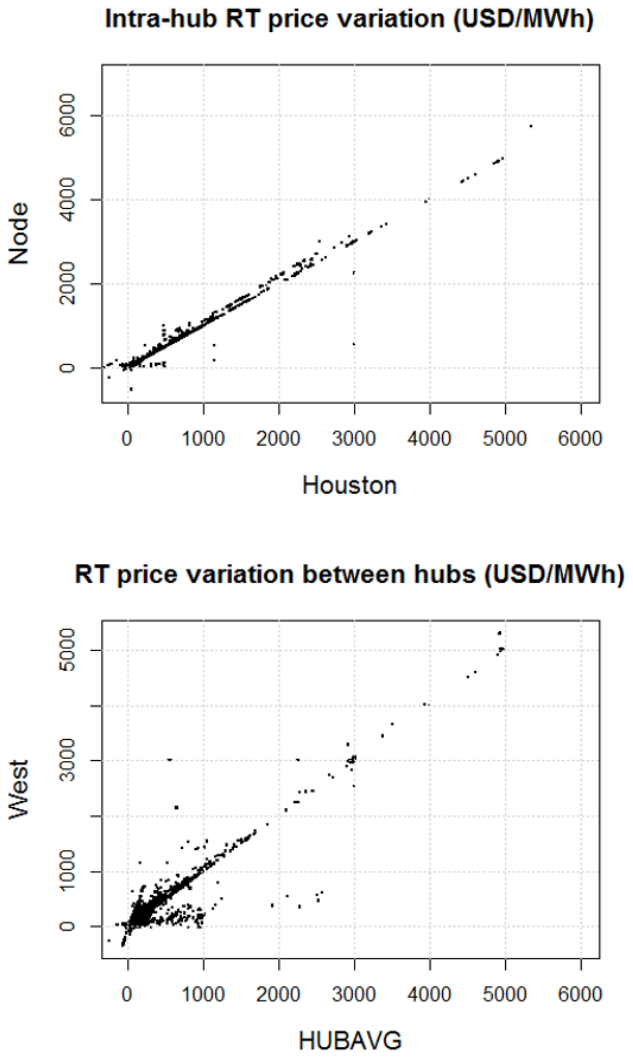

Figure 3.33: Comparison of example hub prices and node prices. Top graph shows RT price of Houston hub on horizontal axis and RT price of one node in Houston hub on vertical axis. Prices differ only a little from each other. Prices of hub average and West hub, shown on bottom graph, differ more often than intra-hub prices on top graph.

FCA and DA price DA to determine spike starting probability with limits 7000 MW and 30 USD/MWh, respectively.

We also found that RT price being on non-spike level, variation can be explained by forecast error of surplus capacity $FCFE$, change speed of net-load NLC, DA price DA, and price class of previous time-step. Limits for

Table 3.2: Division points of variables used in non-spike situations.

$FCFE$	-1800 MW
NLC	-200 MW
DA	[25,40] USD/MWh
$RT(t-1)$	$[16.7, 19.1, 21.3, 22.9, 24.5, 26.2, 28.8, 32.6, 39.8]$
$\Delta(t-1)$	$[-13.7, -8.6, -6.0, -4.2, -2.7, -1.5, -0.4, 0.7, 2.4]$

explanatory variables $FCFE, NLC$, and DA that make clear difference in price distribution are shown in table 3.2. Table also shows limits of price classes and delta classes determined as deciles of hub average RT price and Δ.

Strong autocorrelation exhibited by RT price is supposedly obtained in simulated time series as a result of strongly autocorrelated explanatory variables, by taking into account the price of previous time-step when sampling prices, and by sampling prices for spike periods as blocks of historical spike prices instead of sampling each price separately.

Chapter 4

Stochastic RT price simulation

In this chapter we introduce a simulation method and use it to obtain future time series of RT price. The main idea of our simulation methodology is to determine historical correlations between the explanatory variables and dependent variable RT price, and then use these correlations together with simulated future values of explanatory variables in stochastic simulation of future RT price. By correlation we do not mean Pearson correlation coefficient or any other specific statistic, but any quantity that can be used to statistically measure the association between two variables. The most commonly used measure of correlation in our simulation is (empirically estimated) conditional probability.

First we divide the historical RT price data into price classes and estimate empirical probabilities of price belonging to each price class. Then we simulate future time series of explanatory variables using DA market modelling software and stochastic methods. Finally, we conduct a stochastic simulation of RT price based on simulated explanatory variables and estimated probabilities.

As explained in Chapter 3, we can divide the time steps to two subsets by DA. When $DA < 40$, we will simulate Δ, i.e. the spread between DA and RT. Otherwise we simulate directly RT. To take into account the impact of explanatory variables on RT price, we introduce a concept of *dynamic time step category* C, which is calculated for all time steps. In our modelling approach C is the reason behind different price patterns occurring in different conditions. Not all variation of RT price can be explained by C, but we have used all information that was determined useful in our analysis. Biggest weakness of our simulation approach that is based on bootstrap method is imperfect prediction accuracy. It results from the wide intervals in which the explanatory variables can vary within a dynamic time step category. Assuming that the values of explanatory variables are precise, accuracy could

be improved by using for example some regression method. However, our approach enables us to create thousands of future RT price time series, and individual errors caused by random selection of values can be assumed to cancel in large number of simulations. Secondly, the future time series of explanatory variables are far from precise. Therefore, greater accuracy in prediction would not help to improve accuracy of RT price prediction. For a thorough discussion of bootstrap methods in time series simulation, see [5].

4.1 Calculating needed historical inputs

We start our simulation by dividing the historical RT prices RT into groups corresponding to n_{RT_C} price classes $G_{RT_C} = \{RT|RT \in RT_C\}, RT_C = 0, 1, ..., n_{RT_C} - 1$ and historical deltas $\Delta = RT - DA$ into groups corresponding to n_{Δ_C} delta classes $G_{\Delta_C} = \{\Delta|\Delta \in \Delta_C\}, \Delta_C = 0, 1, ..., n_{\Delta_C} - 1$. A simple way to determine delta classes and price classes is to use $p_\Delta\%$ and $p_{RT}\%$ percentiles of historical data, respectively, where $p_\Delta = \frac{100}{n_{\Delta_C}}$ and $p_{RT} = \frac{100}{n_{RT_C}}$. For example, if $n_{\Delta_C} = 4$, we will have all Δ values below the first quartile of Δ in $\Delta_C = 0$, values of Δ below median, but above the first quartile in $\Delta_C = 1$, etc. The prices of time steps t with $DA(t) \geq DA_{limits,C}(1)$ are divided into price groups G_{RT_C}, and deltas of time steps t with $DA(t) < DA_{limits,C}(1)$ are divided into delta groups G_{Δ_C}.

We group the values of explanatory variables into variable *dynamic time step category* $C = 0, 1, ..., n(C)$, which is calculated separately for each time step and includes all needed information about the conditions that apply during time step t and previous price class or delta class. The largest C, C_{spike} corresponds to spike time step. Definitions of all dynamic time step categories C for $n_{\Delta_C} = n_{RT_C} = 12$ are shown in table 6.1 in appendix.

When we have determined $C(t)$ and $\Delta_C(t)$ or $RT_C(t)$ for each historical time step (5-minute period) t, we can calculate empirical conditional probability distributions of (i) Δ belonging to each Δ_C, $P(\Delta_C|C)$, and similarly of (ii) RT belonging to each price class RT_C, $P(RT_C|C)$ conditional on C of that time-step. Conditional probabilities are calculated in the usual way by dividing the number of time-steps with each C and Δ_C or RT_C by total number of periods with that C. The probability distribution of delta classes becomes

$$P(\Delta_C|C) = \frac{n(\Delta_C \cap C)}{n(C)}, C \neq C_{spike},$$

where $n(x)$ is the number of periods in the history data with the quality x.

The probability distribution of price classes becomes

$$P(\Delta_C|C) = \frac{n(\Delta_C \cap C)}{n(C)}, C \neq C_{spike}.$$

Because of huge variation in RT and DA, we treat periods of spiky price separately. We build a vector $G_{RT_{spike}}$ containing the prices of spike periods in a chronological order. We also determine empirical probabilities of spike starting conditional on the amount of surplus capacity FCA, and DA price DA. These probabilities will be used later in simulation of future spike periods.

To give a precise definition of spike period, we first specify a limit $l_{spike,RT}$, above which all RT are spikes, and a similar limit $l_{spike,DA}$ for DA. We denote by binary X_{spike} each time step being spike ($X_{spike} = 1$) or non-spike ($X_{spike} = 0$). If RT or DA (or both) of time-step t exceeds the limit, $RT(t) \geq l_{spike,RT}$ or $DA(t) \geq l_{spike,DA}$, a spike period starts. The spike period length is determined as the longest possible L such that the share of spike prices in the spike period is at least S_{spike}, a percentage parameter specified in advance.

Exact description of methodology is shown in algorithm 1.

Algorithm 1 Historical correlations

Data Historical time series of RT price and explanatory variables

RT_t = Real-time price
DA_t = Day-ahead price
IGA_t = Intermittent generation, actual
$NICA_t$ = Available non-intermittent generation capacity, actual
LA_t = Load, actual
IGF_t = Intermittent generation, forecast
$NICF_t$ = Available non-intermittent generation capacity, forecast
LF_t = Load, forecast
NLC_t = Net-load change from previous time step

Output

Grouped RT prices, deltas, and spike lengths

$G_{\Delta_C} = \{\Delta_{hist} | \Delta_{hist} \in \Delta_C\}$ = Historical Δ values of delta class $\Delta_C \in \{0, 1, 2, ..., n_{\Delta_C} - 1\}$

$G_{RT_C} = \{RT_{hist} | RT_{hist} \in RT_C\}$ = Historical RT prices of price class $RT_C \in \{0, 1, 2, ..., n_{RT_C} - 1\}$

$G_{RT_{spike}} = [RT_{hist}(t) | X_{spike,hist}(t) = 1]$ = Historical RT prices of spike periods in chronological order

G_L = Lengths of historical spikes

Estimated probabilities

$P(\Delta_C|C)$ Estimated probabilities of delta classes $\Delta_C \in \{0, 1, 2, ..., n_{\Delta_C} - 1\}$ conditional on dynamic time step category C

$P(RT_C|C)$ Estimated probabilities of price classes $RT_C \in \{0, 1, 2, ..., n_{RT_C} - 1\}$ conditional on dynamic time step category C

$P(X_{start} | I_{DA_{limit,prob}}(DA), I_{FCA_{limit,prob}}(FCA))$ Estimated probabilities of spike starting conditional on day-ahead price and actual surplus capacity

Parameters

$l_{spike,RT}$ limit above which all RT prices are spike prices

$l_{spike,DA}$ limit above which all DA prices are spike prices

$DA_{limit,prob}$ = Division point of DA for conditional probability of spike period starting

$FCA_{limit,prob}$ = Division point of FCA for conditional probability of spike period starting

S_{spike} = Share of time steps that must have spike price within a spike period

l_{RT_C} = vector that contains limits of price classes

l_{Δ_C} = vector that contains limits of delta classes

$DA_{limits,C}$ = vector that contains limits of DA for determining C

$L_{history}$ = number of historical time steps

$FCFE_{limit}$ = limit of $FCFE$ for determining C

NLC_{limit} = limit of NLC for determining C

procedure

Spike periods

$X_{spike} \leftarrow 1 - (1 - I_{l_{spike,RT}}(RT))(1 - I_{l_{spike,DA}}(DA))$

$t \leftarrow 1$

while $t \leq L_{history}$ **do** ▷ Go through all history data

$C^0(t) = 4[I_{DA_{limits,C}(1)}(DA(t)) + I_{DA_{limits,C}(2)}(DA(t))] + 2I_{FCFE_{limit}}(FCFE(t)) + I_{NLC_{limit}}(NLC(t))$

if $X_{spike}(t) = 0$ **then** ▷ Non-spike time step
if $t > 1$ **then**
if $X_{spike}(t-1) = 1$ **then**
$C(t) = C^0(t) + 12(n_{\Delta_C} + n_{RT_C})$
else
$C(t) = C^0(t) + 12\Big[\Delta_C(t-1) + I_{DA_{limits,C}(1)}(DA(t-1))[n_{\Delta_C} + RT_C(t-1) - \Delta_C(t-1)]\Big]$
end if
else ▷ First time step
$C(t) = C^0(t) + 12\Big[\Delta_C(t) + I_{DA_{limits,C}(1)}(DA(t))[n_{\Delta_C} + RT_C(t) - \Delta_C(t)]\Big]$
end if
if $DA(t) \geq DA_{limits,C}(1)$ **then**
$RT_C(t) = I_{l_{RT_C}(1)}(RT(t)) + I_{l_{RT_C}(2)}(RT(t)) + ... + I_{l_{RT_C}(n_{RT_C}-1)}(RT(t))$
$G_{RT_C(t)} \leftarrow [G_{RT_C(t)}, RT(t)]$
else
$\Delta_C(t) = I_{l_{\Delta_C}(1)}(\Delta(t)) + I_{l_{\Delta_C}(2)}(\Delta(t)) + ... + I_{l_{\Delta_C}(n_{\Delta_C}-1)}(\Delta(t))$
$G_{\Delta_C(t)} \leftarrow [G_{\Delta_C(t)}, \Delta(t)]$
end if
$t \leftarrow t + 1$
else ▷ Spike period starts
for all $\tau = 1, 2, ..., L_{history} - t + 1$ **do**
$S(\tau) = \frac{X_{spike}(t) + X_{spike}(t+1) + ... + X_{spike}(t+\tau-1)}{\tau}$
end for
$L_s \leftarrow \min(\tau | S(\tau) < S_{spike}) - 1$ ▷ Last consecutive time step with share of spike prices large enough
$t_d \leftarrow 0$
while $t_d \leq L_s$ **do** ▷ Cut off non-spike prices from end
if $X_{spike}(t + L_s - 1 - t_d) = 0$ **then**
$t_d \leftarrow t_d + 1$
else
$L \leftarrow L_s - t_d$
$X_{spike}(r), r = t, t+1, ..., t+L-1 \leftarrow 1$
$t_d \leftarrow \infty$
end if
end while
$G_L \leftarrow [G_L, L]$ ▷ Save length and prices
$G_{RT_{spike}} \leftarrow [G_{RT_{spike}}, RT(t), RT(t+1), ..., RT(t+L-1)]$
$t \leftarrow t + L$ ▷ Go to the end of this spike and continue
end if
end while
end procedure

4.2 Creating future time series of explanatory variables

We will use actual values of the explanatory variables and day-ahead forecasts of them to create forecast electricity price scenarios. In particular, we will use the difference between the simulated day-ahead forecast and simulated actual value of several explanatory variables (non-intermittent generation capacity, wind generation, demand) to determine the probability of price being on different levels. One possible way to create future time series of explanatory variables would be simulating future actual values first, and then bootstrapping forecast errors from historical forecast errors to achieve future forecast values of explanatory variables. However, since we can simulate the forecast values using commercial electricity market modelling software, we will do the same procedure in the opposite order. We simulate forecast values first and then we simulate actual values of explanatory variables.

It is important to understand that we do not try to predict what the demand and values of other explanatory variables will be each hour of the next years in ERCOT. Instead, we will create several different forecast time series, none of which is likely to be precisely accurate. However, all simulated time series will approximately exhibit the statistical quantities and seasonal patterns that have occurred in history.

4.2.1 Day-ahead market simulation

As explained in Chapter 3, there are relatively accurate modelling methods for forecasting DA price, when several explanatory variables are known. We can use a commercial modelling software to future DA price. The same software can be used to create day-ahead forecasts of demand and generation, based on historical forecast data and estimate of future changes in several meaningful market quantities such as development of ERCOT wind generation capacity. As DA market modelling is relatively straightforward and established methods exist, we focus in this study on creating real-time scenarios based on simulated DA prices and forecasts. We take DA electricity price and day-ahead forecasts of demand and generation as given. However, it is possible that simulated time series would be biased as a result of randomness in DA market simulation. We will carry out several stochastic simulations of explanatory variables to reduce impact of such bias on results of our RT price simulation.

The outputs of DA market simulation that we will use as inputs for our RT market simulation are processes of day-ahead forecasts of intermittent

generation IGF_t, demand LF_t, non-intermittent generation capacity $NICF_t$, and DA electricity price DA_t.

4.2.2 Simulating actual values of explanatory variables through Markov chain and bootstrap methods

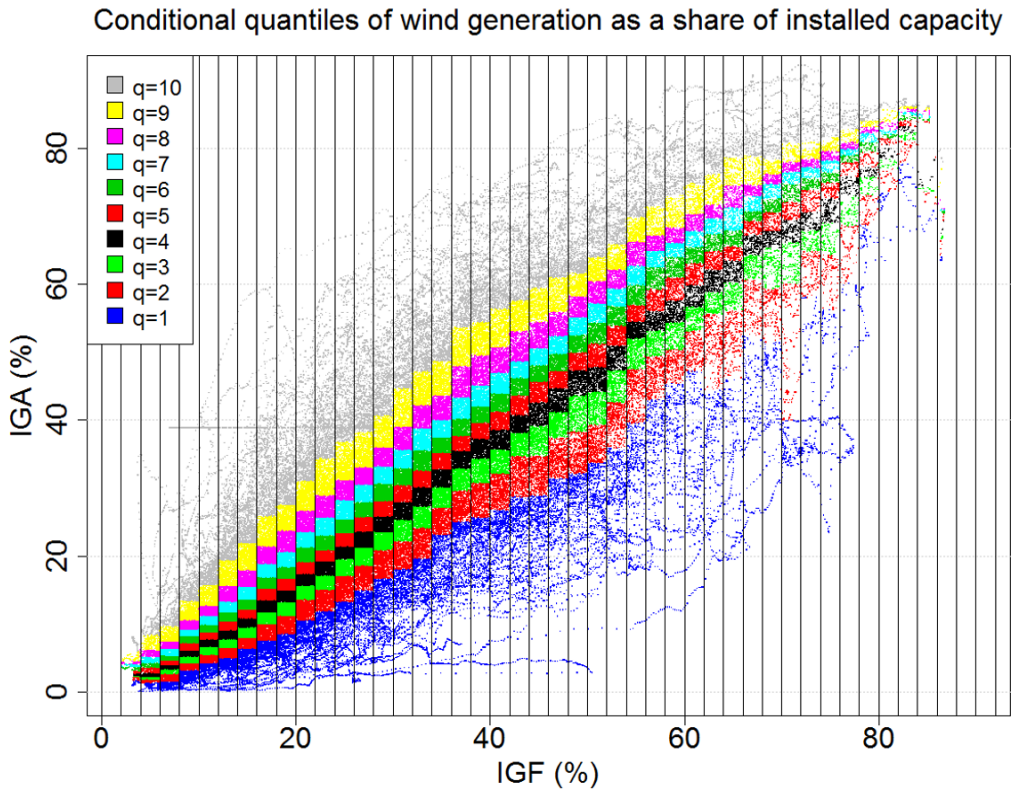

Figure 4.1: Classes of day-ahead forecast are separated by black vertical lines. Different colours represent the groups of actual values separated by conditional quantiles. Within each interval separated by two vertical lines, the ten colours correspond to deciles of values of intermittent generation in that interval.

Forecast values obtained using DA market modelling software are used as a starting point when we create future time series of actual values of the explanatory variables. We use bootstrap method to choose an actual value corresponding to each forecast value. To avoid incredibly large differences between simulated day-ahead and actual values, we divide the forecast values

into equally wide classes XF_C. We sample each actual value from the historical actual values occurred when the historical forecast was in the same class. To preserve the strong autocorrelation of actual values in history, we divide the historical actual values occurred while XF was in class XF_C to q sets $G_{XA_{XF_C,Q_{XA}}}, XF_C \in \{0, 1, 2, ..., n_{XF_C} - 1\}, Q_{XA} \in \{1, 2, ..., q\}$ of equal number of observations. We simulate the future groups Q_{XA} so that historical autocorrelation is preserved. We use Markov chain approach with transition matrix estimated from historical observations. Then we simply sample the actual value $XA(t)$ from the set corresponding to group Q_{XA} and class XF_C.

Figure 4.1 shows the groups of IGA separated by conditional quantiles $Q_{XA} = 1, ..., q = 10$ (colours). $Q_{XA} = 1$ corresponds to 10% conditional quantile, $Q_{XA} = 2$ corresponds to 20% conditional quantile, etc. Black vertical lines represent limits of classes XF_C.

Next, we define the algorithm that can be used to simulate future actual values of any variable X. We will use this algorithm for IGA, $NICA$, and LA. Simulation methodology is shown in algorithm 2.

Algorithm 2 Actual future values of explanatory variables

Data

Processes of variable X, i.e. the variable of interest

$XA_{hist.}$ = historical actual values of X

$XF_{hist.}$ = historical day-ahead forecast values of X

$XF_{fut.}$ = future day-ahead forecast values of X

Output

$XA_{fut.}$ = Simulated future scenario of X

Parameters

q = number of groups separated by conditional quantiles per XF_C used in simulation

n_{XF_C} = number of equally wide intervals XF is divided into

l_{XF_C} = vector of limits of XF_C

$l_{XA_{XF_C}}$ = vector of conditional quantiles of XA conditional on XF_C

$L_{history}$ = number of historical time steps

L_{future} = number of time steps to simulate

procedure

Grouping values of $XA_{hist.}$

for all $t = 1, 2, ..., L_{history}$ **do**

$XF_C(t) \leftarrow \sum_{j=1}^{n_{XF_C}-1} I_{l_{XF_C}(j)}(XF_{hist.}(t))$ ▷ Class of forecast

$Q_{XA}(t) \leftarrow \sum_{k=1}^{q-1} I_{l_{XA_{XF_C(t)}}(k)}(XA_{hist.}(t))$ ▷ Conditional quantile

$G_{XA_{XF_C(t),Q_{XA}(t)}} \leftarrow [G_{XA_{XF_C(t),Q_{XA}(t)}}, XA_{hist.}(t)]$ ▷ Actual value saved to right group

end for

Estimating empirical transition probabilities between groups

$P(Q_{XA}(t)|Q_{XA}(t-1)) \leftarrow \frac{n(Q_{XA}(t) \cap Q_{XA}(t-1))}{n(Q_{XA}(t-1))}$

Simulating $XA_{fut.}$

for $t = 1$ **do**

Sample group $Q_{XA}(t)$ with equal probabilities for all q groups

Sample $XA(t)$ from $Q_{XA}(t)$ with equal probabilities

end for

for all $t = 2, 3, ..., L_{future}$ **do**

Sample group $Q_{XA}(t)$ with probabilities $P(Q_{XA}(t)|Q_{XA}(t-1))$

$XF_C(t) \leftarrow \sum_{j=1}^{n_{XF_C}-1} I_{l_{XF_C}(j)}(XF_{fut.}(t))$

$XA_{fut.}(t)$ sampled from $G_{XA_{XF_C(t),Q_{XA}(t)}}$ with equal probabilities

end for

end procedure

4.2.2.1 Wind generation

New wind generation units are often built in same windy regions as the existing ones. Therefore, their output fluctuations are highly similar as wind speed changes. The difference between forecast and actual output of one wind generating unit is strongly correlated with another that is situated in the same region. Therefore, output powers of units can be treated as strongly positively correlated random variables. Thus, it can be assumed that increasing wind generation capacity will result in forecast errors increasing linearly in the same proportion.

To account for this linear impact of new-build generation capacity on total output and forecast errors we divide historical $IGA_{hist.}$ and $IGF_{hist.}$ values by historical system-wide installed wind generation capacity of each time step. We obtain forecast and actual wind generation of each time step as percentages $IGF_{hist.} - \%$ and $IGA_{hist.} - \%$, respectively, of installed capacity. We also simulate future day-ahead forecast by DA market modelling software as percentage of installed capacity $IGF_{fut.} - \%$. Now the simulation algorithm for actual values used with inputs $XF_{hist.} = IGF_{hist.} - \%$, $XA_{hist.} = IGA_{hist.} - \%$ and $XF_{fut.} = IGF_{fut.} - \%$ gives also actual wind generation time series as percentage of installed capacity $IGA_{fut.} - \%$. To transform it into MW-value we can simply multiply by an estimate of future installed capacity. This provides also a natural way to study the impact of increasing or decreasing installed wind generation capacity on RT price. We can create scenarios with same $IGA_{fut.} - \%$, but different installed capacity and compare simulated RT price time series.

4.2.2.2 Available non-intermittent generation capacity

In case of available non-intermittent generation capacity forecast errors are mainly due to changes in plans of generating companies. They may be a result of e.g. unexpected outages or changes in near-term demand forecasts that are used as inputs for decision making of power plant operation (Rhodri Williams, Regional director - ERCOT, Genscape, 10 February 2016, Boston MA). We assume that the forecast errors of non-intermittent generation capacity will be linearly dependent on the total installed capacity of non-intermittent generating units. Therefore we can use a similar methodology to create future time series of actual available non-intermittent generation capacity $NICA_{fut.}$ as we used for actual wind generation $IGA_{fut.}$. Obviously, values are not calculated as percentages of installed wind generation capacity, but of installed non-intermittent generation capacity. Data for the simulation algorithm is then $XF_{hist.} = NICF_{hist.} - \%$, $XA_{hist.} = NICA_{hist.} - \%$

and $XF_{fut.} = NICF_{fut.} - \%$. The output time series $NICA_{fut.} - \%$ is actual available non-intermittent generation capacity time series as percentage of installed capacity. It can be transformed into absolute MW-value in the same way as in case of wind generation by multiplying by an estimate of future installed capacity. We can also study the impact of increasing or decreasing installed non-intermittent generation capacity on RT price by creating scenarios with same $NICA_{fut.} - \%$, but different installed capacity and compare simulated RT price time series.

4.2.2.3 Demand

Since forecast errors of demand mainly result from errors in weather forecasts, we can assume that forecast errors will be linearly dependent on annual peak load, i.e. the highest demand of a year. We can use the same method to create future series of actual demand, $LA_{fut.}$, that we use for actual wind generation $IGA_{fut.}$ and available non-intermittent generation capacity $NICA_{fut.}$. Input values are calculated as percentages of annual peak load. Data for the simulation algorithm is then $XF_{hist.} = LF_{hist.} - \%$, $XA_{hist.} = LA_{hist.} - \%$ and $XF_{fut.} = LF_{fut.} - \%$. The output series $LA_{fut.} - \%$ is actual demand time series as percentage of annual peak load and it can be transformed into absolute MW-value by multiplying by an estimate of annual future peak load. Again, we can study the impact of increasing or decreasing demand on RT price by creating scenarios with same $LA_{fut.} - \%$, but different peak load growth rate, and compare simulated RT price time series.

4.3 Stochastic simulation methodology for future RT price

We use a stochastic simulation methodology to create simulated future RT price time series. First we simulate prices for future spike periods, and then we simulate prices of rest of the time steps. We use groups G_{RT_C} and G_{Δ_C} corresponding to RT and Δ values of each price class and delta class from simulation 1 as the populations from which future RT prices and Δ values are sampled. Price class or delta class of each future time step is selected randomly using estimated conditional probabilities of classes conditional on values of explanatory variables and the price class or delta class of previous time step. Therefore, the simulation model could be classified as a dynamic bootstrap method.

4.3.1 Simulating RT prices of future spike periods

We start by sampling a binary variable X_{start} for each future time step indicating whether a spike period starts on that time step. Sampling is performed with conditional probabilities determined in 1. Then we sample duration of each spike period, L_{spike}, from a group of historical spike lengths G_L. The prices for spike periods are obtained using block sampling method. Instead of sampling each price separately, we sample as many blocks of length L_B with replacement until all prices of the spike period are sampled. This method helps to get the autocorrelation of sampled time series as strong as in history. A small correction to keep the distribution of prices of spike periods same as historical is done in block sampling by using wrapping method. In this method a block can start near the end of and "continue" in the beginning of vector $G_{RT_{spike}}$.

Algorithm 3 Spike simulation

Data

Simulated processes of explanatory variables

DA_t = Day-ahead price

IGA_t = Intermittent generation, actual

$NICA_t$ = Non-intermittent generation capacity, actual

LA_t = Demand, actual

Sampling populations

$G_{RT_{spike}}$ = RT prices of historical spike periods in chronological order

G_L = Lengths of historical spike periods

Estimated probabilities

$P(X_{start}|I_{DA_{limit,prob}}(DA), I_{FCA_{limit,prob}}(FCA))$ Estimated probabilities of spike starting conditional on DA and FCA

Parameters

L_B = Price block length

$DA_{limit,prob}$ = Division point of DA for conditional probability of spike starting

$FCA_{limit,prob}$ = Division point of FCA for conditional probability of spike starting

L_{future} = number of future time steps to simulate

procedure SIMULATION

$t \leftarrow 0$

while $t \leq L_{future}$ **do**

$FCA(t) \leftarrow NICA(t) + IGA(t) - LA(t)$

$X_{start}(t)$ chosen randomly with $P(X_{start}|I_{DA_{limit,prob}}(DA), I_{FCA_{limit,prob}}(FCA))$

if $X_{start} = 1$ **then**

L_{spike} sampled from G_L

$RT_{temp} \leftarrow L_{spike}$ values chosen randomly from $G_{RT_{spike}}$ using block sampling method with block length L_B with wrapping

for all $\tau = t, t+1, ..., t+L_{spike}-1$ **do**

$X_{spike}(\tau) \leftarrow 1$

$RT(\tau) = RT_{temp}(\tau - t + 1)$

end for

$t \leftarrow t + L_{spike}$

else

$t \leftarrow t + 1$

end if

end while

end procedure

4.3.2 Simulating other RT prices than spike periods

Now that we have simulated prices of future spike periods, we can simulate prices of non-spike time steps. Dynamic time step category C is determined first based on values of explanatory variables. Price class, if $DA \geq DA_{limits,C}$, or delta class, otherwise, is sampled using class probabilities conditional on C. Then RT or Δ is sampled from chosen class with equal probabilities. Simulation method for non-spike time steps is shown in algorithm 4.

Algorithm 4 Non-spike simulation

Data

Simulated processes of explanatory variables

DA_t = Day-ahead price

IGA_t = Intermittent generation, actual

$NICA_t$ = Non-intermittent generation capacity, actual

LA_t = Load, actual

IGF_t = Intermittent generation, forecast

$NICF_t$ = Non-intermittent generation capacity, forecast

LF_t = Load, forecast

X_{spike_t} = Binary vector telling whether each time step is spike or not

Sampling populations

$G_{\Delta_C} = \{\Delta_{hist} | \Delta_{hist} \in \Delta_C\}$ = Historical Δ values of delta class $\Delta_C \in \{0, 1, 2, ..., n_{\Delta_C} - 1\}$

$G_{RT_C} = \{RT_{hist} | RT_{hist} \in RT_C\}$ = Historical RT prices of price class $RT_C \in \{0, 1, 2, ..., n_{RT_C} - 1\}$

$G_{RT_{spike}} = [RT_{hist} | X_{spike,hist} = 1]$ = RT prices of historical spike periods in chronological order

Estimated probabilities

$P(\Delta_C|C)$ Probabilities of delta classes $\Delta_C \in \{0, 1, 2, ..., n_{\Delta_C} - 1\}$ conditional on dynamic time step category C

$P(RT_C|C)$ Probabilities of price classes $RT_C \in \{0, 1, 2, ..., n_{RT_C} - 1\}$ conditional on dynamic time step category C

Parameters

$DA_{limits,C}$ = vector that contains limits of DA for determining C

$FCFE_{limit}$ = limit of $FCFE$ for determining C

NLC_{limit} = limit of NLC for determining C

n_{Δ_C} = number of delta classes

n_{RT_C} = number of price classes

procedure SIMULATION
for $t = 0$ **do**
$\Delta_C(0)$ chosen randomly with equal probabilities
$\Delta(0)$ chosen randomly from $G_{\Delta_C(0)}$ with equal probabilities
end for
for all $t > 0$ **do**
$C^0(t) = 4[I_{DA_{limits,C}(1)}(DA(t)) + I_{DA_{limits,C}(2)}(DA(t))] + 2I_{FCFE_{limit}}(FCFE(t)) + I_{NLC_{limit}}(NLC(t))$
if $X_{spike}(t) = 0$ **then**
if $X_{spike}(t-1) = 1$ **then** ▷ Previous was spike
$C(t) = C^0(t) + 12(n_{\Delta_C} + n_{RT_C})$
if $DA(t) \geq DA_{limits,C}(1)$ **then**
$RT_C(t)$ chosen randomly with probabilities $P(RT_C|C(t))$
$RT(t)$ chosen randomly with equal probabilities from $G_{RT_C(t)}$
else
$\Delta_C(t)$ chosen randomly with probabilities $P(\Delta_C|C(t))$
$\Delta(t)$ chosen randomly with equal probabilities from $G_{\Delta_C(t)}$
$RT(t) = DA(t) + \Delta(t)$
end if
else ▷ Previous was not spike
$C(t) = C^0(t) + 12\Big[\Delta_C(t-1) + I_{DA_{limits,C}(1)}(DA(t-1))[n_{\Delta_C} + RT_C(t-1) - \Delta_C(t-1)]\Big]$
if $DA(t) \geq DA_{limits,C}(1)$ **then**
$RT_C(t)$ chosen randomly with probabilities $P(RT_C|C(t))$
$RT(t)$ chosen randomly with equal probabilities from $G_{RT_C(t)}$
else
$\Delta_C(t)$ chosen randomly with probabilities $P(\Delta_C|C(t))$
$\Delta(t)$ chosen randomly with equal probabilities from $G_{\Delta_C(t)}$
$RT(t) = DA(t) + \Delta(t)$
end if
end if
end if
end for
end procedure

4.4 Validating model functioning by comparing historical and simulated values

4.4.1 Actual values of explanatory variables

We can create simulated time series of actual values of explanatory variables LA, IGA, and $NICA$ using historical forecast values of LF, IGF, and $NICF$ as inputs. If time series simulated in this way are similar to their historical correspondents, we can conclude that our simulation methodology can be used to create simulated future scenarios based on forecast values simulated by commercial DA market modelling software. We conduct three separate simulations, one for each variable, with same parameters in each case. We divide the historical data to $n_{XF_C} = 50$ intervals each 2% wide compared to maximum value of forecast. We divide actual values to $q = 10$ groups, i.e. conditional deciles of each 2 % interval of forecast. Example weeks of all three variables are shown in figures 4.2, 4.3, and 4.4. There are differences between historical and simulated actual values. However, the differences are so small that they can be supposed to be a result of random variation that is natural and desired in stochastic simulation.

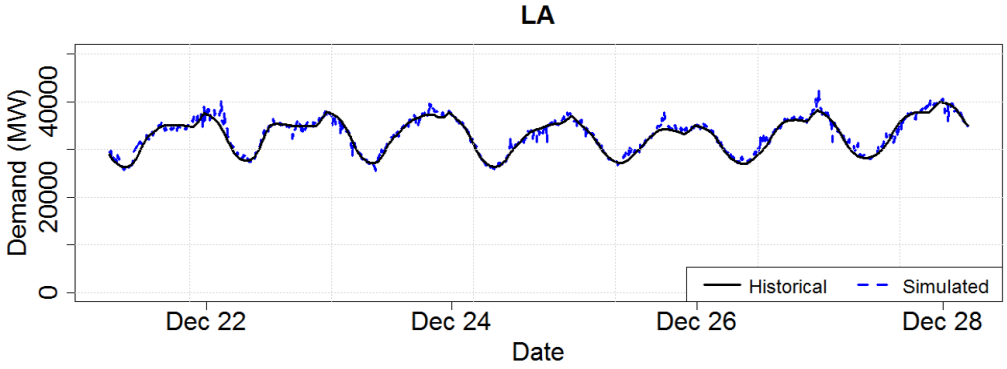

Figure 4.2: Black solid line shows actual historical load of Christmas week 2015. Blue dashed line shows actual load simulated using historical inputs. Simulated and historical demand are almost exactly equal.

Historical and simulated duration curves for each variable are shown in figure 4.5. They coincide almost perfectly, meaning that distribution of simulated values is highly similar to that of historical.

Another measure for performance of our wind generation simulation is capacity factor, i.e. mean of wind generation as percentage of nameplate

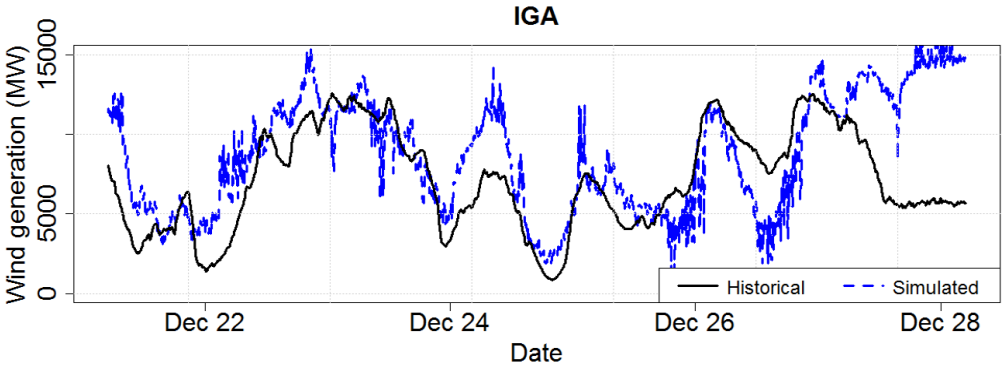

Figure 4.3: Example week of actual wind generation. Black solid line represents historical actual wind generation of Christmas week 2015. Blue line represents actual values simulated with historical inputs. The movements of the two lines are highly similar. Perfect matching is not desired in stochastic simulation.

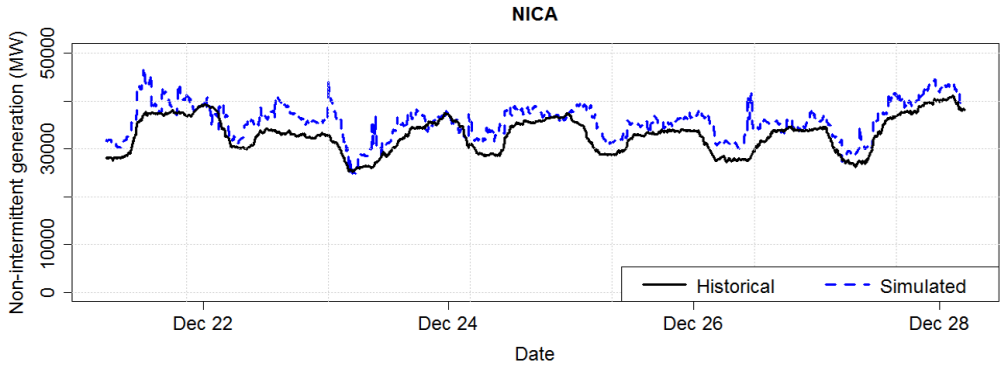

Figure 4.4: Example week of actual non-intermittent generation capacity. Black line represents actual historical values of Christmas week 2015 and blue line represents simulated actual values with inputs of year 2015.

capacity. Historical capacity factor of years 2011-2015 is 32 %. Histogram of capacity factors of 300 simulation runs in figure 4.6 shows that they are generally almost exactly equal to historical.

Analysis of explanatory variables in Chapter 3 showed that they all exhibit strong autocorrelation. Top row of figure 4.7 shows empirical autocorrelation functions of historical and bottom row of simulated actual values. All

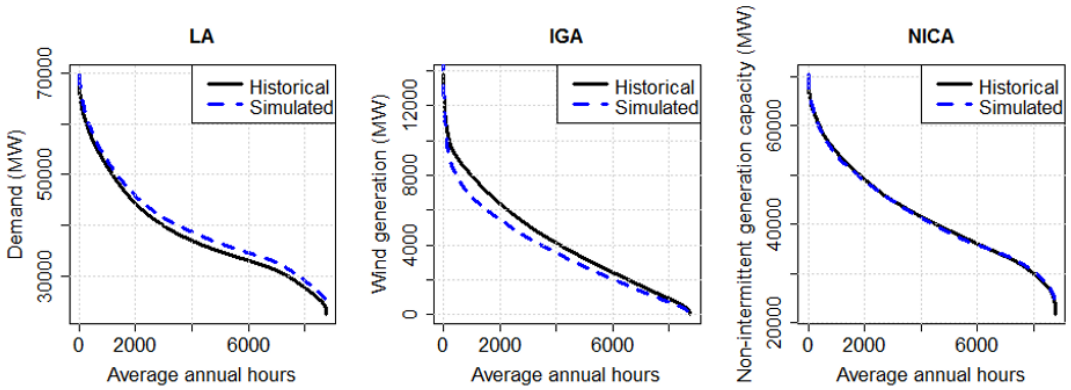

Figure 4.5: Duration curves of explanatory variables. Black solid lines represent actual historical values of 2015 and blue dashed lines represent simulated actual values with inputs of year 2015. Simulated duration curves are very close to their historical correspondents.

three simulated variables exhibit similar simple and seasonal autocorrelation as historical.

The three graphs on left side of figure 4.8 show scatter plot of historical forecast and actual values of load, intermittent generation, and available on-line non-intermittent generation capacity. Graphs on the right side show similar graphs of simulated values. Simulated time series of all three variables exhibit similar strong positive correlation between forecast and actual value as historical series do.

Comparisons between historical and simulated actual values of demand, wind generation, and non-intermittent generation capacity show very close resemblance. Based on analysis, we can conclude that our stochastic simulation methodology for actual values of explanatory variables is sufficient. Next we need to confirm that other parts of simulation method are valid, too. After that we can create simulated RT price scenarios.

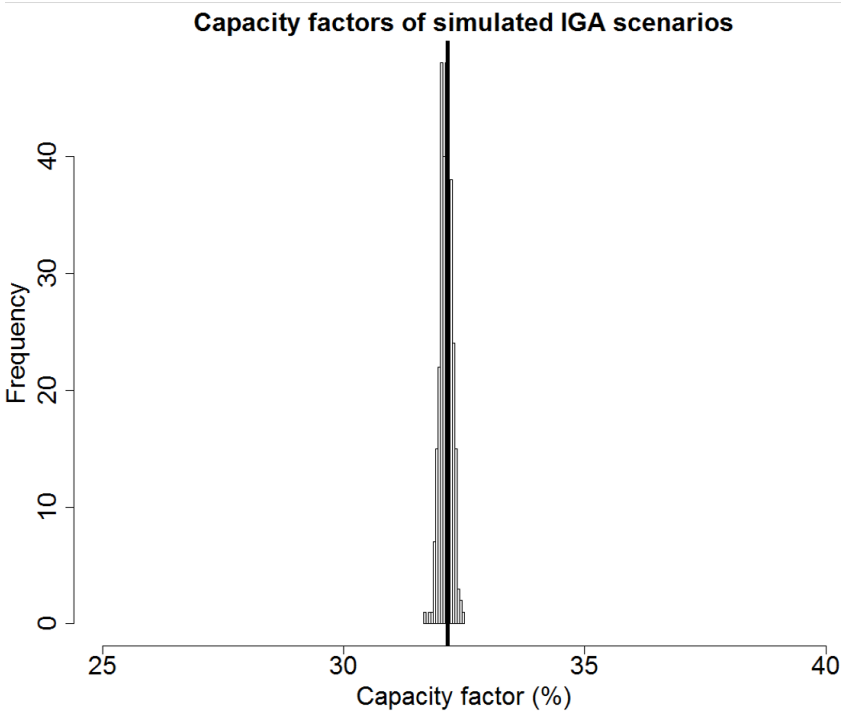

Figure 4.6: Histogram of capacity factors of 300 simulation runs of actual wind generation. Black vertical line represents historical capacity factor 32 %. Simulated capacity factors are close to historical. Small differences are desired and result from randomness in simulation.

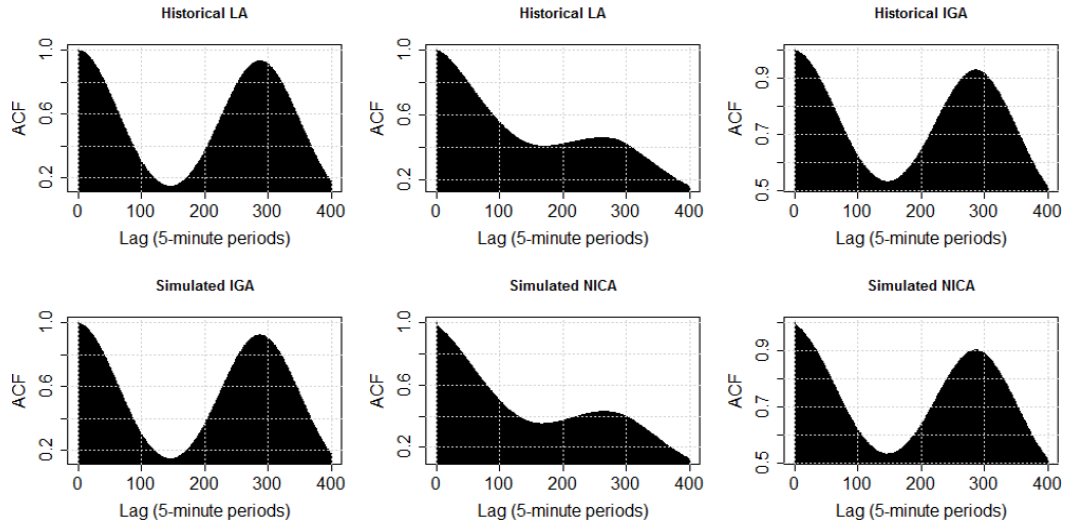

Figure 4.7: Empirical autocorrelation functions of historical (top) and simulated (bottom) LA, IGA, and $NICA$. Simulate series imitate autocorrelation structures of their historical correspondents almost perfectly.

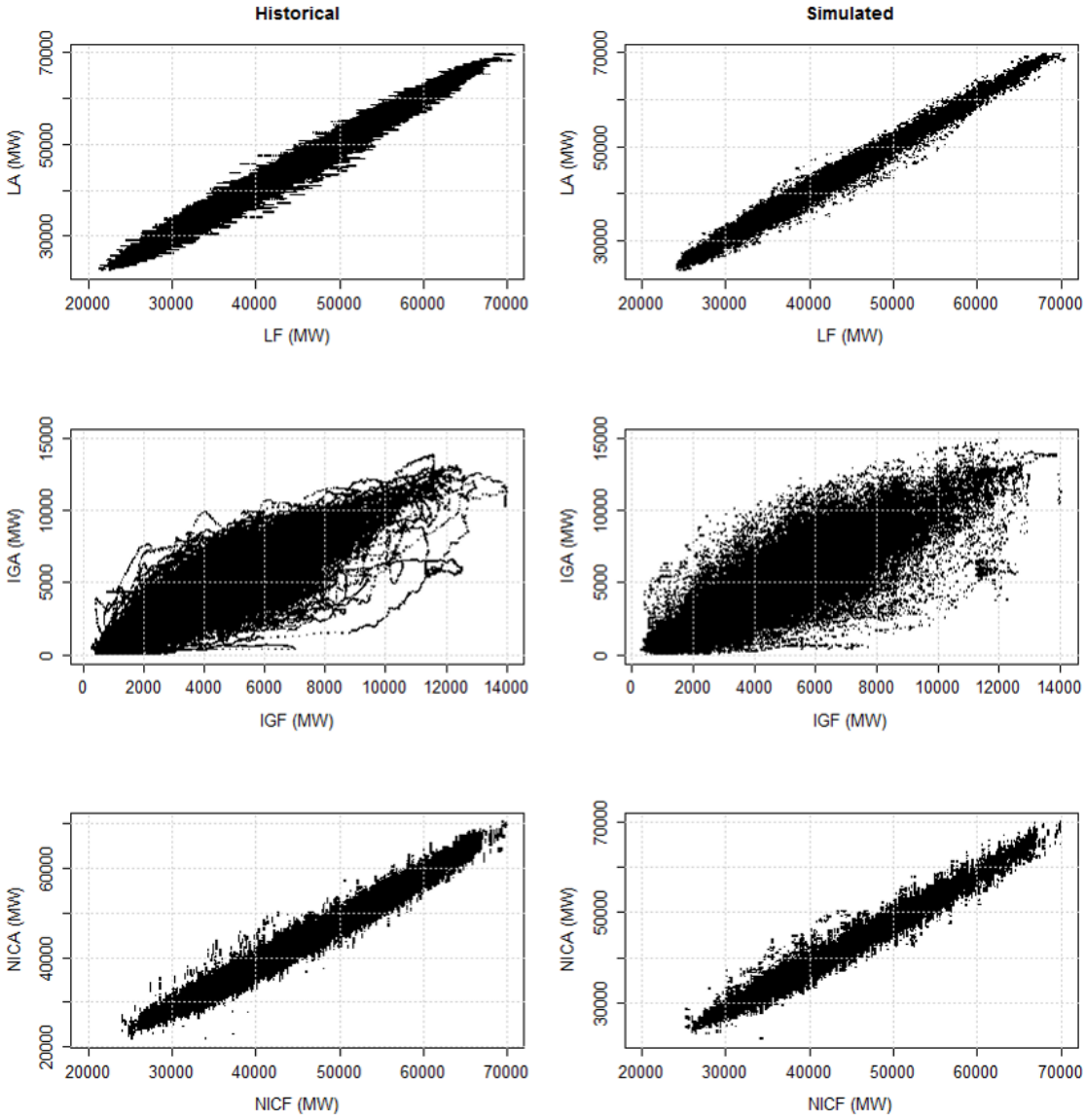

Figure 4.8: Scatter plots of DA-forecast and actual values of historical (left) and simulated (right) demand, LF and LA, wind generation, IGF and IGA, and available non-intermittent generation capacity, $NICF$ and $NICA$. Conditional distributions of actual values as functions of forecasts closely resemble those of historical values.

Table 4.1: Parameters used in algorithm 1

$l_{spike,RT}$	93.19 USD/MWh
$l_{spike,DA}$	300 USD/MWh
$DA_{limit,prob}$	30 USD/MWh
$FCA_{limit,prob}$	7 GW
S_{spike}	0.5
l_{RT_C}	$[16.7, 19.1, 21.3, 22.9, 24.5, 26.2, 28.8, 32.6, 39.8]$
l_{Δ_C}	$[-13.7, -8.6, -6.0, -4.2, -2.7, -1.5, -0.4, 0.7, 2.4]$
$DA_{limits,C}$	$[25, 40]$
$FCFE_{limit}$	-1800 MW
NLC_{limit}	200 MW
$L_{history}$	525600

4.4.2 RT price

To ensure that the model described above can be used to create future RT price time series, we apply it with historical input data and compare the output series with actual historical occurred series. If simulated RT price time series resemble historical price, we can conclude that the model functions as expected. All market data that we use was collected by ERCOT and processed and delivered to us by Genscape.

Historical data that we use for simulation includes historical time series of DA and RT, explanatory variables LA, LF, IGA, IGF, $NICA$, $NICF$, and time series of installed wind generation capacity and installed non-intermittent generation capacity from the years 2011-2015. Individual missing data points and periods of missing data with length less than or equal to two hours are filled by linearly interpolating. In case of time series of dependent variables DA and RT interpolation is not used. As changes in the values of explanatory variables from time step to time step are often relatively small, linear interpolation is a safe choice for short periods of missing data. For large periods of missing data, interpolation is not used. Graphs of input variables as a function of time are shown in chapter 3. All parameters used in this simulation are shown in table 4.1.

Algorithm 1 run with inputs and parameters as specified, gives us the empirical conditional probabilities $P(\Delta_C|C)$ of delta classes and price classes as functions of dynamic time step category C. Explanations of all dynamic time step categories C in case of 12 price classes and delta classes are shown in table 6.1 in appendix. An example of conditional probabilities of delta classes are shown in table 4.3 and conditional probabilities of price classes are shown in table 4.4. Conditional spike starting probabilities are shown in

Table 4.2: Conditional probabilities of price spike at different values of explanatory variables DA price DA and actual surplus capacity FCA. All probabilities are rather low, but at $FCFE$ values above 7000 MW (right column) probabilities are a lot lower.

	$FCFE$ (MW)	
DA (USD/MWh)	< 7000	≥ 7000
< 30	0.002	0.0001
≥ 30	0.006	0

table 4.2.

Now we have determined all needed quantities from historical data and can start to simulate future RT price. In this validation we do not carry out DA market simulation or simulation of actual values of explanatory variables to obtain simulated series of DA, LA, LF, IGA, IGF, $NICA$, and $NICF$ since we do not want randomness in these simulations to impact our simulated RT price. Therefore, we use historical values of explanatory variables as inputs for RT price simulation. In addition we use grouped history data G_{Δ_C}, G_{RT_C}, $G_{RT_{spike}}$, G_L and conditional probabilities $P(X_{start}|I_{DA_{limit,prob}}(DA), I_{FCA_{limit,prob}}(FCA))$, $P(\Delta_C|C)$, and $P(RT_C|C)$ output in the previous step by algorithm 1 for algorithms 3 and 4. The number of time steps to simulate, L_{future}, is 1051200, i.e. 10 years. Parameter values used for spike simulation are shown in table 4.5 and values used for simulation of non-spike RT prices are shown in table 4.6.

One simulated week of RT price and the corresponding historical RT price are shown in figure 4.9. Behaviour of RT prices is highly similar. It is clear that the spike periods do not coincide and it is not desirable in our stochastic simulation approach. There are a few price spikes in both historical and simulated RT price. Spikes occur mainly when DA price is close to its daily maximum. There are also short periods when RT price is close to zero. In this example week such situations only occur at nights when demand is low and stable. Our simulated RT price seems to preserve this quality well. Visual inspection does not show large differences in autocorrelation between the price series.

Median of historical RT price is 25.23 and median prices of 100 simulation runs are shown in histogram in figure 4.10. Medians are very near to each other, again indicating that simulated price series resemble the historical price series. It is not desirable in our stochastic approach that median price would be precisely same.

Duration curves of historical and simulated RT price are shown in figure 4.11. They coincide almost perfectly. Small differences are natural and de-

Table 4.3: Example of estimated conditional probabilities of delta classes $P(\Delta_C = i|C)$ conditional on dynamic time step categories C 0-7 and 12-19. Majority of dynamic time step categories is excluded from this example.

	delta class i									
C	0	1	2	3	4	5	6	7	8	9
0	0.43	0.09	0.01	0.03	0.06	0.1	0.07	0.09	0.07	0.04
1	0.28	0.04	0.03	0.04	0.15	0.1	0.09	0.1	0.04	0.1
2	0.64	0.09	0.02	0.03	0.05	0.04	0.05	0.02	0.03	0.03
3	0.23	0.06	0.03	0.09	0.09	0.14	0.11	0.03	0.23	0
4	0.48	0.1	0.07	0.06	0.07	0.06	0.03	0.03	0.03	0.06
5	0.54	0.14	0.08	0.04	0.02	0.04	0.04	0.01	0.01	0.08
6	0.55	0.07	0.09	0.07	0.05	0.03	0.03	0.03	0.03	0.06
7	0.52	0.12	0.04	0.09	0	0.06	0.01	0.04	0.04	0.09
12	0.07	0.64	0.1	0.05	0.02	0.04	0.02	0.02	0.02	0.02
13	0.1	0.8	0.1	0	0	0	0	0	0	0
14	0.11	0.68	0.09	0.03	0.03	0.02	0.02	0.01	0.01	0
15	0.09	0.45	0.14	0.05	0.05	0.14	0.09	0	0	0
16	0.02	0.86	0.09	0.02	0.01	0	0	0	0	0
17	0.01	0.87	0.09	0.01	0	0	0.01	0	0	0
18	0.02	0.85	0.08	0.03	0.01	0.01	0	0	0	0
19	0.03	0.81	0.16	0	0	0	0	0	0	0

Table 4.4: Example of estimated conditional probabilities of price classes $P(RT_C = i|C)$ conditional on dynamic time step category $C \in \{8, 9, 10, 11, 20, 21, 22, 23, 32, 33, 34, 35, 44, 45, 46, 47, 56, 57, 58, 59\}$. Majority of dynamic time step categories is excluded from this example.

	price class i									
C	0	1	2	3	4	5	6	7	8	9
8	0	0	0	0	0.03	0.02	0.03	0.22	0.33	0.37
9	0	0.05	0	0	0.15	0.05	0.1	0.05	0.3	0.3
10	0	0.01	0.01	0.01	0.02	0.02	0.04	0.08	0.29	0.54
11	0	0	0	0.08	0.08	0	0	0.15	0.23	0.46
20	0	0.02	0	0	0.15	0.26	0.39	0.17	0	0
21	0	0	0	0.14	0	0	0.43	0.43	0	0
22	0	0	0.03	0.05	0.16	0.08	0.51	0.16	0	0
23	0	0	0	0	0.5	0	0	0.5	0	0
32	0	0	0	0.04	0.04	0.07	0.21	0.61	0.04	0
33	0	0	0	0	0.06	0.06	0.22	0.5	0.17	0
34	0	0	0	0	0.04	0.2	0.28	0.44	0.04	0
35	0	0	0	0	0	0	0.14	0.43	0.43	0

Table 4.5: Parameters used in spike simulation.

$DA_{limit,prob}$	30 USD/MWh
$FCA_{limit,prob}$	7 GW

Table 4.6: Parameters used in simulation of non-spike RT prices.

$DA_{limits,C}$	[25, 40] USD/MWh
$FCFE_{limit}$	-1800 MW
NLC_{limit}	200 MW
n_{Δ_C}	10
n_{RT_C}	10

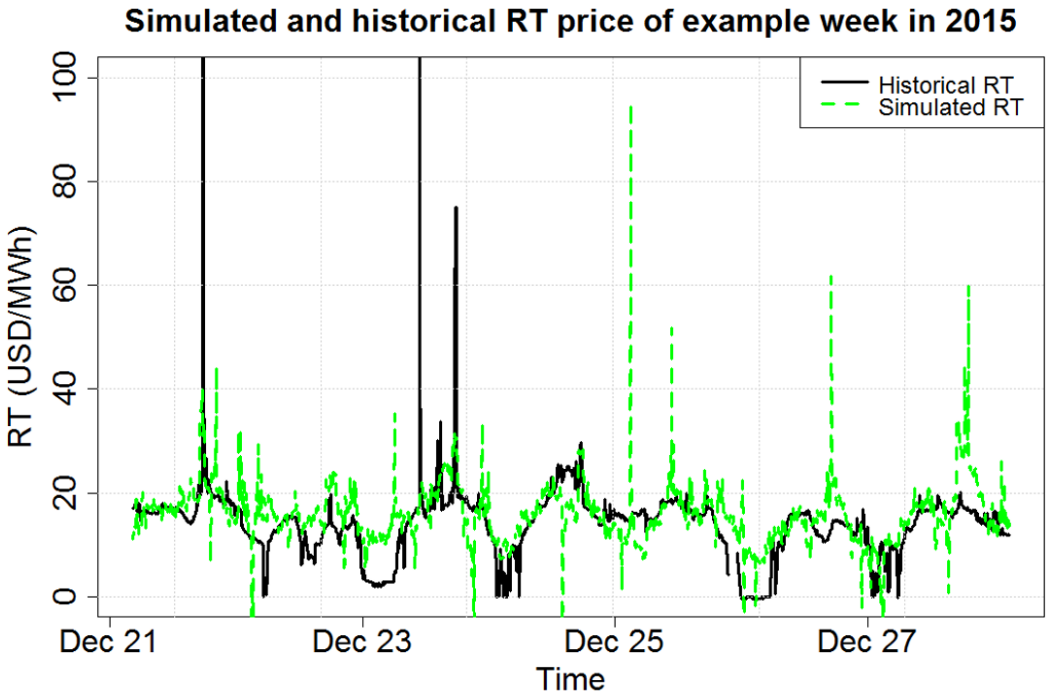

Figure 4.9: Example week. Movements of historical (black solid line) and simulated (blue dashed line) RT price are highly similar. They do not coincide perfectly and, in fact, it is not desirable in our stochastic simulation approach.

sired result of randomness in simulation. Distribution of simulated RT price matches very closely with that of historical RT price.

Empirical autocorrelation functions of both RT price time series are shown in figure 4.12. It can be seen from the two graphs on the left that the strong seasonal autocorrelation of historical RT price on lags approximately 288 (24 hours) is not exhibited by simulated series. However, the graphs on the right show that autocorrelation of simulated series on short lags resembles that of historical price, being only slightly weaker on lags 1-10. For the purpose of profitability calculation for a flexible power plant it does not have any impact that the long lag autocorrelation is not preserved by our simulation. Fast-ramping power plants can change their operating plans many times within a day and they are not dependent on events of previous or next day.

From all graphs and quantities compared above we can conclude that our simulation methodology can create simulated time series of RT price highly similar to historical series in all qualities and quantities that are important

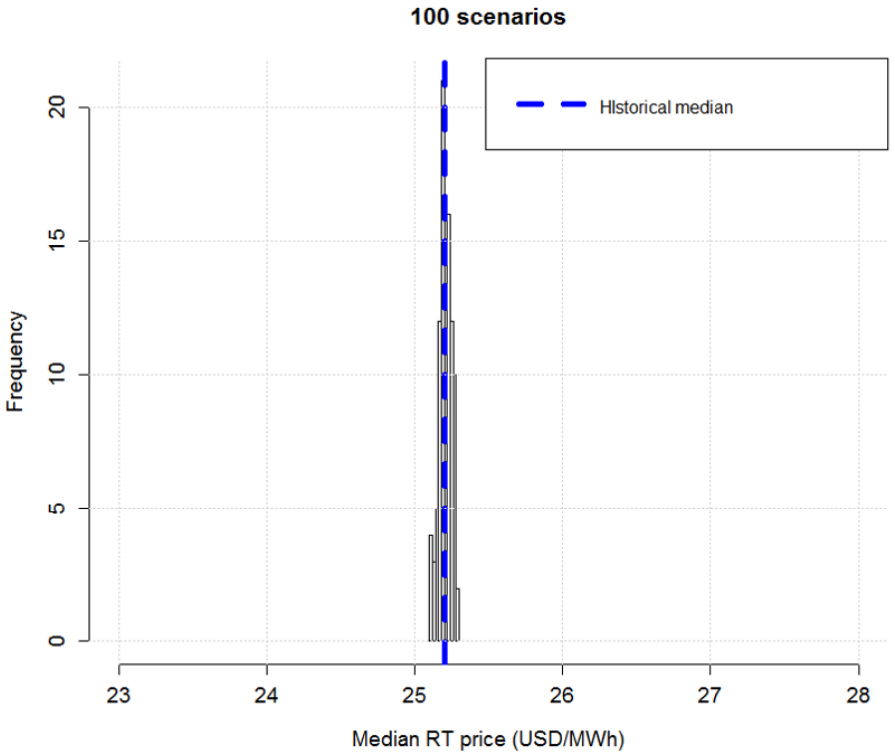

Figure 4.10: Histogram of median prices from 100 simulated scenarios using historical values of input variables. Median prices are generally close to historical median price 25.23 USD/MWh.

for the purpose of our simulation. Next, we will use the simulation method to study the impact of increasing relative share of wind generation capacity as a percentage of demand on RT price in ERCOT.

4.5 Case studies

Next we will use our forecasting model to create simulated future RT price time series. We consider two cases: (i) base scenario, where all important quantities develop as predicted by ERCOT or stay stable, and (ii) growth scenario where installed wind generation capacity and annual peak load grow slightly faster. We run 100 stochastic simulations of each scenario using exactly same historical inputs, RT and DA prices and explanatory variables. We only vary the future values of installed wind generation capacity and

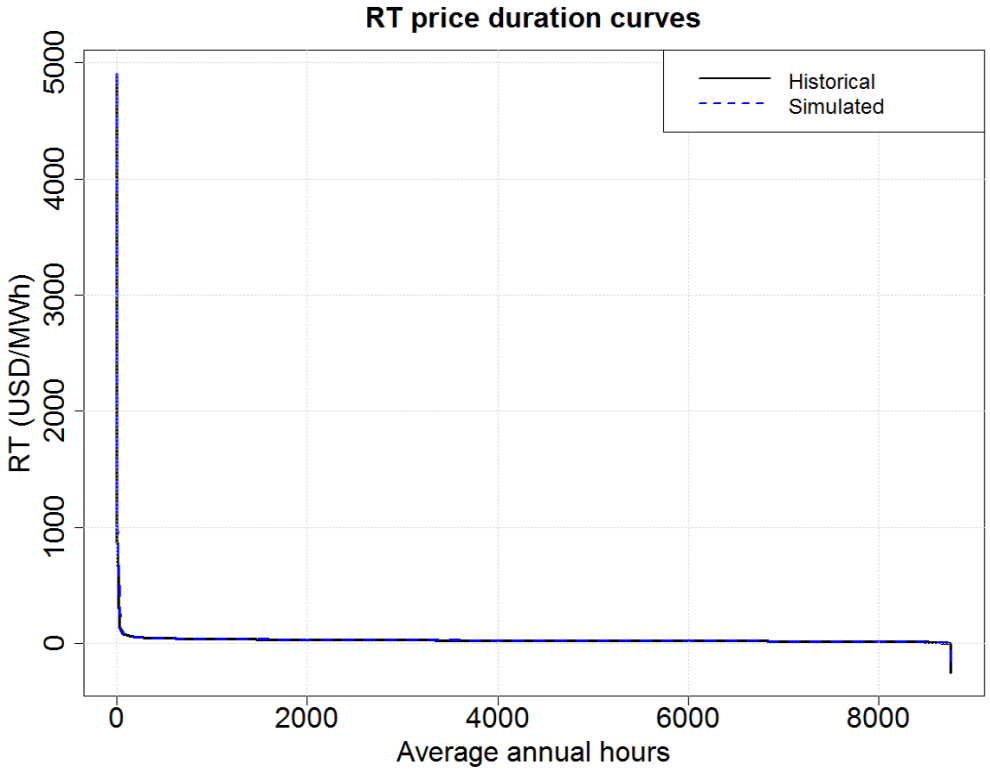

Figure 4.11: Duration curves of historical (black solid line) and simulated (blue dashed line) RT price match almost perfectly.

annual peak load. Installed non-intermittent generation capacity is assumed to develop similarly in both scenarios. We compare our simulation results of both scenarios to reveal any differences in RT price behaviour caused by differences in input variables.

4.5.1 Base scenario and growth scenario

Our historical explanatory variable values consist of ERCOT market data 2012-2015 published by ERCOT, processed and delivered to us by Genscape. Data includes historical time series of DA and RT, explanatory variables LA, LF, IGA, IGF, $NICA$, $NICF$, and time series of installed wind generation capacity and installed non-intermittent capacity. Individual missing data points and periods of missing data with length less than or equal to two hours are filled by linearly interpolating, except in case of time series of dependent variables DA and RT. Interpolation method is same that we used earlier

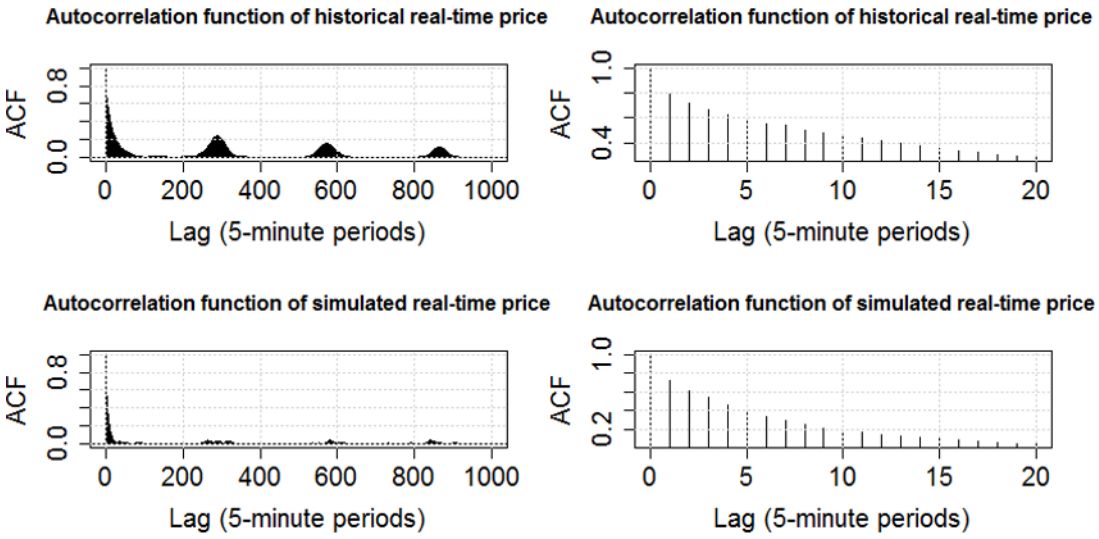

Figure 4.12: Both historical (top) and simulated (bottom) RT price exhibit strong autocorrelation on short lags (right). Simulated price, however, does not exhibit as strong seasonal autocorrelation (left) on lags approximately 288 time steps (24 hours).

in validation of model functioning. As representations of RT and DA prices we use hub average price, which is, as explained earlier in chapter 3, simple average of all hub prices in ERCOT. Parameters used are same as those used for validation runs shown in tables 4.1, 4.5, and 4.6.

Possible differences between scenarios can only result from differences in simulation of future values of explanatory variables LA, LF, IGA, IGF. We consider two scenarios that slightly differ from each other. As explained earlier, our simulation methodology for future values of explanatory variables enables us to create different scenarios of demand and wind generation development. We simply change the time series of peak load and installed wind generation capacity that we use to multiply the simulated values of wind generation as a percentage of installed capacity and demand as a percentage of annual peak load.

In base scenario we use stable installed wind generation capacity as input of future development. Installed wind generation capacity grows from 18 GW in 2016 to 23 GW in 2017 and stays stable thereafter. In growth scenario, capacity grows 8 % per year faster. Growth speed is still slow compared to 12 % geometric mean annual growth in 2011-2015. Difference between capacity

in the two scenarios grows gradually from zero in 2016 to 28 GW in 2025. Growth scenario is more realistic in installed wind generation capacity growth (8 % vs. 0 %, historical 12 %) but it is still slightly conservative estimate. However, the difference between scenarios is important in our simulation, not absolute values. Graphs of development of installed wind generation capacity in both scenarios are shown in figure 4.13.

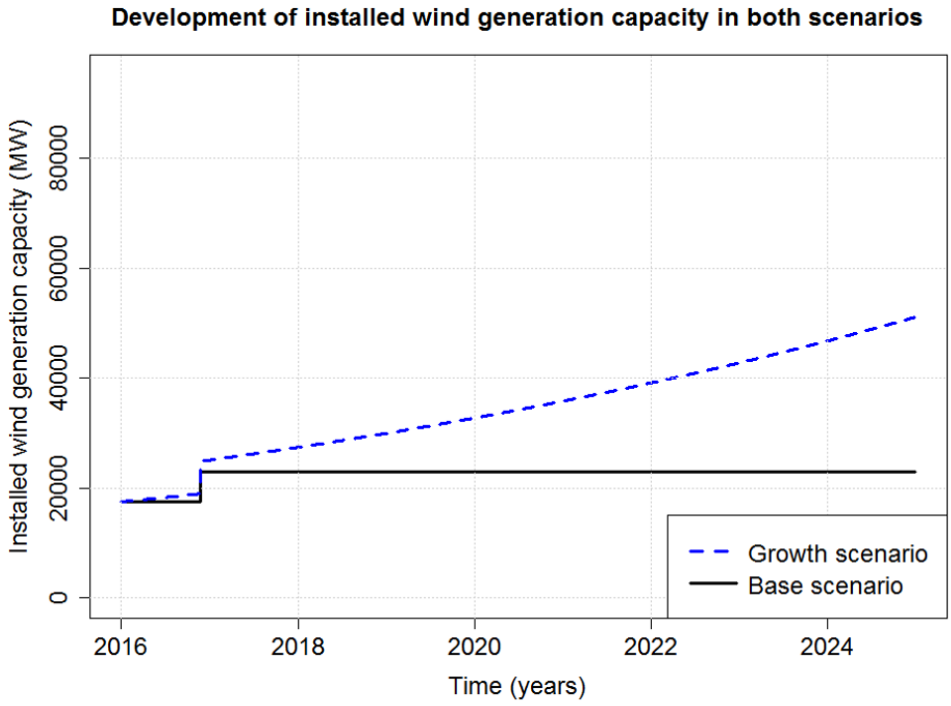

Figure 4.13: Installed wind generation capacity grows faster in growth scenario (dashed blue line) than in base scenario (solid black line). Difference grows from zero in 2016 to 28 GW in 2025.

ERCOT estimates that peak load is 70 GW in 2016 and grows thereafter steadily to 78 GW in 2025. We use this estimate in our base scenario. The geometric mean of yearly peak load growth will be 1.2 % per year in 2016-2025 in base scenario. We modify that number to 3 % in growth scenario, which means that peak load in 2025 is 91 GW. Graphs of development of peak load in both scenarios are shown in figure 4.14. It can be seen that growth of peak load is stronger in growth scenario than in base scenario. Difference grows gradually from zero in the year 2016 to 13 GW in 2025.

Growth scenario has both demand and installed wind generation capac-

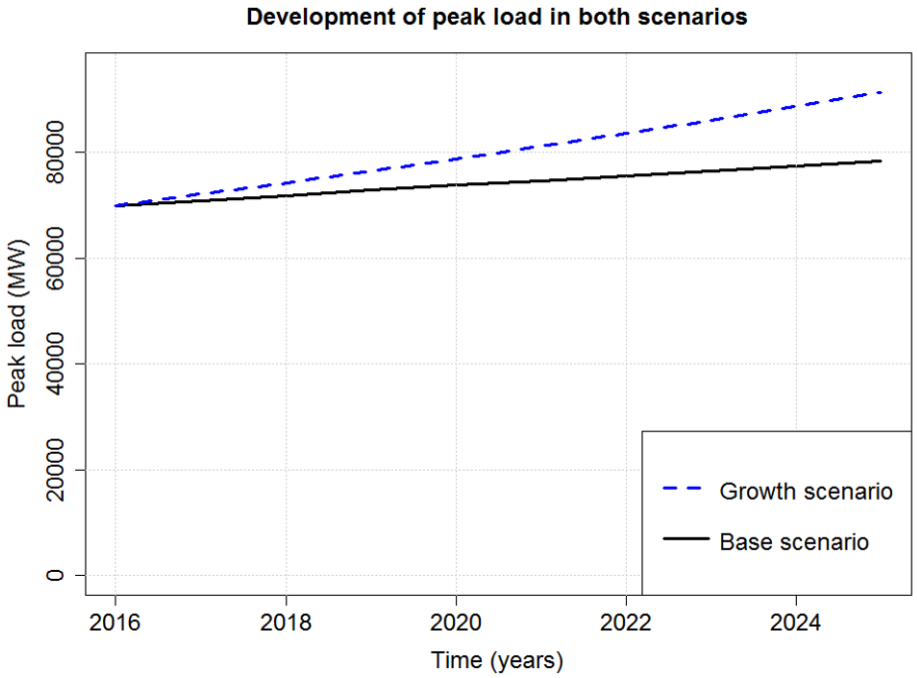

Figure 4.14: Peak load growth in growth scenario (dashed blue line) is faster than in base scenario (solid black line). Peak-load difference between scenarios grows from zero in 2016 to 13 GW in 2025.

ity growing faster than in base scenario. Growth scenario will have greater peak load, but also greater installed wind generation capacity. Higher demand, other things being equal, naturally increases electricity price as there is less surplus capacity and power plants with higher offer prices need to be dispatched. On the other hand, higher wind generation can be expected to decrease prices. These two differences between base and growth scenario can be expected to cancel impact of each other on RT price to some extent. However, there may be some differences in simulation results, e.g. as a result of more frequent and larger net-load ramps caused by intermittent nature of wind generation.

4.5.2 Comparison of results in different scenarios

100 simulations were run of both scenarios and three example RT price time series of both scenarios for years 2016-2025 are shown in figure 4.15. They resemble each other visually, but high prices are slightly more frequent in

growth scenario RT price time series.

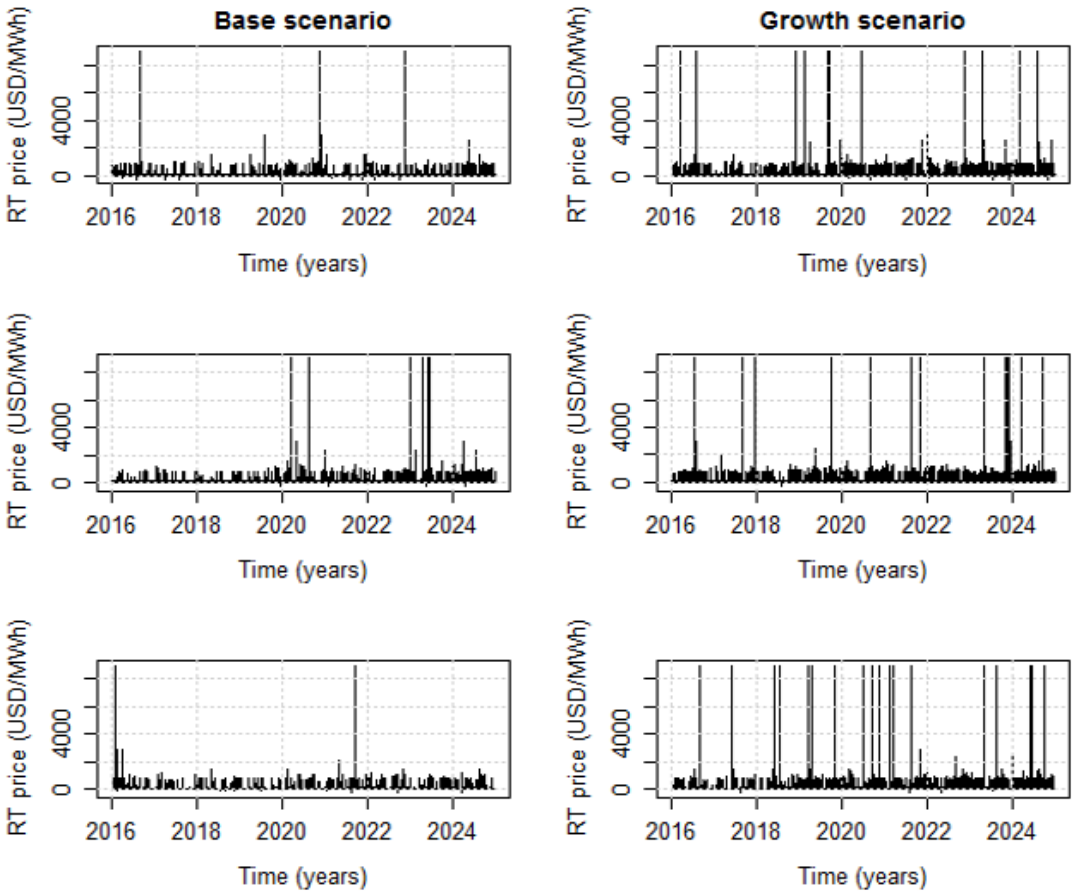

Figure 4.15: Three simulated RT price series of base scenario (left) are mostly similar to series of growth scenario (right). Growth scenario time series have a few more 9000 USD/MWh cap prices.

Histogram of median prices of the 100 simulation runs for each scenario are shown in figure 4.16. Median prices of growth scenario are consistently a little higher than median prices of base scenario simulation runs. However, difference is only a few cents per MWh. Duration curves of RT prices of all 200 simulation runs are shown in figure 4.17. There are no large differences in graphs. Figure 4.18 shows only 200 most expensive hours of year and only prices below 1000 USD/MWh. The dashed blue lines representing duration curves of growth scenario are clearly higher than solid black lines of base scenario. For example, RT price in base scenario simulations is only 40-70 hours

above 400 USD/MWh, whereas RT price in growth scenario simulations is 100-160 hours above 400 USD/MWh.

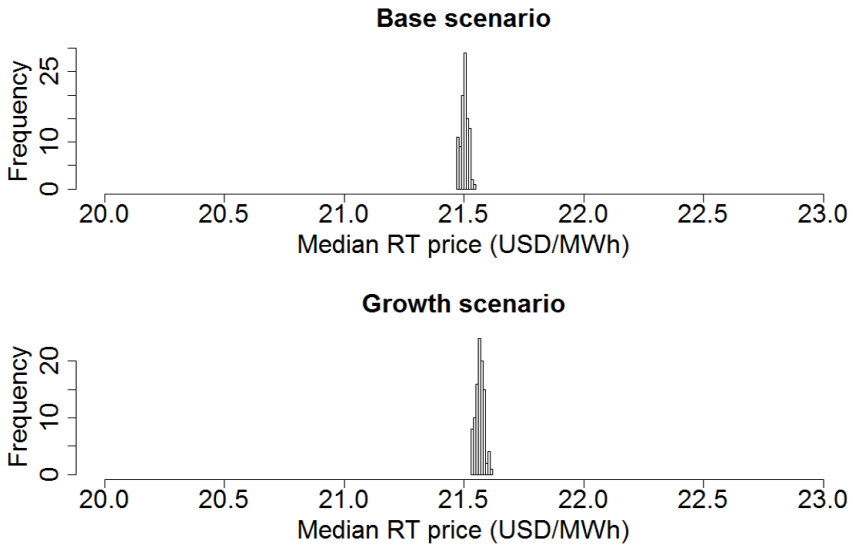

Figure 4.16: Median RT prices of 100 simulation runs of both scenarios show slight difference between scenarios. Median prices of growth scenario are generally a little higher (21.53-21.63 USD/MWh) than median prices of base simulation runs (21.47-21.56 USD/MWh). Difference in median prices is so small that it has practically no impact on market participants.

The differences in duration curves suggest that the impact of wind generation capacity increasing faster in growth scenario creates more price spikes in RT market. Increased number of high-price hours without increase in average price means that price volatility is greater in growth scenario than in base scenario.

Systematic differences in the simulated RT price time series of different scenarios result from differences in demand and wind generation capacity. We can conclude that increasing wind generation capacity share of all generation capacity increases volatility of RT price in ERCOT. This increase in volatility may give opportunities for flexible power plants to sell energy in RT market and benefit from transactions between DA and RT markets.

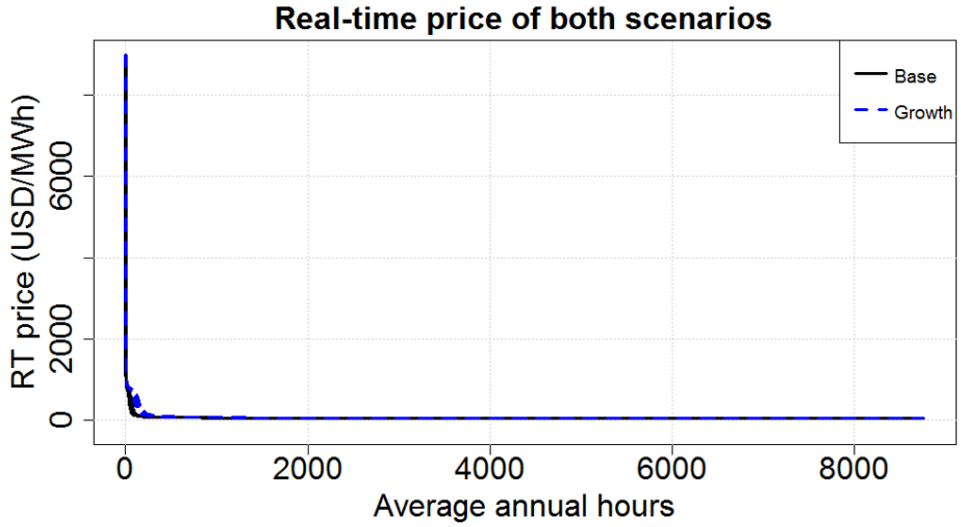

Figure 4.17: 200 duration curves, 100 of each scenario, show hardly any variation. The graphs are almost similar, except in the highest-price hours (left end).

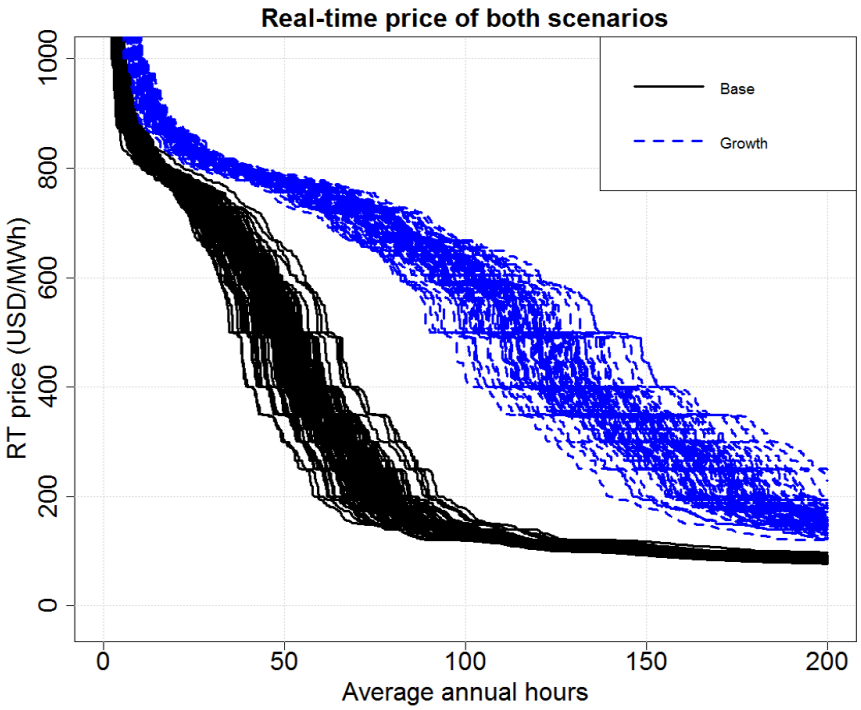

Figure 4.18: Duration curves of 100 simulation runs of base scenario and growth scenario. Blue dashed lines are clearly higher than black solid lines, indicating that high prices are more frequent in growth scenario.

Chapter 5

Conclusions

The purpose of this work was to understand real-time (RT) price creation in ERCOT market and build a simulation methodology that can be used to create simulated RT price series for several future years. We conducted statistical analyses to identify price drivers and constructed a simulation methodology using bootstrap method with several modifications.

We began with a short introduction to electricity markets and took a closer look at ERCOT market. Next, we examined the impact of different factors on RT price and selected explanatory variables to be used as inputs of our forecasting model based on results of statistical analysis. Future values of the chosen explanatory variables can be simulated by conducting a day-ahead (DA) market simulation using a dedicated commercial software. We introduced a stochastic model to create simulated RT price time series. We validated its functioning by using historical inputs and comparing output RT price series to historical RT price. Moreover, we carried out future simulations to compare RT price in two future scenarios, (i) base scenario in which installed wind generation capacity is stable and peak-load develops in future as predicted by ERCOT and (ii) growth scenario in which installed wind generation capacity and load grow 8 % and 3 % per year, respectively. By comparing results of the two simulations we saw that increasing share of wind generation capacity increases RT price volatility in ERCOT market.

The statistical analysis of RT price and its potential drivers showed us that occurrence of price spikes can be largely explained by explanatory variables surplus capacity and DA price. We found that probability of price spike is immensely higher when surplus capacity is below the limit of 7 GW than when it is above the limit. On normal (non-spike) RT price level, forecast error of surplus capacity, change speed of net-load, DA price, and previous behaviour of RT price explain RT price movements.

We chose to build the forecasting method using bootstrap method since

it makes no assumptions of specific distributions or the stochastic processes that generate prices. There is also a strong theoretical ground for forecasting using bootstrap method and it has been used in many applications. We made several modifications to account for impact of explanatory variables. New explanatory variables can be easily added to the model, if identified. We divided historical RT prices to several sampling groups. Conditions of each future time-step determine probabilities used in stochastic selection of sampling groups.

5.1 Topics for future research

The model introduced here could be modified in several ways. Simulation of actual demand could be done taking into account different user groups and modelling their demands separately, as ERCOT does. Autocorrelation of simulated actual wind generation and other explanatory variables could be made stronger. One possibility would be using block bootstrap method, instead of Markov chain and simple bootstrap of actual value as a percentage, in a similar way that we simulate RT price within spike periods.

Impact of solar generation could be added to the forecasting model as one explanatory variable that varies through the day stochastically, taking into account the impact of season and hour of day. Also, differences between node and hub prices could be studied thoroughly to see impact of congestion on electricity price.

Accuracy of forecast RT price conditional on explanatory variable values could be improved by increasing the number of different dynamic time step categories by adding division points to explanatory variables. However, there is limit to accuracy improvement using this method, since sampling groups simultaneously decrease in size, reducing important randomness in simulation. Highest accuracy could be achieved by building a fundamental forecasting model based on explanatory variables predicted accurately, but it would be very difficult to construct. RT price is much more volatile than DA price and often exceeds significantly marginal generation costs of power plants, that form the basis for DA market simulations. Also, predicting explanatory variables precisely is impossible task.

Methodology developed in this work could also be applied to an other market with some modifications. It would only make sense in a market where functioning RT market exists, of course. It is likely that price drivers would be different in some other market, and same parameter values as here could certainly not be used. Steps needed to create a stochastic RT price forecasting method are:

1. Interview experts to understand RT price creation in the chosen market. Hypothesize how RT price could be modelled and what are the most important drivers of RT price.

2. Conduct statistical analysis to decide whether RT price or spread between RT and DA prices has to be simulated. Possibly both are needed, but in different situations depending on DA price as we did in this study. Decision depends on which variable has more uniform distribution.

3. Conduct statistical analysis (scatter plots, empirical autocorrelation function, empirical probabilities of price groups conditional on values of explanatory variables) to identify price drivers

4. Conduct statistical analysis (scatter plots, conditional histograms of RT price) to select class limits for identified price drivers and needed lags of dependent variable as explanatory variable.

5. Build simulation methodology based on historical and future time series of price drivers in the same way as we did in this study.

6. Validate model functioning by comparing simulated and historical RT price time series, when historical inputs are used for simulation.

7. Simulate future inputs using DA market modelling software and stochastic modelling.

8. Forecast future RT price using possibly more than one scenario of development of explanatory variables in future. Compare results to draw conclusions of future RT price volatility in the chosen market.

Bibliography

[1] Bidding areas. http://www.nordpoolspot.com/How-does-it-work/Bidding-areas/, (accessed April 3, 2016).

[2] Breiman, L. Random forests. *Machine learning 45*, 1 (2001), 5–32.

[3] Carr, B. Day ahead LMP forecasting with Aurora OPF, 2008.

[4] Cruz, A., Munoz, A., Zamora, J. L., and Espínola, R. The effect of wind generation and weekday on spanish electricity spot price forecasting. *Electric Power Systems Research 81*, 10 (2011), 1924–1935.

[5] Davison, A. C., and Hinkley, D. V. *Bootstrap methods and their application*, vol. 1. Cambridge university press, 1997.

[6] Dumas, J. Operating reserve demand curve, 2014.

[7] Efron, B. Bootstrap methods: another look at the jackknife. *The annals of Statistics* (1979), 1–26.

[8] Efron, B., and Tibshirani, R. J. *An introduction to the bootstrap*. CRC press, 1994.

[9] ERCOT. 2014 state of the grid report. http://www.ercot.com/content/news/presentations/2015/2014%20State_of_the_Grid_Web_21015.pdf, 2014 (accessed February 3, 2016).

[10] ERCOT. Business practice manual. http://www.ercot.com/mktrules/bpm, 2015 (accessed April 3, 2016).

[11] ERCOT. ERCOT nodal 101. http://www.ercot.com/services/training/course/14#materials, 2015 (accessed February 3, 2016).

[12] ERCOT. ERCOT quick facts. http://www.ercot.com/content/news/presentations/2015/ERCOT_Quick_Facts_122115.pdf, 2015 (accessed February 3, 2016).

[13] ERCOT. ORDC workshop, ERCOT market training. `http://www.ercot.com/content/wcm/training_courses/107/ordc_workshop.pdf`, 2015 (accessed February 3, 2016).

[14] ERCOT. Nodal protocols. `http://www.ercot.com/mktrules/nprotocols/current`, 2016 (accessed April 3, 2016).

[15] ERCOT. Texas nodal market guide. `www.ercot.com/content/services/rq/ERCOT_Nodal_Market_Guide_v3.0.doc`, 2016 (accessed April 3, 2016).

[16] Garrison, J. B. *A grid-level unit commitment assessment of high wind penetration and utilization of compressed air energy storage in ERCOT*. PhD thesis, 2014.

[17] Hodge, B.-M., Florita, A., Sharp, J., Margulis, M., and Mcreavy, D. The value of improved short-term wind power forecasting.

[18] Huang, S.-H., Dumas, J., González-Pérez, C., and Lee, W.-J. Grid security through load reduction in the ERCOT market. *Industry Applications, IEEE Transactions on 45*, 2 (2009), 555–559.

[19] IEA. The power of transformation - wind, sun and the economics of flexible power systems.

[20] Javanainen, T. Analysis of short-term hydro power production in the Nordic electricity.

[21] Klimstra, J. *Power supply challenges*. Wärtsilä Finland Oy, 2014.

[22] Klimstra, J., and Hotakainen, M. *Smart power generation*. Avain, 2011.

[23] Koopman, S. J., Ooms, M., and Carnero, M. A. Periodic seasonal Reg-ARFIMA–GARCH models for daily electricity spot prices. *Journal of the American Statistical Association 102*, 477 (2007), 16–27.

[24] Laboratory, N. R. E. Solar power and the electric grid, energy analysis.

[25] Leino, J., Rautkivi, M., Heikkinen, J., and Hultholm, C. Enabling major national savings through a new approach to system security, 2013.

[26] McBride, D. G., Dietz, M. J., Vennemeyer, M. T., Meadors, S. A., Benfer, R. A., and Furbee, N. L. Bootstrap methods for sex determination from the os coxae using the id3 algorithm. *Journal of forensic sciences 46*, 3 (2001), 427–431.

[27] Newell, S., Spees, K., Pfeifenberger, J., Mudge, R., DeLucia, M., and Carlton, R. ERCOT investment incentives and resource adequacy. *The Brattle Group, prepared for the Electric Reliability Council of Texas* (2012).

[28] Pascual, L., Romo, J., and Ruiz, E. Bootstrap prediction for returns and volatilities in garch models. *Computational Statistics & Data Analysis 50*, 9 (2006), 2293–2312.

[29] Potomac Economics, L. 2014 state of the market report for the ERCOT wholesale electricity markets.

[30] Rautkivi, M. Case ERCOT. `http://www.smartpowergeneration.com/ercot`, 2015 (accessed April 3, 2016).

[31] Savolainen, A., et al. The role of nuclear and other conventional power plants in the flexible energy system.

[32] Vehviläinen, I., and Pyykkönen, T. Stochastic factor model for electricity spot price—the case of the Nordic market. *Energy Economics 27*, 2 (2005), 351–367.

[33] Virasjoki, V., et al. Market impacts of storage in a transmission-constrained power system.

[34] Zareipour, H. Short-term electricity market prices: A review of characteristics and forecasting methods. In *Handbook of Networks in Power Systems I*. Springer, 2012, pp. 89–121.

Chapter 6

Dynamic time step categories

Table 6.1: Values of explanatory variables and category of previous time step (delta class, price class or spike) of each dynamic time step category $C \in 0, 1, 2, ..., 300$. $C = 300$ corresponds to price spike.

C	$X_{spike}(t)$	$\Delta_C(t-1)$	$RT_C(t-1)$	$X_{spike}(t-1)$	DA	$FCFE$	$NLCA$
0	0	0		0	$DA < 25$	< -1800	< 200
1	0	0		0	$DA < 25$	< -1800	≥ 200
2	0	0		0	$DA < 25$	≥ -1800	< 200
3	0	0		0	$DA < 25$	≥ -1800	≥ 200
4	0	0		0	$25 \leq DA < 40$	< -1800	< 200
5	0	0		0	$25 \leq DA < 40$	< -1800	≥ 200
6	0	0		0	$25 \leq DA < 40$	≥ -1800	< 200
7	0	0		0	$25 \leq DA < 40$	≥ -1800	≥ 200
8	0	0		0	$40 \leq DA < 300$	< -1800	< 200
9	0	0		0	$40 \leq DA < 300$	< -1800	≥ 200
10	0	0		0	$40 \leq DA < 300$	≥ -1800	< 200
11	0	0		0	$40 \leq DA < 300$	≥ -1800	≥ 200
12	0	1		0	$DA < 25$	< -1800	< 200
13	0	1		0	$DA < 25$	< -1800	≥ 200
14	0	1		0	$DA < 25$	≥ -1800	< 200
15	0	1		0	$DA < 25$	≥ -1800	≥ 200
16	0	1		0	$25 \leq DA < 40$	< -1800	< 200
17	0	1		0	$25 \leq DA < 40$	< -1800	≥ 200
18	0	1		0	$25 \leq DA < 40$	≥ -1800	< 200
19	0	1		0	$25 \leq DA < 40$	≥ -1800	≥ 200
20	0	1		0	$40 \leq DA < 300$	< -1800	< 200
21	0	1		0	$40 \leq DA < 300$	< -1800	≥ 200
22	0	1		0	$40 \leq DA < 300$	≥ -1800	< 200
23	0	1		0	$40 \leq DA < 300$	≥ -1800	≥ 200
24	0	2		0	$DA < 25$	< -1800	< 200
25	0	2		0	$DA < 25$	< -1800	≥ 200

Continued on next page

Table 6.1 – continued from previous page

C	$X_{spike}(t)$	$\Delta_C(t-1)$	$RT_C(t-1)$	$X_{spike}(t-1)$	DA	$FCFE$	$NLCA$
26	0	2		0	$DA < 25$	≥ -1800	< 200
27	0	2		0	$DA < 25$	≥ -1800	≥ 200
28	0	2		0	$25 \leq DA < 40$	< -1800	< 200
29	0	2		0	$25 \leq DA < 40$	< -1800	≥ 200
30	0	2		0	$25 \leq DA < 40$	≥ -1800	< 200
31	0	2		0	$25 \leq DA < 40$	≥ -1800	≥ 200
32	0	2		0	$40 \leq DA < 300$	< -1800	< 200
33	0	2		0	$40 \leq DA < 300$	< -1800	≥ 200
34	0	2		0	$40 \leq DA < 300$	≥ -1800	< 200
35	0	2		0	$40 \leq DA < 300$	≥ -1800	≥ 200
36	0	3		0	$DA < 25$	< -1800	< 200
37	0	3		0	$DA < 25$	< -1800	≥ 200
38	0	3		0	$DA < 25$	≥ -1800	< 200
39	0	3		0	$DA < 25$	≥ -1800	≥ 200
40	0	3		0	$25 \leq DA < 40$	< -1800	< 200
41	0	3		0	$25 \leq DA < 40$	< -1800	≥ 200
42	0	3		0	$25 \leq DA < 40$	≥ -1800	< 200
43	0	3		0	$25 \leq DA < 40$	≥ -1800	≥ 200
44	0	3		0	$40 \leq DA < 300$	< -1800	< 200
45	0	3		0	$40 \leq DA < 300$	< -1800	≥ 200
46	0	3		0	$40 \leq DA < 300$	≥ -1800	< 200
47	0	3		0	$40 \leq DA < 300$	≥ -1800	≥ 200
48	0	4		0	$DA < 25$	< -1800	< 200
49	0	4		0	$DA < 25$	< -1800	≥ 200
50	0	4		0	$DA < 25$	≥ -1800	< 200
51	0	4		0	$DA < 25$	≥ -1800	≥ 200
52	0	4		0	$25 \leq DA < 40$	< -1800	< 200
53	0	4		0	$25 \leq DA < 40$	< -1800	≥ 200
54	0	4		0	$25 \leq DA < 40$	≥ -1800	< 200
55	0	4		0	$25 \leq DA < 40$	≥ -1800	≥ 200
56	0	4		0	$40 \leq DA < 300$	< -1800	< 200
57	0	4		0	$40 \leq DA < 300$	< -1800	≥ 200
58	0	4		0	$40 \leq DA < 300$	≥ -1800	< 200
59	0	4		0	$40 \leq DA < 300$	≥ -1800	≥ 200
60	0	5		0	$DA < 25$	< -1800	< 200
61	0	5		0	$DA < 25$	< -1800	≥ 200
62	0	5		0	$DA < 25$	≥ -1800	< 200
63	0	5		0	$DA < 25$	≥ -1800	≥ 200
64	0	5		0	$25 \leq DA < 40$	< -1800	< 200
65	0	5		0	$25 \leq DA < 40$	< -1800	≥ 200
66	0	5		0	$25 \leq DA < 40$	≥ -1800	< 200
67	0	5		0	$25 \leq DA < 40$	≥ -1800	≥ 200
68	0	5		0	$40 \leq DA < 300$	< -1800	< 200
69	0	5		0	$40 \leq DA < 300$	< -1800	≥ 200
70	0	5		0	$40 \leq DA < 300$	≥ -1800	< 200
71	0	5		0	$40 \leq DA < 300$	≥ -1800	≥ 200

Continued on next page

Table 6.1 – continued from previous page

C	$X_{spike}(t)$	$\Delta_C(t-1)$	$RT_C(t-1)$	$X_{spike}(t-1)$	DA	$FCFE$	$NLCA$
72	0	6		0	$DA < 25$	< -1800	< 200
73	0	6		0	$DA < 25$	< -1800	≥ 200
74	0	6		0	$DA < 25$	≥ -1800	< 200
75	0	6		0	$DA < 25$	≥ -1800	≥ 200
76	0	6		0	$25 \leq DA < 40$	< -1800	< 200
77	0	6		0	$25 \leq DA < 40$	< -1800	≥ 200
78	0	6		0	$25 \leq DA < 40$	≥ -1800	< 200
79	0	6		0	$25 \leq DA < 40$	≥ -1800	≥ 200
80	0	6		0	$40 \leq DA < 300$	< -1800	< 200
81	0	6		0	$40 \leq DA < 300$	< -1800	≥ 200
82	0	6		0	$40 \leq DA < 300$	≥ -1800	< 200
83	0	6		0	$40 \leq DA < 300$	≥ -1800	≥ 200
84	0	7		0	$DA < 25$	< -1800	< 200
85	0	7		0	$DA < 25$	< -1800	≥ 200
86	0	7		0	$DA < 25$	≥ -1800	< 200
87	0	7		0	$DA < 25$	≥ -1800	≥ 200
88	0	7		0	$25 \leq DA < 40$	< -1800	< 200
89	0	7		0	$25 \leq DA < 40$	< -1800	≥ 200
90	0	7		0	$25 \leq DA < 40$	≥ -1800	< 200
91	0	7		0	$25 \leq DA < 40$	≥ -1800	≥ 200
92	0	7		0	$40 \leq DA < 300$	< -1800	< 200
93	0	7		0	$40 \leq DA < 300$	< -1800	≥ 200
94	0	7		0	$40 \leq DA < 300$	≥ -1800	< 200
95	0	7		0	$40 \leq DA < 300$	≥ -1800	≥ 200
96	0	8		0	$DA < 25$	< -1800	< 200
97	0	8		0	$DA < 25$	< -1800	≥ 200
98	0	8		0	$DA < 25$	≥ -1800	< 200
99	0	8		0	$DA < 25$	≥ -1800	≥ 200
100	0	8		0	$25 \leq DA < 40$	< -1800	< 200
101	0	8		0	$25 \leq DA < 40$	< -1800	≥ 200
102	0	8		0	$25 \leq DA < 40$	≥ -1800	< 200
103	0	8		0	$25 \leq DA < 40$	≥ -1800	≥ 200
104	0	8		0	$40 \leq DA < 300$	< -1800	< 200
105	0	8		0	$40 \leq DA < 300$	< -1800	≥ 200
106	0	8		0	$40 \leq DA < 300$	≥ -1800	< 200
107	0	8		0	$40 \leq DA < 300$	≥ -1800	≥ 200
108	0	9		0	$DA < 25$	< -1800	< 200
109	0	9		0	$DA < 25$	< -1800	≥ 200
110	0	9		0	$DA < 25$	≥ -1800	< 200
111	0	9		0	$DA < 25$	≥ -1800	≥ 200
112	0	9		0	$25 \leq DA < 40$	< -1800	< 200
113	0	9		0	$25 \leq DA < 40$	< -1800	≥ 200
114	0	9		0	$25 \leq DA < 40$	≥ -1800	< 200
115	0	9		0	$25 \leq DA < 40$	≥ -1800	≥ 200
116	0	9		0	$40 \leq DA < 300$	< -1800	< 200
117	0	9		0	$40 \leq DA < 300$	< -1800	≥ 200

Continued on next page

Table 6.1 – continued from previous page

C	$X_{spike}(t)$	$\Delta_C(t-1)$	$RT_C(t-1)$	$X_{spike}(t-1)$	DA	$FCFE$	$NLCA$
118	0	9		0	$40 \leq DA < 300$	≥ -1800	< 200
119	0	9		0	$40 \leq DA < 300$	≥ -1800	≥ 200
120	0	10		0	$DA < 25$	< -1800	< 200
121	0	10		0	$DA < 25$	< -1800	≥ 200
122	0	10		0	$DA < 25$	≥ -1800	< 200
123	0	10		0	$DA < 25$	≥ -1800	≥ 200
124	0	10		0	$25 \leq DA < 40$	< -1800	< 200
125	0	10		0	$25 \leq DA < 40$	< -1800	≥ 200
126	0	10		0	$25 \leq DA < 40$	≥ -1800	< 200
127	0	10		0	$25 \leq DA < 40$	≥ -1800	≥ 200
128	0	10		0	$40 \leq DA < 300$	< -1800	< 200
129	0	10		0	$40 \leq DA < 300$	< -1800	≥ 200
130	0	10		0	$40 \leq DA < 300$	≥ -1800	< 200
131	0	10		0	$40 \leq DA < 300$	≥ -1800	≥ 200
132	0	11		0	$DA < 25$	< -1800	< 200
133	0	11		0	$DA < 25$	< -1800	≥ 200
134	0	11		0	$DA < 25$	≥ -1800	< 200
135	0	11		0	$DA < 25$	≥ -1800	≥ 200
136	0	11		0	$25 \leq DA < 40$	< -1800	< 200
137	0	11		0	$25 \leq DA < 40$	< -1800	≥ 200
138	0	11		0	$25 \leq DA < 40$	≥ -1800	< 200
139	0	11		0	$25 \leq DA < 40$	≥ -1800	≥ 200
140	0	11		0	$40 \leq DA < 300$	< -1800	< 200
141	0	11		0	$40 \leq DA < 300$	< -1800	≥ 200
142	0	11		0	$40 \leq DA < 300$	≥ -1800	< 200
143	0	11		0	$40 \leq DA < 300$	≥ -1800	≥ 200
144	0		0	0	$DA < 25$	< -1800	< 200
145	0		0	0	$DA < 25$	< -1800	≥ 200
146	0		0	0	$DA < 25$	≥ -1800	< 200
147	0		0	0	$DA < 25$	≥ -1800	≥ 200
148	0		0	0	$25 \leq DA < 40$	< -1800	< 200
149	0		0	0	$25 \leq DA < 40$	< -1800	≥ 200
150	0		0	0	$25 \leq DA < 40$	≥ -1800	< 200
151	0		0	0	$25 \leq DA < 40$	≥ -1800	≥ 200
152	0		0	0	$40 \leq DA < 300$	< -1800	< 200
153	0		0	0	$40 \leq DA < 300$	< -1800	≥ 200
154	0		0	0	$40 \leq DA < 300$	≥ -1800	< 200
155	0		0	0	$40 \leq DA < 300$	≥ -1800	≥ 200
156	0		1	0	$DA < 25$	< -1800	< 200
157	0		1	0	$DA < 25$	< -1800	≥ 200
158	0		1	0	$DA < 25$	≥ -1800	< 200
159	0		1	0	$DA < 25$	≥ -1800	≥ 200
160	0		1	0	$25 \leq DA < 40$	< -1800	< 200
161	0		1	0	$25 \leq DA < 40$	< -1800	≥ 200
162	0		1	0	$25 \leq DA < 40$	≥ -1800	< 200
163	0		1	0	$25 \leq DA < 40$	≥ -1800	≥ 200

Continued on next page

Table 6.1 – continued from previous page

C	$X_{spike}(t)$	$\Delta_C(t-1)$	$RT_C(t-1)$	$X_{spike}(t-1)$	DA	$FCFE$	$NLCA$
164	0		1	0	$40 \leq DA < 300$	< -1800	< 200
165	0		1	0	$40 \leq DA < 300$	< -1800	≥ 200
166	0		1	0	$40 \leq DA < 300$	≥ -1800	< 200
167	0		1	0	$40 \leq DA < 300$	≥ -1800	≥ 200
168	0		2	0	$DA < 25$	< -1800	< 200
169	0		2	0	$DA < 25$	< -1800	≥ 200
170	0		2	0	$DA < 25$	≥ -1800	< 200
171	0		2	0	$DA < 25$	≥ -1800	≥ 200
172	0		2	0	$25 \leq DA < 40$	< -1800	< 200
173	0		2	0	$25 \leq DA < 40$	< -1800	≥ 200
174	0		2	0	$25 \leq DA < 40$	≥ -1800	< 200
175	0		2	0	$25 \leq DA < 40$	≥ -1800	≥ 200
176	0		2	0	$40 \leq DA < 300$	< -1800	< 200
177	0		2	0	$40 \leq DA < 300$	< -1800	≥ 200
178	0		2	0	$40 \leq DA < 300$	≥ -1800	< 200
179	0		2	0	$40 \leq DA < 300$	≥ -1800	≥ 200
180	0		3	0	$DA < 25$	< -1800	< 200
181	0		3	0	$DA < 25$	< -1800	≥ 200
182	0		3	0	$DA < 25$	≥ -1800	< 200
183	0		3	0	$DA < 25$	≥ -1800	≥ 200
184	0		3	0	$25 \leq DA < 40$	< -1800	< 200
185	0		3	0	$25 \leq DA < 40$	< -1800	≥ 200
186	0		3	0	$25 \leq DA < 40$	≥ -1800	< 200
187	0		3	0	$25 \leq DA < 40$	≥ -1800	≥ 200
188	0		3	0	$40 \leq DA < 300$	< -1800	< 200
189	0		3	0	$40 \leq DA < 300$	< -1800	≥ 200
190	0		3	0	$40 \leq DA < 300$	≥ -1800	< 200
191	0		3	0	$40 \leq DA < 300$	≥ -1800	≥ 200
192	0		4	0	$DA < 25$	< -1800	< 200
193	0		4	0	$DA < 25$	< -1800	≥ 200
194	0		4	0	$DA < 25$	≥ -1800	< 200
195	0		4	0	$DA < 25$	≥ -1800	≥ 200
196	0		4	0	$25 \leq DA < 40$	< -1800	< 200
197	0		4	0	$25 \leq DA < 40$	< -1800	≥ 200
198	0		4	0	$25 \leq DA < 40$	≥ -1800	< 200
199	0		4	0	$25 \leq DA < 40$	≥ -1800	≥ 200
200	0		4	0	$40 \leq DA < 300$	< -1800	< 200
201	0		4	0	$40 \leq DA < 300$	< -1800	≥ 200
202	0		4	0	$40 \leq DA < 300$	≥ -1800	< 200
203	0		4	0	$40 \leq DA < 300$	≥ -1800	≥ 200
204	0		5	0	$DA < 25$	< -1800	< 200
205	0		5	0	$DA < 25$	< -1800	≥ 200
206	0		5	0	$DA < 25$	≥ -1800	< 200
207	0		5	0	$DA < 25$	≥ -1800	≥ 200
208	0		5	0	$25 \leq DA < 40$	< -1800	< 200
209	0		5	0	$25 \leq DA < 40$	< -1800	≥ 200

Continued on next page

Table 6.1 – continued from previous page

C	$X_{spike}(t)$	$\Delta_C(t-1)$	$RT_C(t-1)$	$X_{spike}(t-1)$	DA	$FCFE$	$NLCA$
210	0		5	0	$25 \leq DA < 40$	≥ -1800	< 200
211	0		5	0	$25 \leq DA < 40$	≥ -1800	≥ 200
212	0		5	0	$40 \leq DA < 300$	< -1800	< 200
213	0		5	0	$40 \leq DA < 300$	< -1800	≥ 200
214	0		5	0	$40 \leq DA < 300$	≥ -1800	< 200
215	0		5	0	$40 \leq DA < 300$	≥ -1800	≥ 200
216	0		6	0	$DA < 25$	< -1800	< 200
217	0		6	0	$DA < 25$	< -1800	≥ 200
218	0		6	0	$DA < 25$	≥ -1800	< 200
219	0		6	0	$DA < 25$	≥ -1800	≥ 200
220	0		6	0	$25 \leq DA < 40$	< -1800	< 200
221	0		6	0	$25 \leq DA < 40$	< -1800	≥ 200
222	0		6	0	$25 \leq DA < 40$	≥ -1800	< 200
223	0		6	0	$25 \leq DA < 40$	≥ -1800	≥ 200
224	0		6	0	$40 \leq DA < 300$	< -1800	< 200
225	0		6	0	$40 \leq DA < 300$	< -1800	≥ 200
226	0		6	0	$40 \leq DA < 300$	≥ -1800	< 200
227	0		6	0	$40 \leq DA < 300$	≥ -1800	≥ 200
228	0		7	0	$DA < 25$	< -1800	< 200
229	0		7	0	$DA < 25$	< -1800	≥ 200
230	0		7	0	$DA < 25$	≥ -1800	< 200
231	0		7	0	$DA < 25$	≥ -1800	≥ 200
232	0		7	0	$25 \leq DA < 40$	< -1800	< 200
233	0		7	0	$25 \leq DA < 40$	< -1800	≥ 200
234	0		7	0	$25 \leq DA < 40$	≥ -1800	< 200
235	0		7	0	$25 \leq DA < 40$	≥ -1800	≥ 200
236	0		7	0	$40 \leq DA < 300$	< -1800	< 200
237	0		7	0	$40 \leq DA < 300$	< -1800	≥ 200
238	0		7	0	$40 \leq DA < 300$	≥ -1800	< 200
239	0		7	0	$40 \leq DA < 300$	≥ -1800	≥ 200
240	0		8	0	$DA < 25$	< -1800	< 200
241	0		8	0	$DA < 25$	< -1800	≥ 200
242	0		8	0	$DA < 25$	≥ -1800	< 200
243	0		8	0	$DA < 25$	≥ -1800	≥ 200
244	0		8	0	$25 \leq DA < 40$	< -1800	< 200
245	0		8	0	$25 \leq DA < 40$	< -1800	≥ 200
246	0		8	0	$25 \leq DA < 40$	≥ -1800	< 200
247	0		8	0	$25 \leq DA < 40$	≥ -1800	≥ 200
248	0		8	0	$40 \leq DA < 300$	< -1800	< 200
249	0		8	0	$40 \leq DA < 300$	< -1800	≥ 200
250	0		8	0	$40 \leq DA < 300$	≥ -1800	< 200
251	0		8	0	$40 \leq DA < 300$	≥ -1800	≥ 200
252	0		9	0	$DA < 25$	< -1800	< 200
253	0		9	0	$DA < 25$	< -1800	≥ 200
254	0		9	0	$DA < 25$	≥ -1800	< 200
255	0		9	0	$DA < 25$	≥ -1800	≥ 200

Continued on next page

Table 6.1 – continued from previous page

C	$X_{spike}(t)$	$\Delta_C(t-1)$	$RT_C(t-1)$	$X_{spike}(t-1)$	DA	$FCFE$	$NLCA$
256	0		9	0	$25 \leq DA < 40$	< -1800	< 200
257	0		9	0	$25 \leq DA < 40$	< -1800	≥ 200
258	0		9	0	$25 \leq DA < 40$	≥ -1800	< 200
259	0		9	0	$25 \leq DA < 40$	≥ -1800	≥ 200
260	0		9	0	$40 \leq DA < 300$	< -1800	< 200
261	0		9	0	$40 \leq DA < 300$	< -1800	≥ 200
262	0		9	0	$40 \leq DA < 300$	≥ -1800	< 200
263	0		9	0	$40 \leq DA < 300$	≥ -1800	≥ 200
264	0		10	0	$DA < 25$	< -1800	< 200
265	0		10	0	$DA < 25$	< -1800	≥ 200
266	0		10	0	$DA < 25$	≥ -1800	< 200
267	0		10	0	$DA < 25$	≥ -1800	≥ 200
268	0		10	0	$25 \leq DA < 40$	< -1800	< 200
269	0		10	0	$25 \leq DA < 40$	< -1800	≥ 200
270	0		10	0	$25 \leq DA < 40$	≥ -1800	< 200
271	0		10	0	$25 \leq DA < 40$	≥ -1800	≥ 200
272	0		10	0	$40 \leq DA < 300$	< -1800	< 200
273	0		10	0	$40 \leq DA < 300$	< -1800	≥ 200
274	0		10	0	$40 \leq DA < 300$	≥ -1800	< 200
275	0		10	0	$40 \leq DA < 300$	≥ -1800	≥ 200
276	0		11	0	$DA < 25$	< -1800	< 200
277	0		11	0	$DA < 25$	< -1800	≥ 200
278	0		11	0	$DA < 25$	≥ -1800	< 200
279	0		11	0	$DA < 25$	≥ -1800	≥ 200
280	0		11	0	$25 \leq DA < 40$	< -1800	< 200
281	0		11	0	$25 \leq DA < 40$	< -1800	≥ 200
282	0		11	0	$25 \leq DA < 40$	≥ -1800	< 200
283	0		11	0	$25 \leq DA < 40$	≥ -1800	≥ 200
284	0		11	0	$40 \leq DA < 300$	< -1800	< 200
285	0		11	0	$40 \leq DA < 300$	< -1800	≥ 200
286	0		11	0	$40 \leq DA < 300$	≥ -1800	< 200
287	0		11	0	$40 \leq DA < 300$	≥ -1800	≥ 200
288	0			1	$DA < 25$	< -1800	< 200
289	0			1	$DA < 25$	< -1800	≥ 200
290	0			1	$DA < 25$	≥ -1800	< 200
291	0			1	$DA < 25$	≥ -1800	≥ 200
292	0			1	$25 \leq DA < 40$	< -1800	< 200
293	0			1	$25 \leq DA < 40$	< -1800	≥ 200
294	0			1	$25 \leq DA < 40$	≥ -1800	< 200
295	0			1	$25 \leq DA < 40$	≥ -1800	≥ 200
296	0			1	$40 \leq DA < 300$	< -1800	< 200
297	0			1	$40 \leq DA < 300$	< -1800	≥ 200
298	0			1	$40 \leq DA < 300$	≥ -1800	< 200
299	0			1	$40 \leq DA < 300$	≥ -1800	≥ 200
300	1						